Mathematical Biosciences and Engineering

Research article

Image quality assessment based on the perceived structural similarity index of an image

Juncai Yao[1,2,†,*], Jing Shen[1,†] and Congying Yao[1]

[1] School of Computer Engineering, Nanjing Institute of Technology, Nanjing 211167, China
[2] School of Information and Communication Engineering, Xi'an Jiaotong University, Xi'an 710049, China

† The authors contributed equally to this work.

* **Correspondence:** Email: yaojcnj@njit.edu.cn.

Abstract: Image quality assessment (IQA) has a very important role and wide applications in image acquisition, storage, transmission and processing. In designing IQA models, human visual system (HVS) characteristics introduced play an important role in improving their performances. In this paper, combining image distortion characteristics with HVS characteristics, based on the structure similarity index (SSIM) model, a novel IQA model based on the perceived structure similarity index (PSIM) of image is proposed. In the method, first, a perception model for HVS perceiving real images is proposed, combining the contrast sensitivity, frequency sensitivity, luminance nonlinearity and masking characteristics of HVS; then, in order to simulate HVS perceiving real image, the real images are processed with the proposed perception model, to eliminate their visual redundancy, thus, the perceived images are obtained; finally, based on the idea and modeling method of SSIM, combining with the features of perceived image, a novel IQA model, namely PSIM, is proposed. Further, in order to illustrate the performance of PSIM, 5335 distorted images with 41 distortion types in four image databases (TID2013, CSIQ, LIVE and CID) are used to simulate from three aspects: overall IQA of each database, IQA for each distortion type of images, and IQA for special distortion types of images. Further, according to the comprehensive benefit of precision, generalization performance and complexity, their IQA results are compared with those of 12 existing IQA models. The experimental results show that the accuracy (PLCC) of PSIM is 9.91% higher than that of SSIM in four databases, on average; and its performance is better than that of 12 existing IQA models. Synthesizing experimental results and theoretical analysis, it is showed that the proposed PSIM model is an effective and excellent IQA model.

1. Introduction

With the rapid development of information, communication and artificial intelligence technology, image technologies have been widely applied. And in order to achieve their own purposes, people often process images in various ways using these image technologies. However, in the process of image acquisition, storage, transmission, processing and reproduction, image distortion and quality degradation are inevitable, due to the imperfections of imaging system, processing method, transmission conditions and storage equipment. Thus, image damage directly affects user's subjective perception effect, use value and quality of service [1,2]. Hence, it is urgent to propose an effective and practical image quality assessment (IQA) method to measure quality of distorted images, so as to control image quality in the process of image processing, in order to better serve image processing, algorithm optimization, network control transmission, and so on [3]. However, so far, in IQA, although more achievements have been made and more IQA methods and models have been proposed, their application is still far from meeting the current growing requirements, due to the gap between the objective IQA scores and subjective perception results, to a certain extent [4,5]. Hence, it is very necessary to propose an effective and convenient IQA metric that is consistent with the subjective HVS perception as much as possible.

In recent years, much work has been done on IQA, and many models have been proposed. Among them, the typical ones are MSE (mean squared error), PSNR (peak signal to noise ratio), VSNR (visual single to noise ratio) [6], SSIM (structural similarity index) [7], FSIMc (feature similarity index) [8], GMSD (gradient magnitude similarity deviation) [9] and VSI (visual loyalty induced index) [10], etc. However, considering the comprehensive benefits of accuracy, generalization and computational complexity of these existing IQA models, SSIM proposed by Wang et al. [7] is generally recognized, and its comprehensive performance is better than that of other existing models. However, the accuracy of SSIM is still not high. Therefore, Wang et al. improved SSIM in 2011, and proposed a IQA model called as IWSSIM (SSIM based on information content weighting) [11]. However, the calculation of IWSSIM is complex, and cannot be applied in practice. So, in 2018, an improved SSIM model, namely CSF+SSIM, is proposed in reference [12]. In CSF+SSIM, the improved parts mainly include two aspects: the brightness in image is processed by perception nonlinear model of human visual system (HVS) characteristics, and the processed brightness is weighted by HVS contrast sensitivity (CS) threshold. However, this model does not consider eliminating the visual redundancy of image and does not improve form of SSIM. Recently, some IQA models have been proposed based on HVS characteristics. In 2022, aiming at the screen content images, by capturing distortions in the horizontal and vertical structures of image, and combining with the characteristic of HVS's trend to analyze a scene in a multi-scale fashion, a IQA model called as WS-HV [13] is proposed, which adopts wavelet transformation. However, its average accuracy PLCC (Pearson linear correlation coefficient) value in LIVE [14], TID2013 [15], and CSIQ [16] image databases can only reach 0.7454. In reference [17], based on the human judgment process, a IQA framework comprising three abstract perception layers (A3L) is proposed. Its IQA accuracy PLCC values in three databases [14−16] can get to 0.9640, 0.8820 and 0.8713, respectively. However, A3L [17] is relatively complex because it needs to deal with many

parameters in 3 layers. In reference [18], unifying structure and texture similarity, a full-reference IQA model with explicit tolerance to texture resampling is developed, called as deep image structure and texture similarity (DISTS) index. Their results show that DISTS [18] have a good applications in IQA on noised, blur, super-resolution, and compressed images, but is insensitive to mild local and global geometric distortions of images. In recent years, by combining machine learning and neural network, many IQA models have been proposed. In 2018, by means of introducing multiple pseudo reference images (MPRIs) by further degrading the distorted image in several ways and to certain degrees, and then comparing the similarities between the distorted image and the MPRIs, a bind MPRIs-based (BMPRI) [19] IQA measure is proposed. Their results show that BMPRI [19] are comparable for IQA of four mainstream natural scene image and screen content image, but its IQA generalization performance for different distortion type is poor. In 2022, a visual compensation restoration network (VCRNet)-based NR-IQA method is proposed [20], which uses a non-adversarial model to efficiently handle the distorted image restoration task. VCRNet [20] consists of a visual restoration network and a quality estimation network. It has a high accuracy, however, it is also very complex. In 2023, By systematically studying the interactions between channel-wise and spatial-wise attention, an adaptive spatial and channel attention merging transformer (ASCAM-Former) [21] is then proposed for IQA. ASCAM-Former [21] can yield accurate prediction on both authentically and synthetically distorted image quality datasets. However, zero-padding is employed to handle images with various resolutions, but it might result in sparse valid features when image resolution is too small. Therefore, in recent years, much work has been done on SSIM, and many improvement methods have been proposed. However, their comprehensive benefits are still unsatisfactory. Analyzing SSIM and these existing models, SSIM has some defects or deficiencies, which are mainly shown as follows:

1) For images with some distortion types, IQA effect of SSIM is poor. The IQA results in LIVE [14], TID2013 [15] and CSIQ [16] databases show that SSIM has a poor IQA effect on some types of distortion images, such as contract change (CC) and change of color saturation (CCS) in TID2013 database. Their IQA accuracy is low, whose PLCC values between the subjective IQA results and objective IQA scores are only 0.6067 and 0.4345, and SROCC (Spearman rank order correlation coefficient) values are only 0.3775 and 0.4141 [15].

2) De-visual redundancy is not fully considered in SSIM. SSIM is mainly based on the feature that HVS adapts well to the structural information in images, while it does not pay much attention to their visual redundancy [7,22]. Thus, there is a certain threshold for HVS perceiving image distortion. When the distortion of images is lower than the distortion perception threshold, HVS cannot detect that image has been distorted [22,23], and the subjective IQA scores MOS (mean opinion score) are basically unchanged. However, for SSIM, as long as image is distorted, the IQA scores using SSIM must change. These cause inconsistency between the subjective IQA results and objective IQA scores.

3) HVS perception characteristics is not combined insufficiently in SSIM. SSIM mainly evaluates the difference between two images from perspective of image structure. HVS characteristics combined in SSIM namely is image structure perception characteristics. So, the existing issues are as follows: (i) Except HVS structural cognitive characteristics, there are still many other HVS characteristics, such as the CS, masking, frequency sensitivity characteristics, etc. They can also be used in IQA [24]. (ii) IQA using SSIM is to aim at real image. However, real image is quite different from image perceived by HVS. In order to achieve higher consistency between the subjective and objective IQA results, it is necessary to eliminate the influence of real image on IQA accuracy. Hence, it is necessary to further introduce other HVS characteristics expect structural similarity, so as to improve the IQA accuracy.

4) The generalization performance of IQA model is often poor. From the IQA results of SSIM in LIVE [14], CSIQ [15] and TID2013 [16] databases, the IQA accuracy (PLCC and SROCC) can reach 0.9449 and 0.9479 in LIVE; while they are only 0.7895 and 0.7417 in TID2013, which show that the generalization performance of SSIM is poor.

5) The complexity of the IQA model is difficult to be solved effectively when designing IQA model. This is especially so when an IQA model is designed by combining with HVS characteristics, and it would be very complex because of the large amount of calculations [25]. However, it can greatly improve IQA accuracy for combining HVS characteristics and its mathematical model. Therefore, building IQA model combining with HVS characteristics has also been more studied, and many IQA models have been proposed too, such as VSNR, FSIMc, and VSI. Hence, the current difficulty focuses on how to combine HVS characteristics, while to make IQA model uncomplicated. Today, common solutions are as follows: First, the features causing distortion in image are extracted and analyzed, and then IQA model is designed according to HVS characteristics and weighting factors. Although the evaluation effect of these IQA methods is significant, it is highly targeted in terms of HVS characteristics. Moreover, to improve IQA accuracy, more image distortion features need to be extracted, which brought much calculations in IQA [26,27].

Based on the above analysis, comprehensively considering the accuracy, complexity and generalization performance of IQA model, combining with the CS, luminance perception and masking effect of HVS characteristics, based on the ideas of designing SSIM model, a novel IQA model based on the perceived structural similarity index (PSIM) is proposed. It will provide a theoretical and technical basis for IQA technology and application.

2. SSIM and HVS model

At present, SSIM has been used widely in IQA. Generally, the final receiver of the image is human. Hence, the performance goal of IQA model is to conform with HVS perception. Therefore, in the current IQA research, combining with HVS characteristics and its models is seen as an effective means to improve accuracy of IQA model. Hence, in this paper, based on SSIM, an IQA model by combining it with HVS characteristics is built.

2.1. SSIM model

SSIM is an index to measure similarity of two images. It is declared that image structure is an attribute that is independent of luminance and contrast, and can reflect object information in the image scene. In SSIM, the combination of luminance, contrast and structure is used as the visual information of image. Namely, the average value is expressed as the estimation of luminance, the standard deviation is expressed as the estimation of contrast, and covariance is used to measure the structural similarity. For two images (x and y, i.e., source image and distorted image), SSIM is calculated as Eq (1) [7].

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\delta_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\delta_x^2 + \delta_y^2 + c_2)}, \tag{1}$$

where μ_x and μ_y are the average values of image x and y, respectively. δ_x^2 and δ_y^2 are the variance of x and y, and δ_{xy} is the covariance of x and y. $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$; here, L is the dynamic range of pixel luminance; $k_1 = 0.01$, $k_2 = 0.03$; c_1 and c_2 are constants, whose range is between 0~1. When two

images are exactly the same, the value of SSIM is 1, and when it has no relationship, it is 0.

2.2. Contrast sensitivity characteristics model of HVS

The contrast sensitivity characteristics is one of four space characteristics of HVS. At present, its mathematical model is often described mainly by the contrast sensitivity function (CSF), which reflects mainly the functional relationship between the HVS sensitivity to targets and the angular frequencies of targets. Hence, CSF can better reflect the CS and frequency sensitivity characteristics of HVS detecting targets. Because CSF can be applied widely, much research about CSF has been carried out, and many CSF models have been proposed [28−30]. Considering the computational complexity of applying CSF in the IQA algorithm and the effectiveness of CSF model reflecting HVS characteristics, the luminance CSF (i.e., CSF_L) proposed by Barten et al. [29] and the chromaticity CSF (i.e., CSF_{rg} and CSF_{by}) proposed by Nadenau et al. [30] are used in the proposed algorithm, and they are as shown in Eqs (2)–(4).

1) CSF model of HVS perceiving luminance

$$CSF_L(f_\theta) = a \cdot f_\theta \cdot exp(-b \cdot f_\theta)[1 + c \cdot exp(b \cdot f_\theta)]^{0.5}, \tag{2}$$

here, $a = 540 \cdot (1 + 0.7/L)^{-0.2} \cdot \{1 + 12[w \cdot (1 + f_\theta/3)^2]^{-1}\}$, $b = 0.3 \cdot (1 + 100/L)^{0.15}$, $c = 0.06$, where, f_θ is the angular frequency, w is the spatial viewing angle and L is the average luminance of detected target [29].

2) CSF models of HVS perceiving opposite colors

CSF model of HVS perceiving the red-green opposite colors is written as CSF_{rg}, which is as follows:

$$CSF_{rg}(f_\theta) = a \cdot exp^{b(f_\theta)^c} \qquad a = 1, b = -0.152, c = 0.893, \tag{3}$$

CSF model of HVS perceiving the blue-yellow opposite colors is written as CSF_{by}, which is as follows:

$$CSF_{by}(f_\theta) = a \cdot exp^{b(f_\theta)^c} \qquad a = 1, b = -0.2041, c = 0.9, \tag{4}$$

where CSF_{rg} and CSF_{by}, respectively, represent chroma CS threshold of HVS perceiving red-green and blue-yellow opposite color targets [30].

The trend of HVS perceiving intensity change from small to large is a about exponential curve change form. The above brightness and chroma HVS models mainly adopt the exponential form under linear modulation, which can better reflect the HVS characteristics.

3. IQA method and its model

At present, SSIM is applied widely by using structural similarity to evaluate the quality of the image. However, its accuracy is often not high, and its generalization performance for partial distortion types is poor. Analyzing the reason, from the characteristics of SSIM itself, the main is that SSIM uses three features, namely luminance, contrast and similar structure, to evaluate image quality. This is a very good idea, though these features are features of real images and there are great differences between real images and images perceived by HVS. Hence, the performance of SSIM is slightly poor. For example, the luminance perceived by HVS is nonlinear, while luminance in real image is linear;

and SSIM uses the standard deviation of real images to describe contrast of image; however, there is a special contrast definition and some visual models in the study of HVS characteristics, which can better reflect contrast perception of HVS. Further, the most direct reason is that the real images generally have more redundancy; therefore, the consistency between the subjective IQA results and the objective IQA scores is often not high, and their accuracy PLCC often do not exceed 0.8 [15]. HVS is a natural high-quality filter, which can better eliminate the visual redundancy of real image when perceiving. In subjective IQA, the object of evaluation is the perceived image; however, in objective IQA using IQA models, the object of evaluation is the real image. Hence, three features selected in SSIM need to be improved in combination with HVS characteristics. Based on these ideas, combining with HVS characteristics and SSIM method, a novel IQA model based on the structure similarity of perceived image is proposed in this paper, which is called as PSIM, namely the perceived structural similarity index of the image.

3.1. Method of applying CSF model to IQA

At present, in the research of IQA combining with HVS characteristics and their models, HVS masking characteristic is often used. However, it mainly reflects a visual masking effect, and there is no specific corresponding HVS characteristic model. Moreover, there are few studies and reports about the real application of HVS characteristics and their models in IQA. The reason is that there is still a difficult problem, that is, it is how to relate HVS model with image features to effectively apply HVS model to IQA. Hence, the following methods are proposed to solve this problem.

In order to relate image features with CSF model, the image is first required to be processed to obtain image features with the sparse representation. Currently, the main method of sparse representation is image spatio-temporal transformation, and discrete cosine transform (DCT) is the most commonly used transformation method. After image is transformed by DCT, the spectrogram in the transform domain is an energy redistribution image composed with the various frequencies and energy intensities. The frequency is generally called as the spatial frequency (f), which is defined as the number of periodic light and dark fringes per unit length. The frequency in CSF model is called as the angular frequency (written as f_θ), which is defined as the period's number of the periodic light and dark fringes in the grating of stimulating human eye in the unit viewing angle, and is expressed as Eq (5) [28].

$$f_\theta = \frac{N}{\theta},\tag{5}$$

where, θ is the observation viewing angle, N is the number of periods number of periodic light and dark stripes in the range θ.
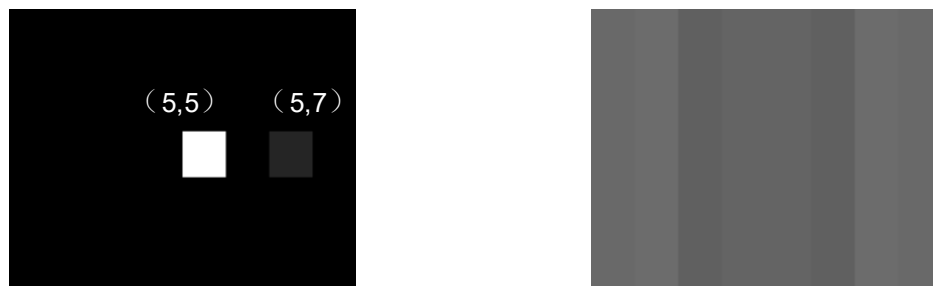
Based on the above description of two frequencies (f and f_θ), in IQA combining with CSF, the specific calculation method of the relationship between the two frequencies is illustrated as follows.

After image is transformed by DCT, the spectral coordinates (f_x, f_y) of any point on the spectrogram is the spatial frequency of that point. By experiments, it is found that, for any point in the frequency domain, inverse DCT transform (IDCT) is displayed as a vertical or horizontal or both interlaced fringe grating, the period's number of periodic fringes is $(f_x^2 + f_y^2)^{1/2}$, and the fringes cover the whole spatial image whose size is the same with the original spectrogram. The schematic diagram of experimental results to reflect the relationship between DCT and IDCT is shown Figure 1. These results show that when HVS perceiving image, HVS perceiving each point on image is essentially to

perceive its fringe grating with the corresponding frequency, and the perception results of the whole image is the synthetic effect of fringe gratings represented by all perceived points. Based on these explanations, the viewing angle corresponding to the size range of observation target is calculated and substituted into Eq (5). The relationship between two frequencies can be described by Eq (6).

$$f_\theta = \frac{N}{\theta} = \frac{\sqrt{f_x^2 + f_y^2}}{2 \times \frac{W \times (\frac{2.54}{P})}{D} \times \frac{180}{\pi}}, \tag{6}$$

where, D is observation distance, generally, it is 50 cm. P is pixel density of image, generally, it is 72. W is pixel number in the range of observation image. According to the size of sub-block of image, the range of each observation image includes 64 pixels. Considering the case of diagonal side of sub-block, the range of frequency f_θ can be taken as 0.546–6.16cpd (cycles per degree).

<div align="center">

(a) two points saved in DCT domain   (b) the stripes obtained by IDCT

</div>

Figure 1. Schematic diagram to reflect the relationship between DCT and IDCT of image.

Based on the above description, combining with Eqs (2)–(4) and (6), the contrast detection threshold and sensitivity threshold of HVS perceiving luminance and chroma of any point in image can be obtained when HVS perceives image, which can better realize the application of CSF in IQA.

3.2. IQA method and model

The IQA results using mathematical models that usually need to conform to the subjective perception of HVS. The perception of HVS detecting image mainly comes from the stimulation of luminance, chroma and contrast. Hence, HVS judging the image quality mainly depends on the measurement of the luminance, chroma and contrast information of perceived image [28,29]. Based on this explanation, the basic idea of the proposed IQA method is as follows:

First, the image is handled with the color space transformation, and is processed using the HVS nonlinear characteristic model of luminance and chroma intensity perception to obtain the results of the perceived luminance and chroma of images.

Then, the spatiotemporal transformation is carried out for the image, and according to the frequency relationship mentioned above, combining with the CSF model of HVS, the contrast detection threshold of each point on the image is calculated. The perceived image is then filtered using these thresholds, so as to remove the visual redundancy of the image.

Then, according to these calculated thresholds, the CS value of each point is obtained, and the

sensitivity is used as the weight factor of each point after filtering, and they are weighted.

After that, the weighted information is transformed by the inverse spatiotemporal transformation and inverse color space transformation to obtain the perceived image from simulating HVS.

Finally, based on the modeling idea of SSIM, the original image is replaced by the perceived image; and the three features, namely, the luminance, contrast and structure in the perceived image, are used as image features. They are substituted into the SSIM model to calculate the image quality scores. Thus, the calculated scores are taken as the IQA results.

The scheme flow for building IQA model is shown in Figure 2, and the specific description is as follows:
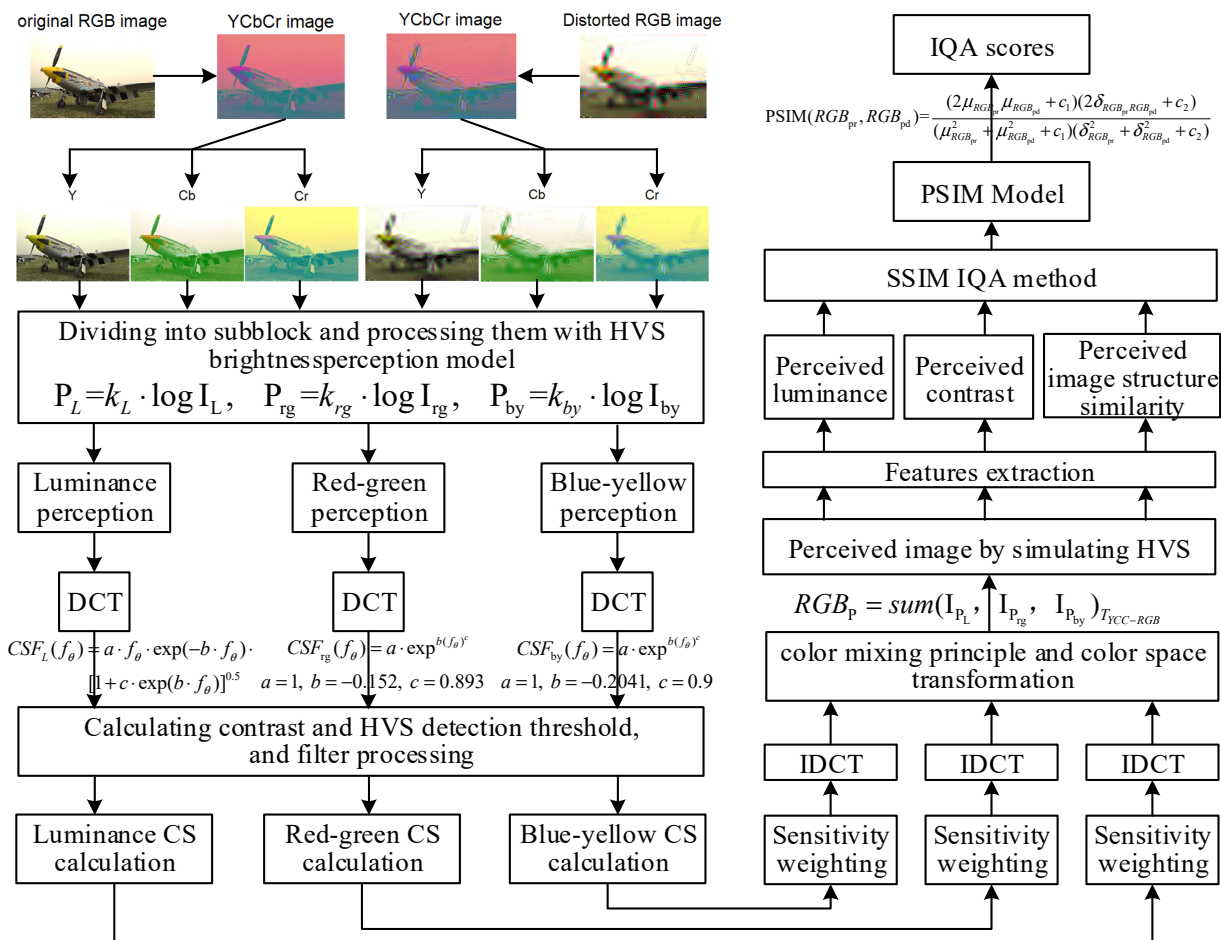
Figure 2. The flow chart of the proposed IQA method.

In SSIM model, the mean of luminance is used as the estimation of luminance, the standard deviation is used to quantify the contrast, and the covariance is used to measure the degree of structural similarity. Based on the above ideas and flow charts, the main idea of building IQA model in this paper is: taking the luminance of the perceived images as the luminance feature in SSIM calculation formula, taking the standard deviation of the image processed by CSF model as the contrast in SSIM model, and the structural similarity is measured by the covariance of the image processed by the HVS perception model. The specific construction methods and steps are as follows:

1) RGB image is transformed with a color space conversion formula (Eq (7)), to obtain the YC_rC_b

color space image. Thus, three-component images, namely luminance (*L*), red-green (*rg*: red-green), and blue-yellow (*by*: blue-yellow), are obtained.

$$Y=0.257 \times R+0.504 \times G+0.098 \times B+16$$

$$C_b=\text{-}0.148 \times R\text{-}0.291 \times G+0.439 \times B+128 \qquad (7)$$

$$C_r=0.439 \times R\text{ -}0.368 \times G\text{-}0.071 \times B+128$$

Here, in order to be consistent with the color space HVS characteristics model to reduce the computational complexity of the proposed IQA model, in color space conversion, it is selected to transform into the YC_rC_b color space.

Then, three-components images are, respectively, divided into the subblocks with 8pixels×8 pixels. After that, the nonlinear characteristic function of the HVS perceived intensity (*I*) (Eq (8)) is used to process each pixel value on each sub block to obtain subjective perception results (written as P_L, P_{rg}, P_{by}, respectively) of HVS perceiving the luminance and chroma intensity of image.

$$P_L = k_L \cdot log\, I_L\,,\ P_{rg} = k_{rg} \cdot log\, I_{rg}\,,\ P_{by} = k_{by} \cdot log\, I_{by}, \qquad (8)$$

where, k_L, k_{rg}, k_{by} all are parameters. For convenience of the IQA model calculation, their values are all 1.

2) When HVS perceives the target, the HVS target detection and obtaining its information is based on the spatial contrast of target, and involves the frequency sensitivity and contrast sensitivity characteristics of HVS. Combining this with visual masking characteristics, HVS sensitivity is based on the contrast of the target, that is, when the contrast of the target information is less than the HVS contrast threshold, HVS cannot detect differences among targets; when it is greater than HVS contrast threshold, it is conducive for HVS to detect and can distinguish targets [30,31]. Based on these analyses, the calculation method of the perceived image is as follows: First, contrast detection threshold is used to filter the perceived results of image luminance and chroma. Then, contrast sensitivity values are taken as the weight factors of the filtered results above, and they are weighted. So, the sum of the weighted results is used as the perceived image of HVS perceiving the real image, whose method is detailed as follows:

i) Each sub block of the subjective perceived value P of the luminance and chroma of image is transformed with DCT, and frequency shift is carried out to make the central spectrum be zero. Then, the spatial frequency of each point on each spectrum is calculated.

ii) According to the relationship between the spatial frequency and angular frequency of the image mentioned above (i.e., Eq (6)), the angular frequency of each point in the transformation domain of the sub block is calculated, and is subsituted into the luminance and chroma CSF models (i.e., Eqs (2)–(4)). Thus, the contrast detection threshold (written as $C_{T\text{-}L}$, $C_{T\text{-}rg}$, and $C_{T\text{-}by}$) of each point on the sub block is calculated. Then, using these contrast detection thresholds, the image is filtered (written as "Fil" in the following formula) to eliminate image visual redundancy.

iii) Meanwhile, based on the calculated contrast detection thresholds, the CS values of each point on the three-components images, namely, CSF_L, CSF_{rg} and CSF_{by}, are calculated and normalized. They are then used as the weight factor of each point on the sub block after being filtered, and are weighted. The weighted results are transformed by IDCT to obtain the sub block of HVS perceiving image in YC_rC_b color space. The calculation is shown in Eq (9).

$$I_{P_{\text{L\_block}}}=\text{IDCT}\{CSF_L \cdot \text{DCT}[Fil(P_L)]\}$$

$$I_{P_{\text{rg\_block}}}=\text{IDCT}\{CSF_{rg} \cdot \text{DCT}[Fil(P_{rg})]\} \tag{9}$$

$$I_{P_{\text{by\_block}}}=\text{IDCT}\{CSF_{by} \cdot \text{DCT}[Fil(P_{by})]\}$$

iv) All perceived sub blocks are sorted and synthesized to obtain the three-components images (written as I_{P_L}, $I_{P_{rg}}$, $I_{P_{by}}$) of HVS perceiving in YC$_r$C$_b$ color space. They are then transformed into the RGB color space. Then, the HVS perceived image in RGB color space is obtained, which is recorded as RGB_p. The calculation is shown in Eq (10).

$$RGB_P = sum(I_{P_L},\ I_{P_{rg}},\ I_{P_{by}})_{TYCC-RGB}, \tag{10}$$

where, when the perceived image is transformed into RGB color space from YC$_r$C$_b$ color space, the relations between Y, C$_r$, C$_b$ and R, G, B, including their parameters, all follow the color mixing principle and their transformation formulas, which are shown in Eq (11).

$$\text{R}=1.164 \times (\text{Y-16})+1.596 \times (C_r\text{-}128)$$

$$\text{G}=1.164 \times (\text{Y-16})\text{-}0.392 \times (C_b\text{-}128)\text{-}0.813 \times (C_r\text{-}128) \tag{11}$$

$$\text{B}=1.164 \times (\text{Y-16})+2.017 \times (C_b\text{-}128)$$

3) Using the above methods, the reference image and the distorted image are processed, respectively, to obtain their perceived images, which are, respectively, recorded as RGB$_{pr}$ and RGB$_{pd}$. Based on the idea of SSIM, the luminance, contrast and structural similarity of the perceived images (RGB$_{pr}$ and RGB$_{pd}$) are used to replace the corresponding three features of the real image required in SSIM, so as to simulate HVS to evaluate the quality of distorted image. That is, in SSIM model (Eq (1)), the *x* and *y* are replaced by RGB$_{pr}$ and RGB$_{pd}$, respectively, and the calculated value is used as the IQA score of the distorted image. This IQA method is called as the perceived structural similarity index (PSIM) of the image, whose calculation is as shown in Eq (12).

$$\text{PSIM}(RGB_{\text{pr}}, RGB_{\text{pd}}) = \frac{\left(2\mu_{RGB_{\text{pr}}}\mu_{RGB_{\text{pd}}}+c_1\right)\left(2\delta_{RGB_{\text{pr}}RGB_{\text{pd}}}+c_2\right)}{\left(\mu_{RGB_{\text{pr}}}^2+\mu_{RGB_{\text{pd}}}^2+c_1\right)\left(\delta_{RGB_{\text{pr}}}^2+\delta_{RGB_{\text{pd}}}^2+c_2\right)}, \tag{12}$$

4. Experiment and result analyses

To test the proposed PSIM model, the images in LIVE [14], TID2013 [15] and CSIQ [16] databases are used for simulation. Three open source databases provide 84 (25, 30 and 29, corresponding to three databases) reference images and 4645 (3000, 866 and 779, corresponding to three databases) distorted images, and include 35 distortion types. In order to better illustrate the performance of the proposed PSIM model, experiments are carried out in the following three ways: 1) overall IQA: that is to evaluate all distorted images in each database, in order to illustrate the overall

IQA effect of the model; 2) IQA of each type of distortion: that is that the images of each distortion type in the database is carried out IQA simulation, and the IQA effect of PSIM on each distortion type is studied to explore its generalization performance for distortion types; 3) quality assessment of the images distorted by CC and CCS: for the images distorted from contrast change and change of color saturation, SSIM evaluation effect is poor; while PSIM is used to evaluate them to illustrate the improvement effect.

1) Overall IQA of all images in each of three databases

The specific methods are as follows: i) First, PSIM is used to evaluate all images in three databases, and the objective IQA scores are obtained. The IQA results of each database are taken as a whole to analyze separately overall IQA effect of each database. Then, when combining this with the subjective IQA scores (MOS or DMOS (Differential MOS)) [14−16] in three open source databases, the scatter plots between the subjective and objective IQA scores are obtained, and the consistency parameters between them are calculated. ii) At the same time, SSIM, MSSIM [7] and FSIMc models are used to evaluate all images in each of the three databases, and the scatter plots and four consistency parameters are also obtained. iii) Finally, based on the scatter plots and four consistency parameters, IQA effects of several IQA models are compared, which are shown in Figure 3. Here, at present, the IQA models with better improvement effect aiming at SSIM mainly include MSSIM and FSIMc. Although the two IQA models are relatively complex, their accuracy has been improved. In order to better illustrate the advantages of the proposed model, their simulation results are compared with the ones of PSIM.

In addition, the consistency parameters between the subjective and objective IQA scores mainly include PLCC, SROCC, RMSE (root mean squared error), and OR (outlier ratio) [32,33]. Among them, the calculation of OR is shown in Eq (13), where N_{img} is the number of distorted images to be evaluated, MOS_P is the predicted IQA scores obtained by fitting with the logic function of 5 parameters. It's fitting curve is in Figure 3.

$$OR = \frac{1}{N_{img}} \sum_{i=1}^{N_{img}} \left| \frac{MOS_p(i) - MOS(i)}{MOS_p(i)} \right| \tag{13}$$

By comparing the experimental results of SSIM, MSSIM and FSIMc models with that of the proposed PSIM model in Figure 3, it can be obtained that: i) analyzing from the dispersion degree of dots in the scatter plots and correlation parameter values, for the overall IQA effect of each database in the three databases, PSIM is better than SSIM, MSSIM and FSIMc, namely, the accuracy of PSIM is higher than that of the three IQA models, and the dispersion degree of the dots in the scatter plots is lower than that of three IQA models. ii) MSSIM, FSIMc and PSIM are all the improved IQA models based on SSIM. Compared with SSIM, using weighting, the accuracy of the PLCC value of PSIM is improved, on average, by 10.6074% in three databases. Compared with MSSIM and FSIMc, in the accuracy PLCC value, PSIM is higher by 5.5812% and 2.2097% than the two IQA models. These results indicate that the proposed PSIM model has the highest accuracy in the current improvement effect based on SSIM. At the same time, it also shows that the introduction of CSF and perceptual image has a good effect on the improvement of SSIM, and can more truly and effectively reflect and measure image quality.

In order to better illustrate accuracy of PSIM, another 9 existing IQA models, namely VSNR [6], GMSD [9], VSI [10], PSNRHVS [14], A3L [17], DISTS [18], BMPRI [19], VCRNet [20] and ASCAM [21], are used to evaluate all images in three databases. Based on PLCC and SROCC

parameters, their results are compared with the above experiment results of PSIM. The comparisons are shown in Table 1.
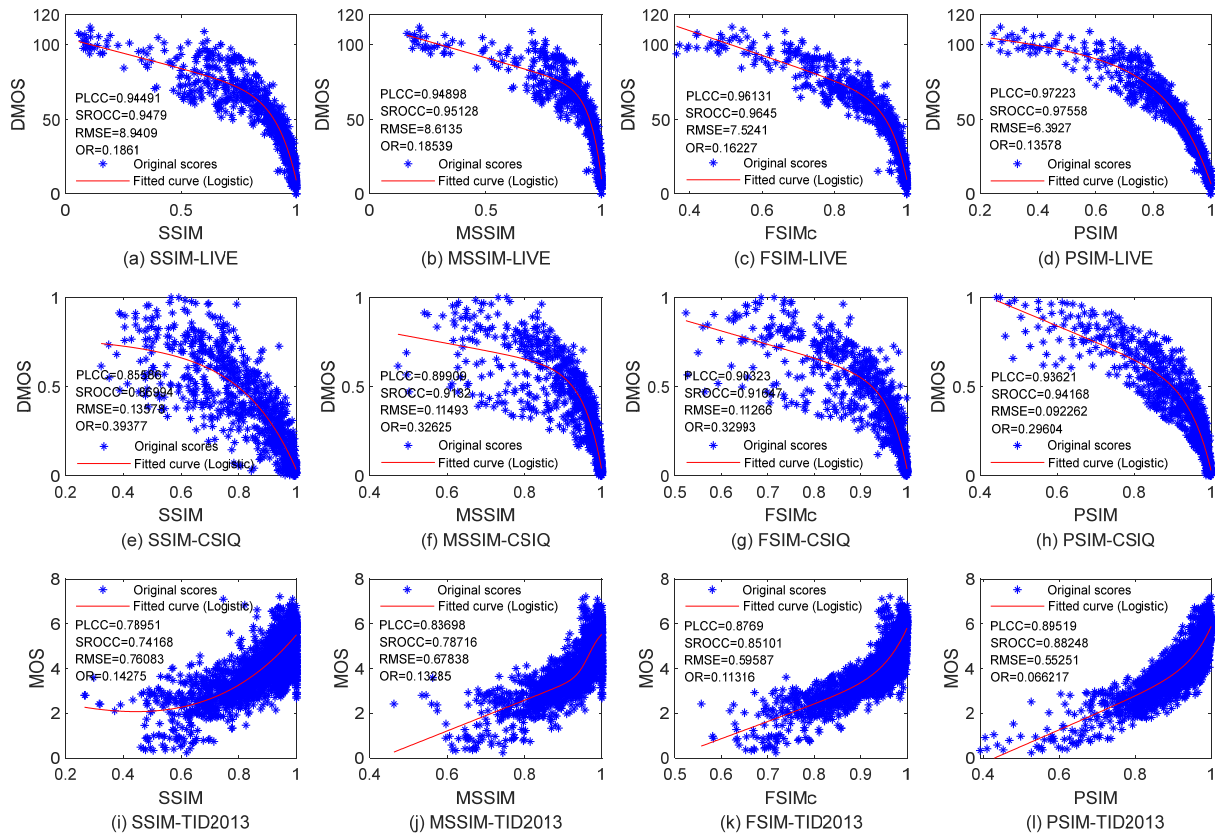
Figure 3. Comparing the accuracy of PSIM with ones of SSIM, MSSIM and FSIMc, based on the IQA results of all images in TID2013, CSIQ and LIVE databases.

Table 1. Performance comparison between PSIM and 9 existing IQA models based on PLCC and SROCC parameters of IQA results in three databases (the top three results are shown in bold).

Databases		TID2013(3000)		CSIQ (866)		LIVE (779)		Weighting	
Parameters		*PLCC*	*SROCC*	*PLCC*	*SROCC*	*PLCC*	*SROCC*	*PLCC*	*SROCC*
IQA models	VSNR [6]	0.7406	0.7197	0.8129	0.8268	0.9231	0.9274	0.7847	0.7745
	GMSD [9]	0.8590	0.8044	**0.9471**	**0.9498**	0.9603	0.9603	0.8924	0.8577
	VSI [10]	**0.9000**	**0.8965**	0.9279	**0.9422**	0.9482	0.9524	**0.9133**	**0.9144**
	PSNRHVS [14]	0.7209	0.7061	0.8238	0.8303	0.9182	0.9234	0.7732	0.7657
	A3L [17]	0.8713	0.8553	—	—	**0.9640**	**0.9667**	0.8904	0.8783
	DISTS [18]	0.8550	0.8300	0.9280	0.9290	0.9540	0.9540	0.8852	0.8693
	BMPRI [19]	**0.9466**	**0.9287**	0.9339	0.9085	0.9329	0.9310	**0.9419**	**0.9253**
	VCRNet [20]	0.8750	0.8460	**0.9550**	**0.9520**	**0.9740**	**0.9730**	0.9065	0.8871
	ASCAM [21]	0.7460	0.7350	—	—	0.8800	0.8720	0.7736	0.7632
	PSIM	**0.8952**	**0.8825**	0.9362	0.9417	**0.9722**	**0.9756**	**0.9158**	**0.9091**

By comparing and analyzing the data in Table 1, we can conclude the following: From both the overall IQA in each of three databases and the weighted IQA results of three databases, when it comes to accuracy, the proposed PSIM model is significantly better than PSNRHVS, VSNR and ASCAM, and is close to GMSD, VSI, A3L, DISTS and VCRNet, though being slightly better than them, and is slightly worse than BMPRI. In addition, the accuracy of PSIM is significantly higher than that of the WS-HV [13] model proposed recently.

By analyzing comprehensively the results in Table 1 and Figure 3, and further comparing PSIM with 12 existing IQA models, it can be obtained that, in the IQA of each of three databases, the PLCC accuracies of PSIM are all over 0.89, with a small fluctuation. Its IQA accuracy and stability are better than the above 11 IQA models, except for BMPRI, and the weighted PLCC of three databases is larger than that of 11 existing IQA models. These results show that PSIM has both high accuracy and good generalization performance.

The reasons are analyzed as follows: BMPRI, VCRNet and ASCAM are IQA models based on the neural networks (IQA-NN), and the other 10 IQA methods (including PSIM) are modeled using the mathematical methods. From their modeling processes and experimental results, the accuracy of IQA-NN is related to the training data. When training and testing IQA models, if the data is from the same database, their IQA accuracy is often higher, such as the results of BMPRI [19] and VCRNet [20]. However, if different databases are used for training and testing, their IQA accuracy will be significantly reduced, such as the results of ASCAM [21]. It shows that the generalization performance of IQA-NN is often not high. The modeling process of SSIM, FSIM, MSSIM, VSNR, GMSD, VSI, PSNRHVS, A3L, DISTS is often not so strong related to the fitting data. Compared with three IQA-NN models, although the highest accuracy of 10 IQA models (including PSIM) using the mathematical methods is lower than them, their accuracy fluctuation of 10 IQA models is small, and their stability is better, which present better generalization performance. In addition, PSIM is based on SSIM, and combines HVS characteristics well. Hence, the accuracy of PSIM is improved well, and higher than the above existing 11 IQA models.

2) IQA for all distortion types in three databases

The proposed PSIM model is used to separately evaluate each type of distorted images of all distortion types (35 types in total) in the three databases, the objective IQA scores of each type of distorted images are predicted. By combining these with the subjective IQA scores (MOS or DMOS) [14−16] provided in the three open source databases, the consistency parameter between the subjective and objective IQA scores of each type of distorted images are calculated and the scatter plot is obtained. These results are shown in Figures 4–6.

At the same time, based on the correlation parameter PLCC values, the results are compared with the experimental results of the above seven common IQA models. The comparison results are shown in Figures 7–9.

In combination with Figures 4–6, and analyzing Figures 7–9, it can be obtained that, for 35 distortion types in three databases, PSIM shows good IQA effect for each type of distorted images. i) Among 35 types, for PSIM, its minimum IQA accuracy PLCC can reach 0.6949 and the maximum value can reach 0.9896. While for seven existing models, the differences between their maximum and minimum IQA accuracy PLCC is more than 0.5. The fluctuation of their IQA accuracy is significantly greater than that of PSIM, which indicates that PSIM has better IQA generalization performance. ii) For each distortion type, IQA accuracy of PSIM can reach the highest or second highest among eight methods, which indicates that PSIM has high accuracy and good generalization performance. iii)

PLCC accuracy values of PSIM are basically all higher than that of PSNRHVS, SSIM and MSSIM, and is higher than that of FSIMc in 21 types of distorted images. These results show that combination of CSF model and perceptual image in IQA can effectively improve IQA performance of SSIM.
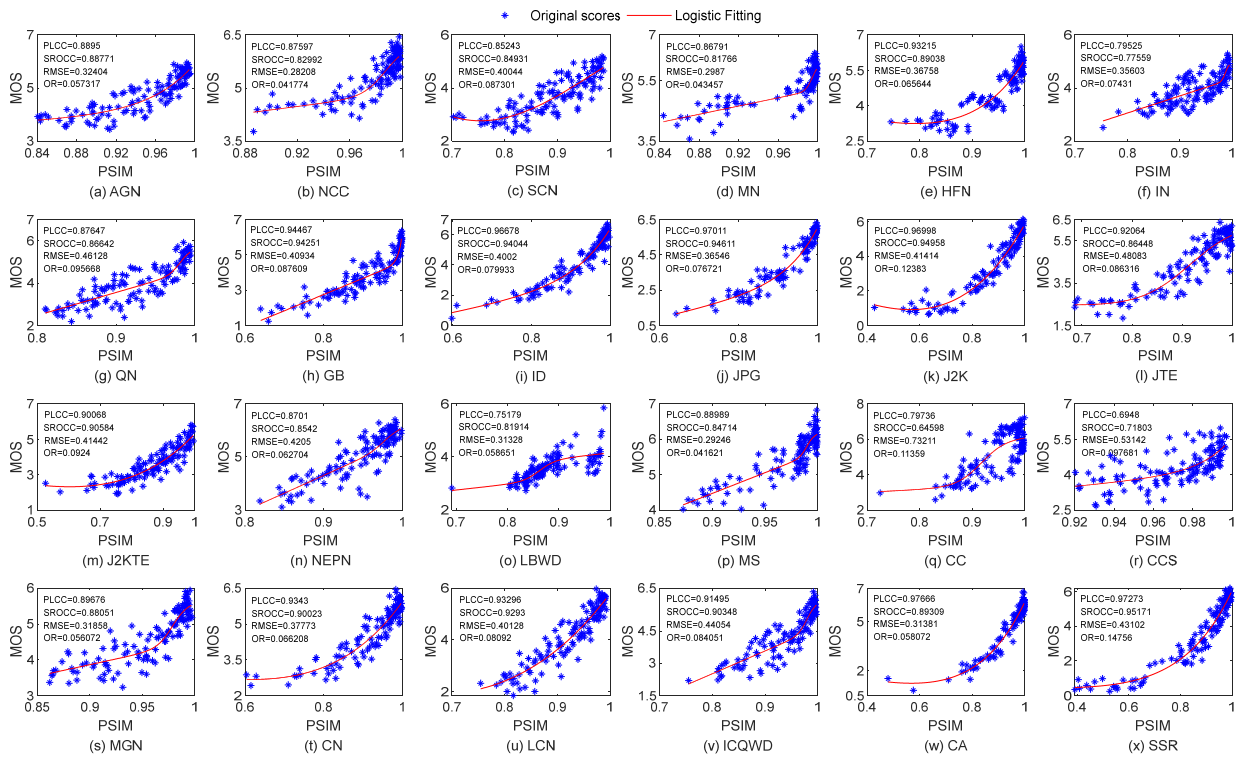

Figure 4. The IQA results of using PSIM to evaluate images from each of 24 distortion types in TID2013 database.
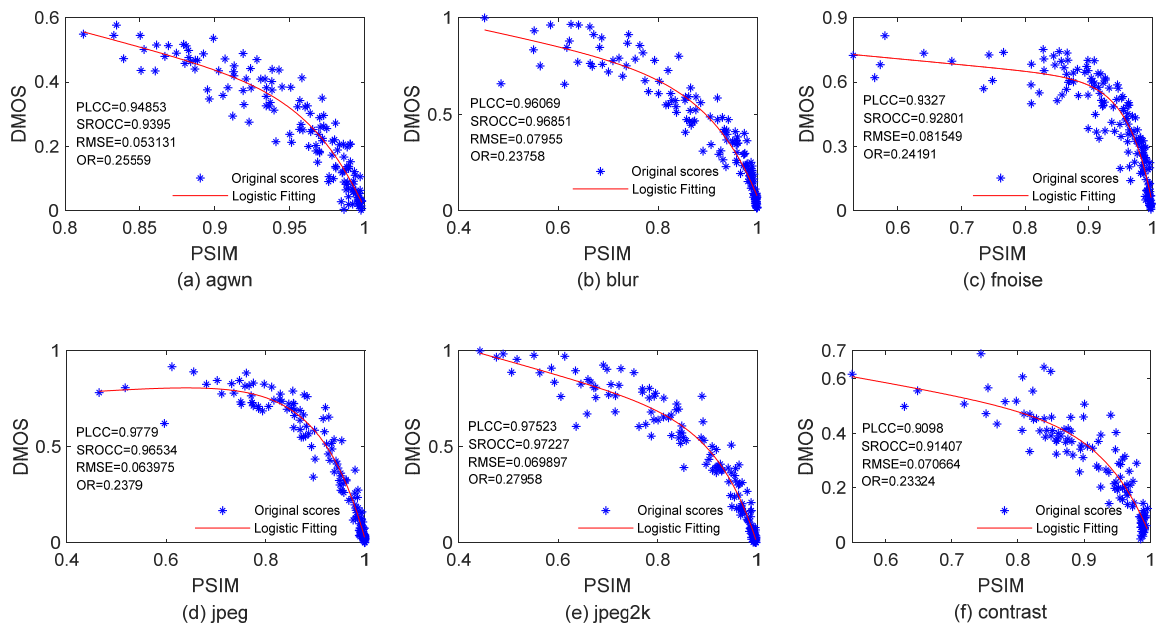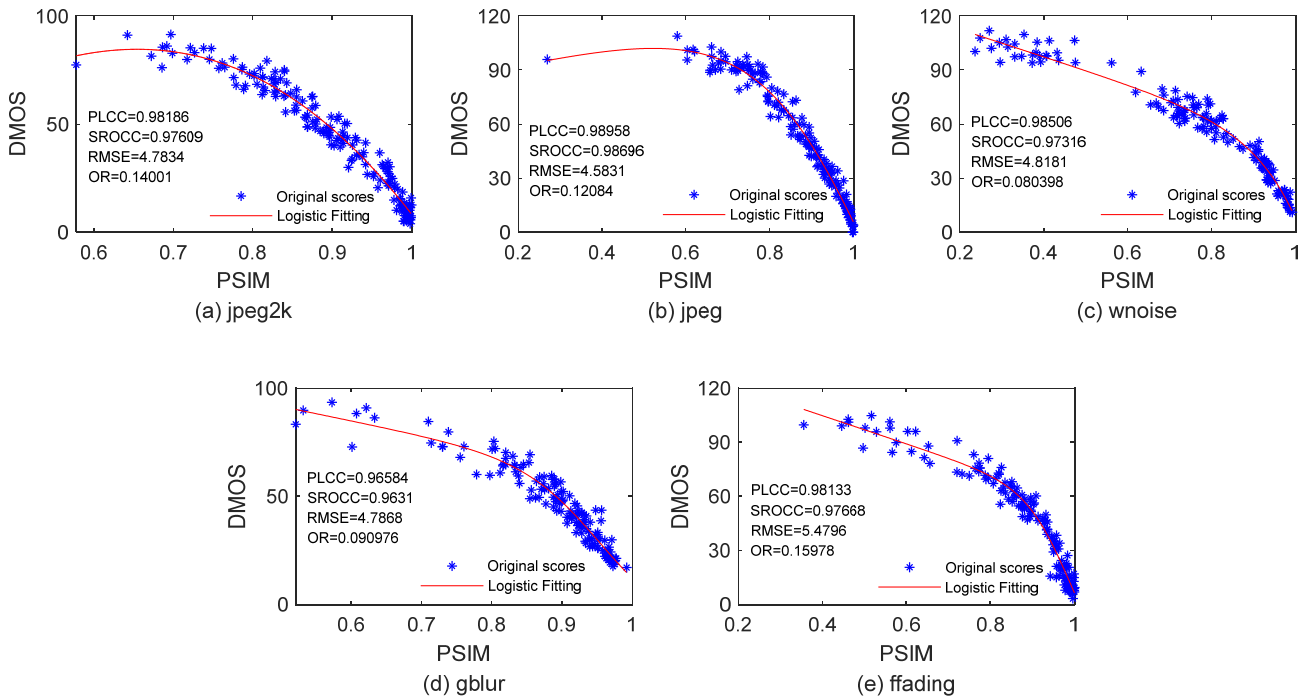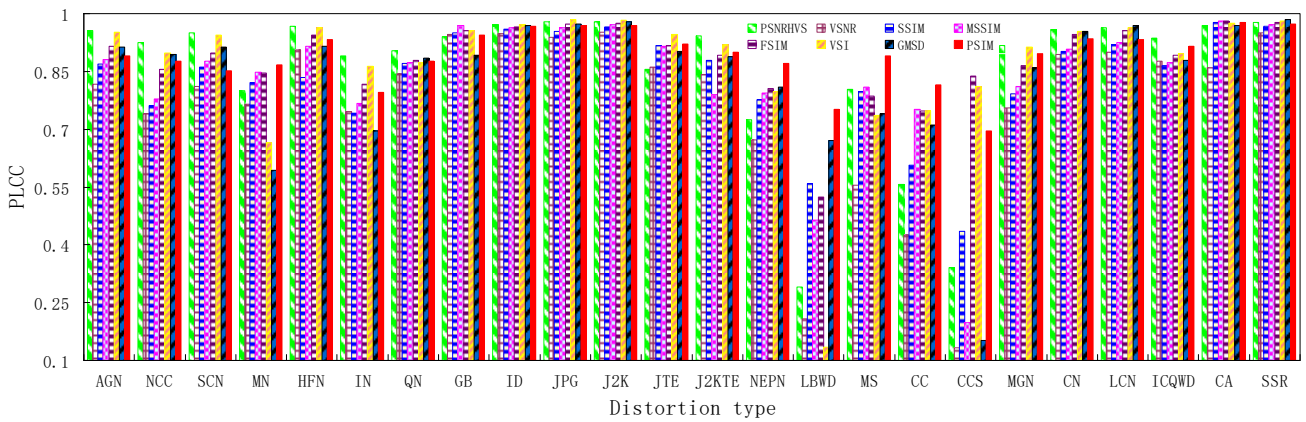

Figure 5. The IQA results of using PSIM to evaluate images from each of 6 distortion types in CSIQ database.
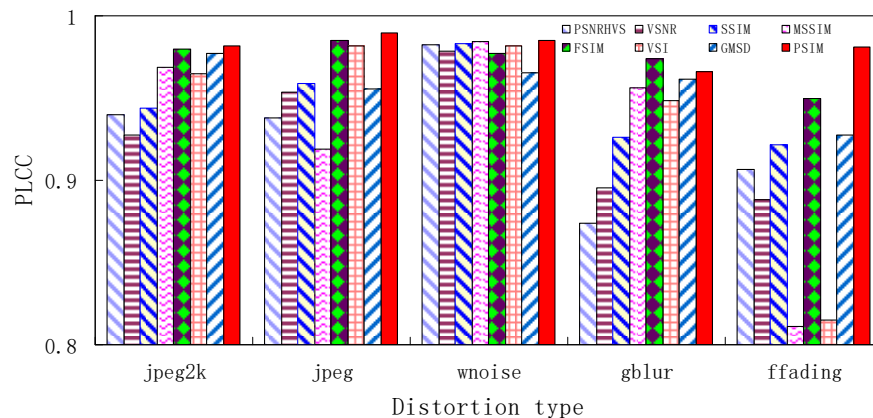
Figure 6. The IQA results of using PSIM to evaluate images from each of 5 distortion types in the LIVE database.

Figure 7. Performance comparison between PSIM and 7 existing models based on PLCC of IQA results of 24 types of distorted images in TID2013 databases.
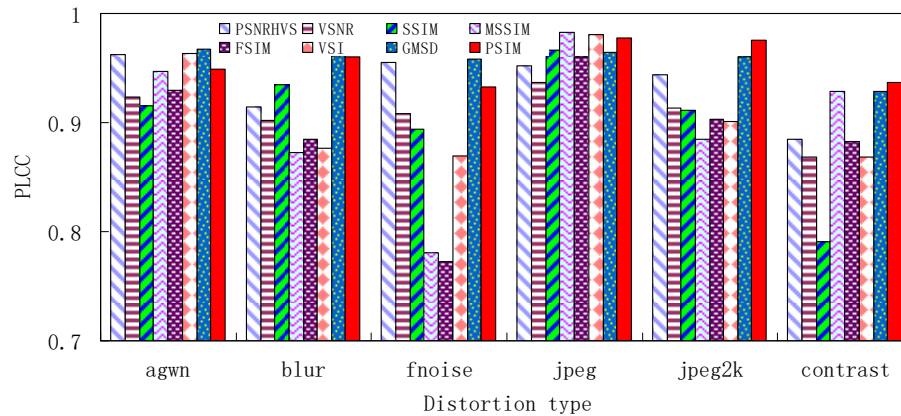
Figure 8. Performance comparison between PSIM and 7 existing models based on PLCC of IQA results of 6 types of distorted images in CSIQ databases.
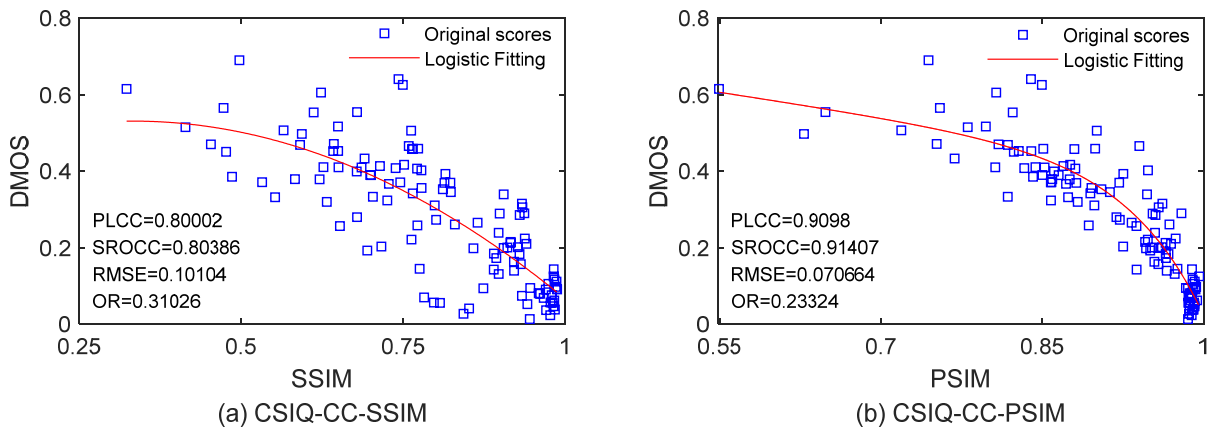

Figure 9. Performance comparison between PSIM and 7 existing models based on PLCC of IQA results of 5 types of distorted images in LIVE databases.

3) IQA of images distorted by CC and CCS

To better illustrate the improvement effect of PSIM on SSIM, two types of distorted images, namely contrast change (CC, in total 241 images in two databases [15,16]) and change of color saturation (CCS, 125 images in TID2013 [15]) distorted images, are selected to be tested, because the accuracy of SSIM for them is the lowest in evaluating the quality of all distortion types of images in TID2013 and CSIQ. They are also evaluated by PSIM. Thus, in combination with the subjective IQA scores [15,16] in the open source databases, four correlation parameters values are calculated. At the same time, SSIM is used to evaluate the 366 distorted images, and the IQA results are compared with the IQA results of PSIM, whose comparison and analysis results are shown in Figures 10 and 11 and Table 2.

From the results in Figures 10 and 11 and Table 2, we can conclude the following: i) For the CC distorted images in two databases, compared with SSIM, the weighted IQA accuracy PLCC value of PSIM is increased by 21.6069%, which indicates that PSIM has a great improvement in accuracy and generalization performance. ii) For the CCS distorted images in TID2013, the IQA accuracy PLCC value of PSIM is 59.7839% higher than that of SSIM, which shows that for the CCS distortion, it is very obvious to use the idea in this paper to improve the IQA effect of SSIM. iii) After weighting three

groups of data in the two databases, the IQA accuracy PLCC value of PSIM is improved by 30.9089% compared with SSIM. The above comparison and analysis show that the proposed model is very effective in improving SSIM, and greatly improves the accuracy and generalization performance.

Figure 10. Performance comparison between SSIM and PSIM based on the IQA results of CC distorted images in the CSIQ database.
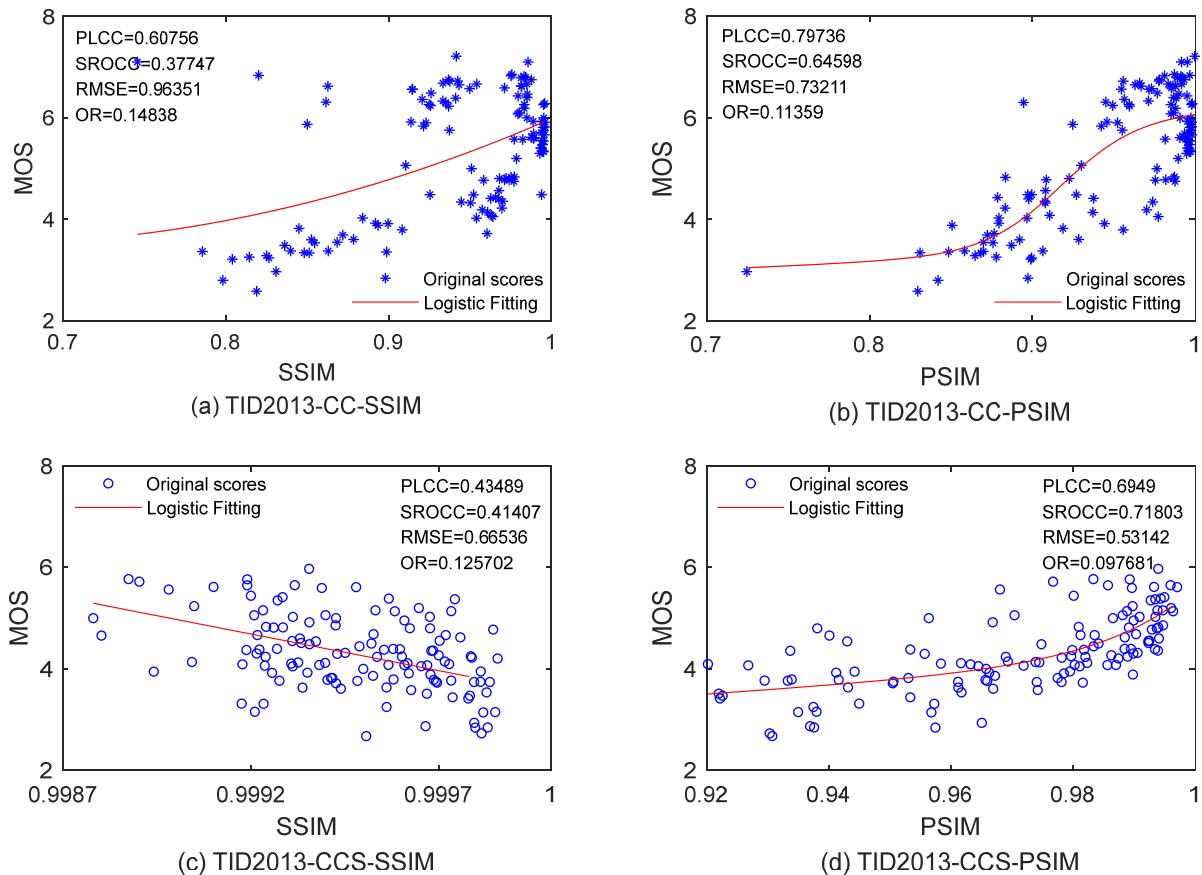
Figure 11. Performance comparison between SSIM and PSIM based on the IQA results of the CC and CCS distorted images in the TID2013 database.

Table 2. Performance comparison between PSIM and SSIM based on the four parameters of IQA results of CC and CCS distorted images.

Images	125 CC distorted images in TID2013		125 CCS distorted images in TID2013		116 CC distorted images in CSIQ		Weighting	
IQA models	SSIM [7]	PSIM	SSIM [7]	PSIM	SSIM [7]	PSIM	SSIM [7]	PSIM
Parameters *PLCC*	0.6076	**0.7974**	0.4349	**0.6949**	0.8000	**0.9098**	**0.6096**	**0.7980**
SROCC	0.3775	**0.6460**	0.4141	**0.7180**	0.8039	**0.9141**	**0.5251**	**0.7555**
RMSE	0.9635	0.7321	0.6654	0.5314	0.1010	0.0707	——	——
OR	0.1484	0.1136	0.1257	0.0977	0.3103	0.2332	——	——

Further, in order to better illustrate the advantages of PSIM, other six existing IQA models are used to evaluate the CC distorted images in the two databases, and their objective IQA scores are obtained, and then four correlation parameters are calculated. Based on the four parameters, the performances of PSIM are compared with that of the six existing models, and the results are shown in Table 3.

By comparing and analyzing the results in Table 3, the following conclusions can be drawn: I) For the CC distorted images in TID2013 database, the IQA accuracy of PSIM is better than that of six existing IQA models. Ii) When evaluating the CC distorted images in CSIQ database, the IQA accuracy of PSIM is significantly better than PSNRHVS and VSNR, and close to four models, namely MSSIM, GMSD, VSI and FSIMc. However, when evaluating the CC distorted images in TID2013 database, the IQA accuracy of four models is significantly lower than that of PSIM. Two results indicate that the generalization performance of PSIM is better than the four models. Iii) In terms of the weighting IQA accuracy of the two databases, PSIM is higher than the six models. Comprehensively analyzing the above IQA results of evaluating CC, CCS distorted images and weighted IQA results, based on the comprehensive benefits of accuracy and generalization performance, it shows that the performance of PSIM is better than the six existing IQA models.

Table 3. Performance comparison between the 6 existing models and PSIM based on the IQA results of the CC distorted images in TID2013 and CSIQ databases.

Images	125 CC distorted images in TID2013				116 CC distorted images in CSIQ				Weighting	
Parameters	PLCC	SROCC	RMSE	OR	PLCC	SROCC	RMSE	OR	PLCC	SROCC
IQA models PSNRHVS [14]	0.5575	0.4428	1.0120	0.1738	0.8847	0.8713	0.0785	0.2645	0.7150	0.6490
VSNR [6]	0.4257	0.3514	1.0977	0.1844	0.8688	0.8801	0.0834	0.2858	0.6390	0.6059
MSSIM [7]	0.7512	0.4684	0.8007	0.1179	0.9281	0.9528	0.0628	0.2313	0.8363	0.7016
FSIMc [8]	0.7481	0.4680	0.8050	0.1184	0.8828	0.9351	0.0792	0.2664	0.8129	0.6928
VSI [10]	0.7474	0.4754	0.8059	0.1200	0.8687	0.9505	0.0834	0.2897	0.8058	0.7041
GMSD [9]	0.7107	0.3235	0.8534	0.1359	0.9284	0.9086	0.0626	0.2301	0.8155	0.6051
PSIM	**0.7974**	**0.6460**	**0.7321**	**0.1136**	**0.9098**	**0.9141**	**0.0707**	**0.2332**	**0.8515**	**0.7814**

5. Discussion

1) IQA model generalization performance

In order to better verify the proposed PSIM model, another image database with obvious image feature difference from the above three databases need to be used to test it to discuss its performance. The images in CID database [34] are significantly different from those in CSIQ, TID2013 and LIVE databases, including image contents, scenes, resolution, and distortion types. Using them as the test object can verify the performance of PSIM. PSIM is used to evaluate images in CID database; and four performance parameters, namely PLCC, SROCC, RMSE and OR, are calculated based on MOS in the CID database. Further, five existing typical IQA models, namely SSIM [7], FSIMc [8], GMSD [9], VSI [10] and MAD (most apparent damage algorithm) [35] are used to evaluate them, and their results are compared with the results of PSIM. The comparison is shown in Figure 12.

By comparing and analyzing the results in Figure 12, we can conclude the following: i) The IQA accuracy of the proposed PSIM model in CID database is the best in six IQA methods, and the effect of its four performance parameters is the best. ii) Although image features in CID database are significantly different from those in the three databases above, the IQA accuracy (PLCC and SROCC) of PSIM in CID database can still reach more than 0.81; iii) Compared with SSIM and FSIMc, the precision of PSIM has increased by 3.66% and 13.29%, in terms of PLCC values. Combined with the IQA results of the above three databases, the precision of PSIM has increased by 9.91%, on average, compared with that of SSIM in the four databases. The analysis results show that PSIM not only has a high accuracy, but also has a good generalization performance. At the same time, it shows that it has a better improvement effect compared with SSIM.
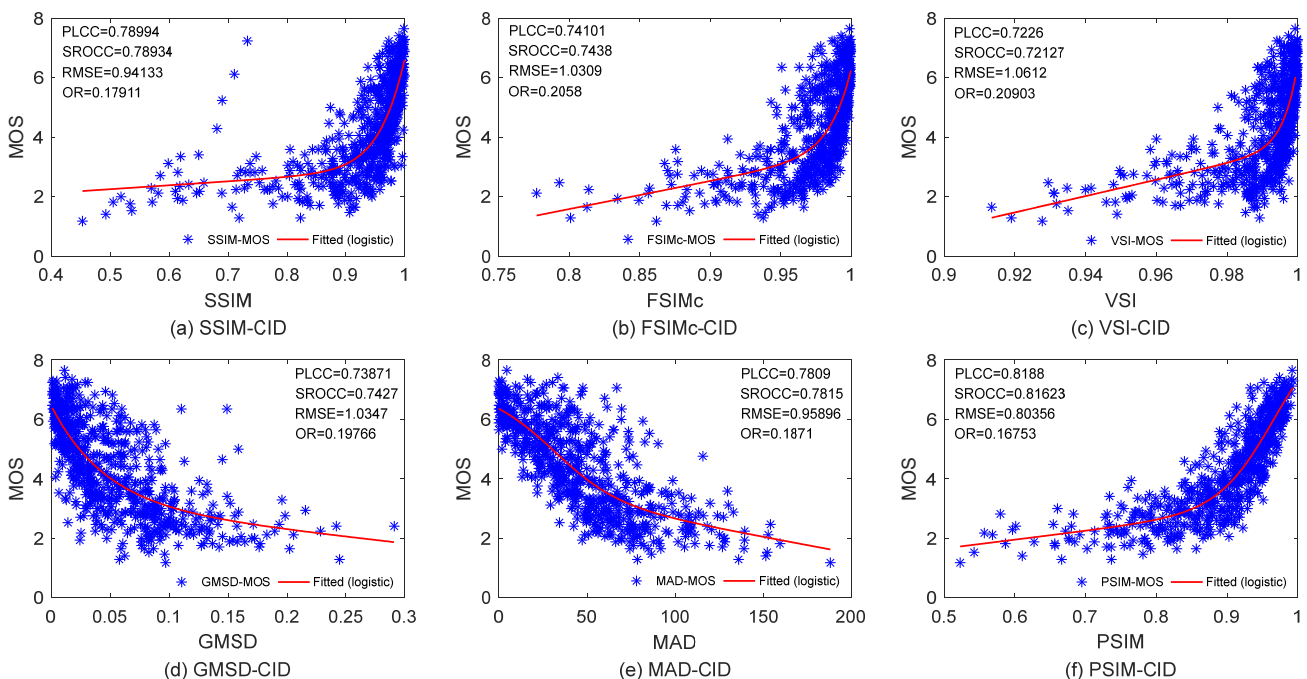
Figure 12. Comparing the performance of PSIM with those of five existing IQA models based on the IQA results in the CID database.

In order to further analyze model performance and discuss the reasons for their advantages and shortcomings, the above IQA experiments and results in four databases using 8 existing IQA models and PSIM are compared and comprehensively analyzed. The main reasons for obtaining the above IQA model performance are as follows:

At present, internationally, many IQA metrics have been proposed, including mainly PSNR, PSNRHVS, SSIM, MSSIM, FSIMc, VSNR, VSI, GMSD and MAD. However, the IQA metrics recognized internationally are fewer. Among these models, currently, PSNR and SSIM (MSSIM) are the most widely applied. Accuracy of SSIM is higher than that of PSNR. However, PSNR is simpler than SSIM, so it is applied more widely. However, in general, their accuracies are not high. Because they are both very simple, however, much work has recently been done to improve them . PSNRHVS and FSIMc are two of the improved models. However, PSNRHVS has a poor generalization performance for different types of distortion, and they are much complex. For GMSD, it has partially been accepted internationally, currently, because of its good performance, and is intended for gray images. Namely, when evaluating color image, converting color images into gray images is necessary. As it is well known, HVS is more sensitive to luminance than to color. Therefore, performance of GMSD would be greatly reduced in practice. VSNR is a metric for quantifying the visual fidelity of natural images based on the near-threshold and suprathreshold properties of HVS. Due to combining with HVS characteristics and a two-stage approach computational process, VSNR is much complex and its accuracy is low. Hence, its current application is very limited. Using VSI to evaluate image quality is mainly to aim at the distortion of image saliency features. Experimental results in some frequently-used image databases show that it has poor generalization performance for some distortion types of images, including CC and CCS. For MAD, it is considered that HVS adopts a dual strategy to determine image quality. That is, for images with serious distortion, HVS will choose to ignore distortion and find image content to obtain perception; for slightly distorted images, HVS focuses on finding distortion parts to obtain perception. MAD mainly use the feature of local brightness, contrast masking, and changes in local statistics of spatial frequency components. The idea of designing MAD is much beneficial to improve performance, but it also leads to more complex calculation. Hence, its application lower.

In this paper, the CS characteristics of HVS and nonlinear characteristics of luminance perception are integrated in improving the SSIM model. Because the luminance perceived by HVS is different from real luminance, and the perceived degradation of luminance is different from that of real changing of luminance, the image perceived by HVS is different from real image and the perceived image distortion is different from true distortion. Hence, in PSIM, it is first combined with HVS characteristics, a model that HVS perceives real image is proposed. This model is then used to simulate HVS to perceive real images, in order to eliminate visual redundancy of real image. Then, the perceived images are evaluated, and the evaluated scores are the objective qualities of real images. Hence, the IQA results obtained by PSIM are more consistent with HVS perception. Thus, in combining with HVS characteristics to build an IQA model, the calculation of PSIM is not much, so its computational complexity is not high. Hence, compared with the above 8 existing IQA models, PSIM has better comprehensive performance.

2) IQA model complexity

In IQA, model complexity is very important and greatly affects the application of the IQA model. Hence, the complexity of PSIM needs to be analyzed and compared with the complexity of the 11 existing models above, namely VSNR [6], SSIM [7], MSSIM [7], FSIMc [8], GMSD [9], VSI [10],

WS-HV [13], PSNRHVS [14], BMPRI [19], VCRNet [20] and ASCAM [21]. In studying the complexity of the IQA model, the calculation running time of evaluating image quality by model is generally used to measure its complexity [36,37]. Hence, matlab2018a programming is used to realize PSIM and 11 IQA models, and 200 images are selected respectively from each of three databases to be evaluated using them. Then, the complexity of the model is quantitatively described by the average calculation running time of evaluating every images. Here, the experimental environment is: notebook with 64 bit operating system, and its processor is Intel (R) core (TM) i7-8550u CPU @ 1.8 GHz, 1.99 GHz. For images with different sizes, their IQA calculation running time is equivalent to the running time of evaluating images that are weighted to 512 pixels × 384 pixels [37,38]. The result is shown in Figure 13.
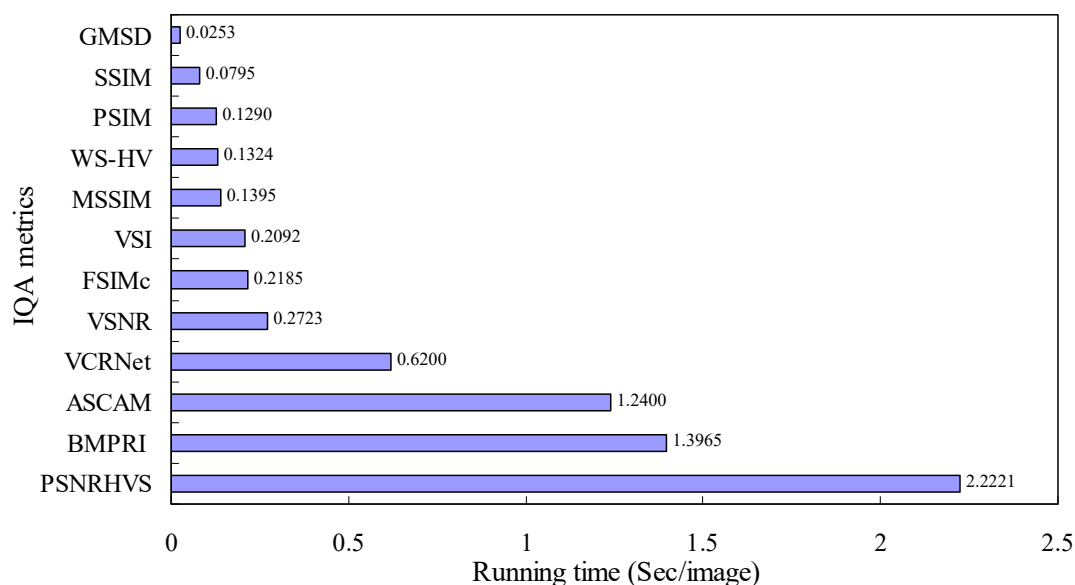
Figure 13. Average operation time of evaluating every image using PSIM and 11 existing models.

By comparing and analyzing the calculation running time of IQA model in Figure 13 and the model accuracy in Tables 1–3, we can conclude the following: i) The complexity of PSIM is lower than that of PSNRHVS, BMPRI, VCRNet, ASCAM, VSNR, FSIMc, VSI and MSSIM, and close to that of WS-HV. Its IQA accuracy in three databases is obviously higher than that of PSNRHVS, VSNR, WS-HV and ASCAM, slightly higher than that of MSSIM, and close to IQA accuracy of VSI, FSIMc, BMPRI and VCRNet. ii) For SSIM and GMSD, PSIM is greater than them in complexity. However, in terms of accuracy, for both the overall IQA of all images in each of three databases and the IQA of each distortion type of images, accuracy of SSIM is significantly lower than that of PSIM. For GMSD, its accuracy is high in three databases, but it shows great instability in the generalization performance of IQA of different distortion types of images. When evaluating quality of CCS distorted images, its accuracy can only reach 0.1516, but in evaluating quality of SSR (sparse sampling and reconstruction) distorted images, its accuracy can reach 0.9838. Furthermore, the accuracy of PSIM, compared with them, whether the overall IQA of all images in each database or the individual IQA of each distortion type of images, shows a good stability and high accuracy. Considering the accuracy, complexity and generalization performance, IQA effect of PSIM is better than that of the above 11 models.

The main reason is that SSIM is the simplest IQA model in the existing IQA models except PSNR.

Using this feature, PSIM adds CSF model and perceptual image model to build IQA model based on the idea of SSIM. Hence, compared with SSIM, the complexity of PSIM is increased, but the accuracy has been significantly improved.

At present, when building the IQA model, its complexity is very important and needs careful consideration due to its practical application. In the paper, HVS characteristics and their mathematical models are introduced to build PSIM IQA models, although its accuracy is improved, it increases the computational complexity of the model, too. PSIM can reach the better level of the above 11 models in terms of complexity, but it is still more complex than PSNR that is currently widely used. Hence, further optimization is needed in the future to make it simpler, and ensure that accuracy will not decline.

6. Conclusions

Aiming at the problem of low IQA accuracy when using SSIM to evaluate images distorted by CC and CCS in TID2013 and CSIQ databases, based on the idea of SSIM, combining with the image contrast distortion features and CS characteristics of HVS, an IQA model based on SSIM (namely PSIM) is designed in this paper. The methods of PSIM are as follows: First, combined with the HVS characteristics and its model, a perception model of HVS perceiving real images is proposed. Then, the perception model is used to simulate HVS to perceive the real image, to eliminate the visual redundancy of the real image and the perceived image is obtained. Finally, based on the idea and form of the SSIM model, combining with perceived image features, a novel IQA model, namely PSIM, is built. Moreover, to illustrate the performance of the proposed IQA model, 5335 distorted images in four databases (TID2013, CSIQ, LIVE and CID) are used for simulation. The IQA effect of each distortion type of images of 41 distortion types in four databases and the overall IQA effect of all images in each database are analyzed, and compared them with that of 12 existing IQA models. The results show that in comparison with SSIM and MSSIM, the IQA scores of PSIM is more consistent with the subjective IQA scores, and the accuracy of the model has been significantly improved. In terms of the comprehensive benefits of accuracy, generalization performance and complexity, PSIM is better than the 12 existing models. The comprehensive results show that PSIM is an effective and excellent IQA model.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. M. Perez, A. Mikhailiuk, E. Zerman, V. Hulusic, G. Valenzise, R. K. Mantiuk, From pairwise comparisons and rating to a unified quality scale, *IEEE Trans. Image Process.*, **29** (2020), 1139–1151. https://doi.org/10.1109/TIP.2019.2936103

2. X. F. Zhang, W. S. Lin, Q. M. Huang, Fine-grained image quality assessment: a revisit and further thinking, *IEEE Trans. Circuits Syst. Video Technol.*, **32** (2022), 2746–2759. https://doi.org/10.1109/TCSVT.2021.3096528

3. Y. M. Fang, R. G. Du, Y. F. Zuo, W. Y. Wen, L. D. Li, Perceptual quality assessment for screen content images by spatial continuity, *IEEE Trans. Circuits Syst. Video Technol.*, **30** (2020), 4050–4063. https://doi.org/10.1109/TCSVT.2019.2951747

4. G. T. Zhai, X. K. Min, Perceptual image quality assessment: a survey, *Sci. China Inf. Sci.*, **63** (2020), 84–135. https://doi.org/10.1007/s11432-019-2757-1

5. X. Y. Huang, L. J. He, Playback experience driven cross layer optimization of APP, TCP and MAC layer for video clients over LTE system, *IET Commun.*, **14** (2020), 2176–2188. https://doi.org/10.1049/iet-com.2019.0645

6. D. M. Chandler, S. S. Hemami, VSNR: A wavelet-based visual signal-to-noise ratio for natural images, *IEEE Trans. Image Process.*, **16** (2007), 2284–2298. https://doi.org/10.1109/TIP.2007.901820

7. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.*, **13** (2004), 600–612. https://doi.org/10.1109/TIP.2003.819861

8. L. Zhang, L. Zhang, X. Q. Mou, D. Zhang, FSIM: a feature similarity index for image quality assessment, *IEEE Trans. Image Process.*, **20** (2011), 2378–2386. https://doi.org/10.1109/TIP.2011.2109730

9. W. F. Xue, L. Zhang, X. Q. Mou, A. C. Bovik, Gradient magnitude similarity deviation: a highly efficient perceptual image quality index, *IEEE Trans. Image Process.*, **23** (2014), 684–695. https://doi.org/10.1109/tip.2013.2293423

10. L. Zhang, Y. Shen, H. Li, VSI: a visual saliency-induced index for perceptual image quality assessment, *IEEE Trans. Image Process.*, **23** (2014), 4270–4281. https://doi.org/10.1109/TIP.2014.2346028

11. Z. Wang, Q. Li, Information content weighting for perceptual image quality assessment, *IEEE Trans. Image Process.*, **20** (2011), 1185–1198. https://doi.org/10.1109/TIP.2010.2092435

12. J. C. Yao, G. Z. Liu, Improved SSIM IQA of contrast distortion based on the contrast sensitivity characteristics of HVS, *IET Image Process.*, **12** (2018), 872–879. https://doi.org/10.1049/iet-ipr.2017.0209

13. P. Cheraaqee, Z. Maviz, A. Mansouri, A. Mahmoudi-Aznaveh, Quality assessment of screen content images in wavelet domain, *IEEE Trans. Circuits Syst. Video Technol.*, **32** (2022), 566–578. https://doi.org/10.1109/TCSVT.2021.3067627

14. Laboratory for Image & Video Engineering, LIVE image and video quality assessment database, The University of Texas at Austin, 2021. Available from: http://live.ece.utexas.edu/research/quality.

15. Tampere University of Technology, Tampere image database 2013 (TID2013), version 1.0, Tampere University of Technology, 2021. Available from: http://www. ponomarenko.info/tid2013.htm.

16. Oklahoma State University, The CSIQ image database, Oklahoma State University, 2021. Available from: http://vision.okstate.edu/csiq.

17. M. H. Khosravi, H. Hamid, A new paradigm for image quality assessment based on human abstract layers of quality perception, *Multimedia Tools Appl.*, **81** (2022), 23193–23215. https://doi.org/10.1007/s11042-022-12478-y

18. K. Y. Ding, K. D. Ma, S. Q. Wang, E. P. Simoncelli, Image quality assessment: unifying structure and texture similarity, *IEEE Trans. Pattern Anal. Mach. Intell.*, **44** (2022), 2567–2581. https://doi.org/10.1109/TPAMI.2020.3045810

19. X. K. Min, G. T. Zhai, K. Gu, Y. T. Liu, X. K. Yang, Blind image quality estimation via distortion aggravation, *IEEE Trans. Broadcast.*, **64** (2018), 508–517. https://doi.org/10.1109/TBC.2018.2816783

20. Z. Q. Pan, F. Yuan, J. J. Lei, Y. M. Fang, X. Shao, S. Kwong, VCRNet: visual compensation restoration network for no-reference image quality assessment, *IEEE Trans. Image Process.*, **31** (2022), 1613–1627. https://doi.org/10.1109/TIP.2022.3144892

21. X. Y. Ma, S. Y. Zhang, Y. Q. Wang, R. Li, X. D. Chen, D. G. Yu, ASCAM-Former: blind image quality assessment based on adaptive spatial & channel attention merging transformer and image to patch weights sharing, *Expert Syst. Appl.*, **215** (2023), 119268. https://doi.org/10.1016/j.eswa.2022.119268

22. Q. Yang, Z. Ma, Y. L. Xu, L. Yang, W. J. Zhang, J. Sun, Modeling the screen content image quality via multiscale edge attention similarity, *IEEE Trans. Broadcast.*, **66** (2020), 310–321. https://doi.org/10.1109/TBC.2019.2954063

23. L. J. He, G. Z. Liu, X. Y. Huang, Playback continuity and video quality driven optimisation for dynamic adaptive streaming over HTTP clients over wireless networks, *IET Commun.*, **12** (2018), 1178–1187. https://doi.org/10.1049/iet-com.2017.0522

24. P. K. Podder, M. Paul, M. Murshed, EMAN: The human visual feature based no-reference subjective quality metric, *IEEE Access*, **7** (2019), 46152–46164. https://doi.org/10.1109/ACCESS.2019.2904732

25. A. Ahar, A. Barri, P. Schelkens, From sparse coding significance to perceptual quality: a new approach for image quality assessment, *IEEE Trans. Image Process.*, **27** (2018), 879–893. https://doi.org/10.1109/TIP.2017.2771412

26. L. Zheng, L. Shen, J. Chen, P. An, J. Luo, No reference quality assessment for screen content images based on hybrid region features fusion. *IEEE Trans. Multimedia*, **21** (2019), 2057–2070. https://doi.org/10.1109/TMM.2019.2894939

27. S. Mahmoudpour, P. Schelkens, Synthesized view quality assessment using feature matching and superpixel difference, *IEEE Signal Process. Lett.*, **27** (2020), 1650–1654. https://doi.org/10.1109/LSP.2020.3024109

28. Y. Jiang, X. Yan, J. Chen, J. Cheng, J. Zhang, Meaningful secret image sharing for JPEG images with arbitrary quality factors, *Math. Biosci. Eng.*, **19** (2022), 11544–11562. https://doi.org/10.3934/ mbe.2022538

29. S. Westland, H. Owens, V. Cheung, Model of luminance contrast-sensitivity function for application to image assessment, *Color Res. Appl.*, **31** (2006), 315–319. https://doi.org/10.1002/col.20230

30. M. J. Nadenau, J. Reichel, M. Kunt, Wavelet-based color image compression: exploiting the contrast sensitivity function, *IEEE Trans. Image Process.*, **12** (2003), 58–70. https://doi.org/10.1109/TIP.2002.807358

31. Y. Z. Niu, H. F. Zhang, W. Z. Guo, R. R. Ji, Image quality assessment for color correction based on color contrast similarity and color value difference, *IEEE Trans. Circuits Syst. Video Technol.*, **28** (2018), 849–862. https://doi.org/10.1109/TCSVT.2016.2634590

32. J. C. Yao, G. Z. Liu, Bitrate-based no-reference video quality assessment combining the visual perception of video contents, *IEEE Trans. Broadcast.*, **65** (2019), 546–557. https://doi.org/10.1109/ TBC.2018.2878360

33. S. Cheng, H. Q. Zeng, J. Chen, J. H. Hou, J. Q. Zhu, K. K. Ma, Screen content video quality assessment: subjective and objective study, *IEEE Trans. Image Process.*, **29** (2020), 8636–8651. https://doi.org/10.1109/TIP.2020.3018256

34. X. Liu, M. Pedersen, J. Y. Hardeberg, CID:IQ–A new image quality database, in *Image and Signal Processing*, Springer International Publishing, Cherbourg, France, 2014. https://doi.org/10.1007/978-3-319-07998-1_22

35. E. C. Larson, D. M. Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy, *J. Electron. Imaging*, **19** (2010), 011006. https://doi.org/10.1117/1.3267105.

36. Z. Yi, D. M. Chandler, Opinion-unaware blind quality assessment of multiply and singly distorted images via distortion parameter estimation, *IEEE Trans. Image Process.*, **27** (2018), 5433–5448. https://doi.org/10.1109/TIP.2018.2857413

37. L. H. Chen, C. G. Bampis, Z. Li, J. Sole, A. C. Bovik, Perceptual video quality prediction emphasizing chroma distortions, *IEEE Trans. Image Process.*, **30** (2021), 1408–1422. https://doi.org/10.1109/TIP.2020.3043127

38. C. Zhang, W. Cheng, K. Hirakawa, Corrupted reference image quality assessment of denoised images, *IEEE Trans. Image Process.*, **28** (2019), 1732–1747. https://doi.org/10.1109/TIP.2018.2878326