International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.11 (2015), pp.179-194
http://dx.doi.org/10.14257/ijsip.2015.8.11.17

A Fast Feature Similarity Index for Image Quality Assessment

Shaoping Xu[1], Xiaoping Liu[1, 2] and Shunliang Jiang[1]

[1]*School of Information Engineering, Nanchang University, NanChang, JiangXi, China, 330031*
[2]*Department of Systems and Computer Engineering, Carleton University, Ottawa, ON Canada, K1S 5B6*
[2]*xushaoping@ncu.edu.cn*

Abstract

The main drawback of the phase congruency feature employed in the feature similarity index (FSIM) image quality assessment (IQA) algorithm is its low computational efficiency. In this paper, a novel fast feature similarity index (FFSIM) for image quality assessment is proposed. Based on the fact that human visual system (HVS) responds to the brightness stimulus mainly complying with Weber's law, the proposed FFSIM only performs spatial filtering to quickly calculate the contrast between the current pixel and its background, which is used to compute Weber visual salience similarity and a weighting coefficient in pooling stage after applied nonlinear mapping. Weber contrast and the gradient magnitude play complementary roles in characterizing the image local quality. After obtaining the local quality map, we use Weber weighting coefficient again as a weighting coefficient to derive a single quality score. As such, the multi-scale version of the FFSIM algorithm, i.e., MS-FFSIM is also proposed, which complies with the spatial frequency response characteristics of the HVS system. Extensive experiments performed on six publicly available IQA databases demonstrate that the proposed FFSIM and MS-FFSIM can achieve higher consistency with the subjective evaluations than state-of-the-art IQA metrics and the computational efficiency is greatly improved as well.

Keywords: *Image Quality Assessment; Phase Congruency; Weber Law; Weighted Coefficient; Gradient Structural similarity; Computational Efficiency*

1. Introduction

Digital images are usually distorted by a wide variety of contaminations during acquisition, compression, transmission or storage, decoding, and display, any of which generally could result in a degradation of visual quality [1-5]. Since the images are ultimately to be viewed by human visual system (HVS), the best method of quantifying visual image quality is through subjective evaluation. However, subjective evaluation is usually time-consuming and impractical in real world applications. Therefore, there has been an increasing push to develop objective measurement approaches that predict image quality automatically. According to the availability of a reference image, the objective IQA algorithms generally fall into three categories: full-reference (FR) [6], reduced-reference (RR) and no-reference (NR) algorithms [3]. These three classes are required at different situations. Although NR-IQA is potentially the most useful goal, the difficulty of creating algorithms that accurately predict visual quality, especially without any information about the original image, still makes it attractive to develop FR-IQA algorithms in practical applications. We focus on FR-IQA in this paper.

Commonly, many state-of-the-art FR-IQA algorithms adopted two-stage framework, viz. local quality measurement stage and pooling stage. The local quality measurement process typically results in a quality map defined either in the spatial domain or in the transform domain. To date, significant progress has been made in the design of the local

ISSN: 2005-4254 IJSIP

quality measurement. For example, the well-known structural similarity (SSIM) [7] is widely accepted due to its reasonably good evaluation accuracy, pixel-wise quality measurement, and simple mathematical formulation. The multi-scale extension of the SSIM, namely MS-SSIM [8], produces better results than its single-scale counterpart. The SSIM and its extensions take advantage of the features, *i.e.*, luminance, contrast and structure (covariance), with a certain degrees of success. Besides, the image features including visual information [9], phase congruency (PC) [10-11], singular value decomposition [12], Harris response [13], gradient [14], *etc* were also employed in developed FR-IQA algorithms. It should be noted that, among these features, gradient feature that conveys important visual information was often utilized to capture local distortion in the image due to relatively low computational complexity. Actually, recent works image quality assessment applied image gradient feature for a variety of FR-IQAs and showed good results in the experiments. For instance, Chen [14] proposed a so-called gradient-based structural similarity (G-SSIM), which compared edge information between the distorted image and the original image. Experimental results show that the G-SSIM algorithm correlates with the perception of visual quality much better than the SSIM and MS-SSIM. In [15], Zhu *et. al.*, presented a multi-scale visual gradient similarity index (VGS) incorporating HVS low-level visual properties, such as visual detection threshold of gradient and visual frequency sensitivity. Experimental results show that VGS is competitive with state-of-the-art algorithms in terms of prediction precision and reliability. However, the main limitations of the VGS are caused by the fixed visible gradient threshold and the optimal values of two main parameters, which need to be trained with existing IQA databases. Thus, the VGS is image-dependent and would be inconvenient in practical applications. In [16], Liu *et. al.*, proposed a new IQA algorithm based on the concept of gradient similarity. They had demonstrated that the gradient similarity measure considering masking effect and visibility threshold can achieve a similar or better performance compared with eight other representative and prominent IQA algorithms. Nevertheless, the performance of the algorithm is heavily affected by parameter values.

While significant progress has been made in the design of the local quality measurement stage in the literature, pooling stage still has not been well explored to convert quality map into a single quality score and has a long way to go before reaching useful levels of performance. For instance, many classical algorithms such as SSIM and MS-SSIM share a common deficiency that, when pooling a single quality score from the local quality map, all positions are considered to have the same importance. In other words, they treat different located distortions equally. Neither SSIM nor MS-SSIM takes into account factors such as the visual importance of image features. In [17], images were decomposed in different sub-bands and these sub-bands can have different weights at the pooling stage. However, within each sub-band, every position was still given the same importance. Such pooling strategies were not consistent with the intuition that different locations on an image can have very different contributions to human visual perception of the image. In [18], Wang *et. al.*, proposed a novel information content weighted SSIM IQA algorithm (IW-SSIM). When pooling the quality map into a quality score, instead of averaging, the IW-SSIM used information content as weights. The success of the IW-SSIM algorithm contributes to an effective combination of local SSIM measurement and multi-scale image decomposition followed by scale-variant weighting. Although it achieves good correlation with the human judgment, it has higher computational complexity. Similarly, motivated by the idea that different image regions (edges, textures and smooth regions) have different perceptual significance relative to quality, Li and Bovik [19] presented a four-component weighted structural similarity index and experimental results show that this kind of region-weighted approach can improve the performance of SSIM, MS-SSIM, and some of their popular variants.

Recently, Zhang *et. al.*, [10] proposed a so-called feature similarity (FSIM) index based on advance local image quality measurement and saliency-based weighting. Specifically, the FSIM algorithm used two kinds of features to measure local image quality, *i.e.*, the phase congruency (PC) and the gradient magnitude (GM), which represent complementary aspects of the visual image quality. Note that PC in FSIM index also provides a saliency measure of image gradient feature, which is used to weight the contribution of each pixel to the overall quality score (We use saliency as a general term that represents low-level local image features that are of perceptual significance). Similar to IW-SSIM, its main drawback is low computational efficiency that makes FSIM not appropriate to real-time applications [11]. In this paper, we focus on reducing the computational complexity of the FSIM, and propose a novel low complexity version of the FSIM, called fast feature similarity index (FFSIM) incorporating gradient visual saliency. First, we use the gradient structural similarity [14] to measure the changes in structure in images. Furthermore, based on the fact that human visual system (HVS) responds to the brightness stimulus mainly complying with Weber's law, the proposed algorithm only performs one pass filtering to quickly calculate the contrast between the current pixel and its background, which is used as a dimensionless measure of the visual significance of the gradient magnitude after applying nonlinear mapping. The gradient magnitude combined with Weber contrast (WC) visual significance, which play similar role as the PC in FSIM algorithm but with significantly higher computational efficiency, serves basis for characterizing the local image quality. Finally, the multi-scale extension of the FFSIM (MS-FFSIM) is proposed, which complies with the spatial frequency response characteristics of the HVS system and has better prediction performance than single scale FFSIM. Extensive experiments conducted with six publicly available subject-rated databases have confirmed the effectiveness, robustness, and efficiency of the proposed algorithm in comparison with the relevant state-of-the-art algorithms.

The remainder of this paper is organized as follows. In Section 2, we provide a brief introduction of the FSIM, as the ground for the following analysis, discussion, and comparison. Section 3 describes the details of the Weber's law and discusses how to combine the Weber contrast and gradient magnitude changes. The experimental results with further discussion are given in Section 4. We will validate our FFSIM and MS-FFSIM algorithms with six publicly available IQA databases for benchmarking. Finally, conclusions are drawn in Section 5.

2. FSIM

For the sake of brevity, this section only presents a brief overview of the FSIM algorithm. For detailed information, readers are referred to [10]. It is found that visually discernable features coincide with those points where the Fourier waves at different frequencies have congruent phases [11]. Highly informative phase congruency (PC) feature plays an important role in FSIM algorithm, thus we start from the definition of PC. Let M_n^e and M_n^o denote the even-symmetric and odd-symmetric wavelet at scale n, respectively, and they form a quadrature pair. The convolution results of the input image I with quadrature pairs of filters at position $I(x, y)$ on scale n will form a response vector that is the basic components to calculate PC:

$$[e_n(x, y), o_n(x, y)] = [I(x, y) * M_n^e, I(x, y) * M_n^o] \quad (1)$$

With respect to the quadrature pair of filters, *i.e.*, M_n^e and M_n^o, FSIM adopts log-Gabor filters due to the following two reasons. First, log-Gabor filters allow arbitrarily bandwidth and maintain a reasonably small DC component in the even-symmetric filter. Second, the transfer function of the log-Gabor filter has an extended tail at the high-

frequency end, which makes it more capable to encode natural images than ordinary Gabor filters. Specifically, the log-Gabor function has the following transfer function:

$$LogGabor(\omega,\theta_j) = \exp(-\frac{(\log(\omega/\omega_0))^2}{2\sigma_r^2}) \cdot \exp(-\frac{(\theta-\theta_j)^2}{2\sigma_\theta^2}) \quad (2)$$

where n is scale, $\theta_j = j\pi / J$, $j = \{0,1,\cdots,J-1\}$ is the orientation angle of the filter, J is the number of orientations, σ_θ determines the filter's angular bandwidth, ω_0 is the filter's center frequency, and σ_r controls the filter's bandwidth.

As such, the 2D PC at position $I(x,y)$ can be expressed as the summation over orientation θ_j and scale n

$$PC(x,y) = \frac{\sum_j E_{\theta_j}(x,y)}{\varepsilon + \sum_n \sum_j A_{n,\theta_j}(x,y)} \quad (3)$$

where $A_{n,\theta_j}(x,y) = \sqrt{e_{n,\theta_j}(x,y)^2 + o_{n,\theta_j}(x,y)^2}$ is the local amplitude on scale n and orientation θ_j, and $E_{\theta_j}(x,y) = \sqrt{F_{\theta_j}^{\ 2}(x,y) + H_{\theta_j}^{\ 2}(x,y)}$ is local energy along orientation θ_j, where $F_{\theta_j}(x,y) = \sum_n e_{n,\theta_j}(x,y)$ and $H_{\theta_j}(x,y) = \sum_n o_{n,\theta_j}(x,y)$. It should be noted that $PC(x,y)$ is a real number within (0-1].

The FSIM local quality measurement at position $I(x,y)$ between reference image R and distorted D is defined as:

$$FSIM(x,y) = S_{GM}(x,y) \cdot S_{PC}(x,y) \quad (4)$$

where $S_{GM}(x,y), S_{PC}(x,y)$ are GM similarity measure and PC similarity measure, respectively. The similarity measure for $G_R(x,y)$ and $G_D(x,y)$ is defined as follows:

$$S_{GM}(x,y) = \frac{2G_R(x,y)\cdot G_d(x,y)+T_1}{G_R^2(x,y)+G_R^2(x,y)+T_1} \quad (5)$$

where T_1 is a positive constant depending on the dynamic range of GM values. Eq. (5) is a commonly used measure to define the similarity of two positive real numbers and its result ranges within (0-1]. Similarly, the similarity measure for $PC_R(x,y)$ and $PC_D(x,y)$ is defined as follows:

$$S_{PC}(x,y) = \frac{2PC_R(x,y)\cdot PC_d(x,y)+T_2}{PC_R^2(x,y)+PC_R^2(x,y)+T_2} \quad (6)$$

where T_2 is a positive constant to increase the stability of $S_{PC}(x,y)$ (such a consideration was also included in SSIM). The determination of T_2 depends on the dynamic range of PC values. Eq. (6) is a commonly used measure to define the similarity of two positive real numbers and its result ranges within (0-1].

Having obtained the similarity $FSIM(x,y)$ at each location $I(x,y)$, the overall similarity between R and D can be calculated. After computing the local similarity map, PC is utilized again as a weighting function to derive a single similarity score. As we know, human visual cortex is sensitive to phase congruent structures, the PC value at a location can reflect how likely it is a perceptibly significant structure point. Intuitively, for a given location $I(x,y)$, if anyone of $R(x,y)$ and $D(x,y)$ has a significant PC

value, it implies that this position $I(x, y)$ will have a high impact on HVS in evaluating the similarity between R and D. Therefore, $PC_m(x, y) = max(PC_R(x, y), PC_D(x, y))$ was used to weight the importance of $FSIM(x, y)$ in the overall similarity between R and D. Accordingly, the FSIM index between R and D is defined as follows [10]:

$$FSIM = \frac{\sum_{(x,y)\in\Omega} FSIM(x, y) \cdot PC_m(x, y)}{\sum_{(x,y)\in\Omega} PC_m(x, y)} \tag{7}$$

where Ω means the whole image spatial domain.

3. Fast FSIM

3.1. Basic Idea

The main limitation of the PC is its low computational efficiency since $n * j$ filtering operations must be conducted when computing visual saliency for the pixel $I(x, y)$. The novel perceptual local image quality measurement is motivated by the fact that perceptual image gradient degradations are not only affected by spatial distributions of pixels but also by background luminance, which is ignored in many IQA algorithms and should be also accounted for a complete and more robust IQA algorithm. In fact, the visual perception to image gradient magnitude is nonlinear, which can be seen from Figure 1. When one gradient magnitude is twice that of another, we can not perceive double intensity from the former. Without loss of generality, in this work the subjective (or perceptual) gradient intensity with raw gradient magnitude is represented by $f(\cdot)$, which will be detailed in the following sections.

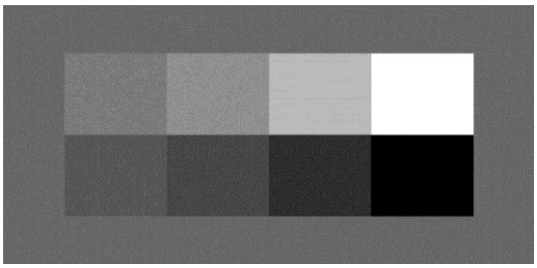


Figure 1. Illustration for Perceptual Intensity of Gradient Magnitude

The back ground gray value is 128 with an 8-bit gray scale. From left to right, the gray values are 128 + [16, 32, 64, 127] for the rectangles in the top row and 128-[16, 32, 64, 128] for the rectangles in the bottom row.

3.2. Weber's Law

It is well known that human vision has a nonlinear perception to different physical stimuli (*e.g.*, luminance), which has been theorized and empirically proven by the psychologist E.H. Weber [20]. In psychophysics, let X and Y denote the raw magnitude and the perceptual intensity of physical stimuli respectively, then we roughly have:

$$\Delta Y = f(\frac{\Delta X}{X_0}) \tag{8}$$

where ΔX represents the increment threshold (just noticeable difference for discrimination), X_0 represents the initial stimulus intensity, $\frac{\Delta X}{X_0}$ is the relative change in

luminance, *i.e.*, Weber's contrast (WC), ΔY is the change in subjective brightness, and f is a mapping function, which should match well a human being's perception.

3.3. Weber Contrast Visual Saliency

The observer's assessment of image quality is prejudiced by the perceived structural distortions in saliency regions. Therefore, a relative measure of the importance of different regions indicated by a saliency map plays an important role in evaluating the image quality. In this paper, we employ Weber's law model to detect the saliency. Weber's Law is computed empirically, its effectiveness had been proven for several application fields. We assume that Weber's law is also suitable for image gradient magnitude, which will be empirically verified later by testing subject-rated IQA databases. In the following, perception nonlinearity of image gradient magnitude is explored and taken into consideration. We use the intensity differences $\xi(x_c)$ between a current pixel x_c and its neighbors as the visual salience of the current pixel $f(\xi(x_c))$. By this means, we hope to find the saliency variations within an image to simulate the gradient magnitude perception of human beings. Specifically, the Weber contrast visual saliency $f(\xi(x_c))$ of a current pixel x_c is computed as:

$$f(\xi(x_c)) = f(\frac{\Delta X}{X}) = f(\sum_{i=0}^{N-1}(\Delta x_i)/x_c = \frac{\sum_{i=0}^{N-1}(x_i - x_c)}{x_c}) \quad (9)$$

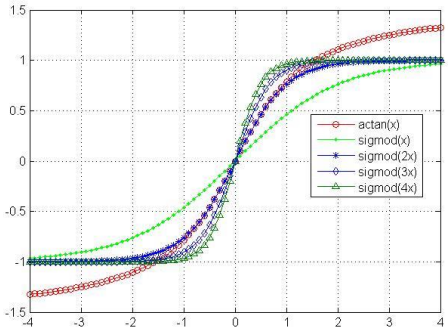


Figure 2. Comparison of the Arctangent Function and Some Sigmoid Functions

Note that the output of arctan () is in radian measure.

It should be noted that, if $\xi(x_c)$ is positive, it simulates the case that the surroundings are lighter than the current pixel. In contrast; if $\xi(x_c)$ is negative, it simulates the case that the surroundings are darker than the current pixel. As we known, logarithm function matches well a human being's perception. However, it can not be used here since many outputs of $\xi(x_c)$ are negative. As shown in Figure 2, we use the arctangent function to compute $f(\xi(x_c))$. We use this function since it can limit the output to prevent it from increasing or decreasing too quickly when the input becomes larger or smaller. By this means, we attempt to preserve more discriminating information in comparison to using the absolute value of $\xi(x_c)$. As such, Weber contrast visual salience similarity is defined as:

$$S_{WC}(x, y) = \frac{2 f_R(x,y) \cdot f_D(x,y) + T_3}{f_R^2(x,y) + {}^2_D(x,y) + T_3} \quad (10)$$

It is well known that edges are crucial for visual perception and play a major role in the recognition of image content. Intensity edges certainly contain considerable image information and are perceptually significant. As shown in Figure 3, by visual examination, we can see that PC map and WC map of the reference image are substantially similar, perceptual differences only exist in the areas of clouds and shadows where have a slight impact on the overall image quality. Figure 3 clearly illustrates that WC play similar role as the PC in FSIM algorithm and is alternative to PC. More statistically convincing results will be presented in the following sections.

a) b) c)

Figure 3. Comparison of the PC map and WC Map

(a) is the reference image; (b) is the PC map of (a); (c) is the Weber visual salience map of (a).

3.4. FFSIM and its Multi-Scale Extension

In our proposed algorithm, $S_{WC}(x, y)$ is used to replace the FSIM local image quality measurement $S_{PC}(x, y)$ to further improve its computational efficiency, which is appealing for visual prediction of human beings.

$$FFSIM = \frac{\sum_{\Omega} S_{GM}(x,y) \cdot S_{WC}(x,y) WC_m(x,y)}{\sum_{\Omega} WC_m(x,y)} \tag{11}$$

where $WC_m(x, y) = \max(WC_R(x, y), WC_D(x, y))$.

Images are naturally multi-scale and image features possess multi-scale attributes. In MS-SSIM, quality assessment is accomplished over multiple scales of the reference and distorted image patches by iteratively low-pass filtering and down-sampling the signals. Similar to MS-SSIM, a multi-scale FFSIM (MS-FFSIM) algorithm is obtained by combining the measurement across scales via:

$$MS - FFSIM = \prod_{j=1}^{M} (FFSIM_j)^{B_j} \tag{12}$$

where M is scale, the values B_j are relative weights between scales that can be obtained through psychophysical measurement [8]. Interestingly, the measured weight function peaks at middle-resolution scales and drops at both low- and high-resolution scales, consistent with the contrast sensitivity function (CSF) extensively studied in the vision literature [21].

4. Experiments and Results

4.1. Databases and Criteria for Comparison

As summarized in Table 1, there are six publicly available human-rated image databases, including LIVE, TID2008, CSIQ, IVC, A57 and MICT [18], which are widely recognized in the IQA research community. We include all of them in our algorithm validation and comparisons. The performance of the proposed FFSIM and MS-FFSIM algorithms will be evaluated and compared with seven representative FR-IQA metrics, including six state-of-the-arts (SSIM, MS-SSIM, IFC, VSNR, NQM, and FSIM) and the classical PSNR. For FFSIM and MS-FFSIM, we implemented it by ourselves. For FSIM, we used the implementation provided by the author, which is available at http://www.comp.polyu.edu.hk/~cslzhang/IQA/FSIM/FSIM.htm. For all the other methods evaluated, we used the public software MeTriXMuX, which is available at http://foulard.ece.cornell.edu/gaubatz/metrix_mux/.

Table 1. Benchmark Test Databases for IQA

Database	Source Images	Distorted Images	Distortion Types	DMOS Rang	Provider
LIVE	29	779	5	0-100	University of Texas
TID2008	25	1700	17	0-9	Joint international effort between Finland、Italy and Ukraine
CSIQ	30	866	6	0-1	Oklahoma University
IVC	10	185	5	0-5	Ecole Polytechnique University
A57	3	54	6	0-1	Cornell University
MICT	14	168	2	1-5	Toyama University

We used three evaluation criteria to compare the performance of the FR-IQA algorithms, *i.e.*, Spearman rank-order correlation coefficient (SROCC), Pearson linear correlation coefficient (PLCC) and Kendall rank-order correlation coefficient (KROCC) between the objective scores after nonlinear regression and the subject scores. Among these criteria, SROCC and KROCC are used to assess prediction monotonicity. They are independent of any monotonic nonlinear mapping and ignore the relative distance between data points. PLCC is adopted to evaluate prediction accuracy for a perfect match between the mapped objective scores and the subjective scores. A better objective FR-IQA algorithm has a higher PLCC, SROCC, and KROCC. The mathematical definition of these three criteria is given as follows [18].

SRCC is defined as:

$$SRCC = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2-1)} \tag{13}$$

where d_i is the difference between the i th image's ranks in subjective and objective evaluations. SRCC is a non-parametric rank-based correlation metric.

KRCC is another non-parametric rank correlation metric given by:

$$KRCC = \frac{N_c - N_d}{0.5N(N-1)} \tag{14}$$

where N_c and N_d are the numbers of concordant and discordant pairs in the data set, respectively.

PLCC can then be computed as:

$$PLCC = \frac{\sum_i (q_i - \bar{q}) * (o_i - \bar{o})}{\sqrt{\sum_i (q_i - \bar{q})^2 * \sum_i (o_i - \bar{o})^2}} \quad (15)$$

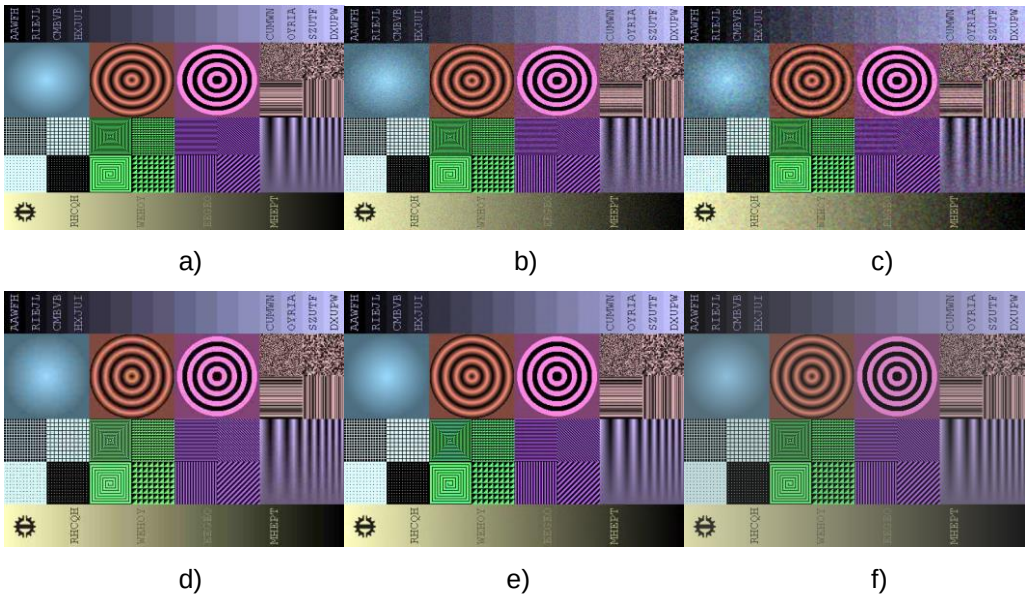
where o_i is subjective score of the i th image in an image database, q_i is the mapping score of the raw objective score r_i using a nonlinear mapping function. We used the following logistic mapping function:

$$q(r) = a_1 (0.5 - \frac{1}{1 + \exp(a_2 (r - a_3))}) + a_4 r + a_5 \quad (16)$$

where a_1 to a_5 are model parameters found numerically using a nonlinear regression process to maximize the correlations between subjective and objective scores.

4.2. Validation and Comparison

In this subsection, we first use an example to demonstrate the effectiveness of FFSIM in evaluating the perceptible image quality. Figure 4 (a) is the I25 reference image in the TID2008 database, and Figures 4(b-f) show five distorted images of I25. Distortion types of Figures 4 (b-f) is additive Gaussian noise, spatially correlated noise, JPEG compression, JPEG 2000 compression, and contrast change, respectively. We compute the image quality of Figures 4 (b-f) using various FR-IQA algorithms, and the results are summarized in Table 2. We also list the subjective scores (extracted from TID2008) of these five images. For each FR-IQA algorithm and the subjective evaluation, higher scores mean higher image quality. In order to show the correlation of each FR-IQA algorithm with the subjective evaluation more clearly, in Table 2, we list SRCC value at the last row. From Table 2, we can see that the quality scores computed by FFSIM and MS-FFSIM correlate with the subjective evaluation much better than the other IQA algorithms. Similar results can also be found in Table 3, where the I04 reference image and its five distorted images are employed as test data.



a) b) c)

d) e) f)

Figure 4. Evaluation of I25 Image Contaminated by Different Distortions

(a) Reference image; (b-f) are the distorted versions of (a) in the TID2008 database. Distortion types of (b-f) are additive Gaussian noise, spatially correlated noise, JPEG compression, JPEG 2000 compression, and contrast change, respectively.

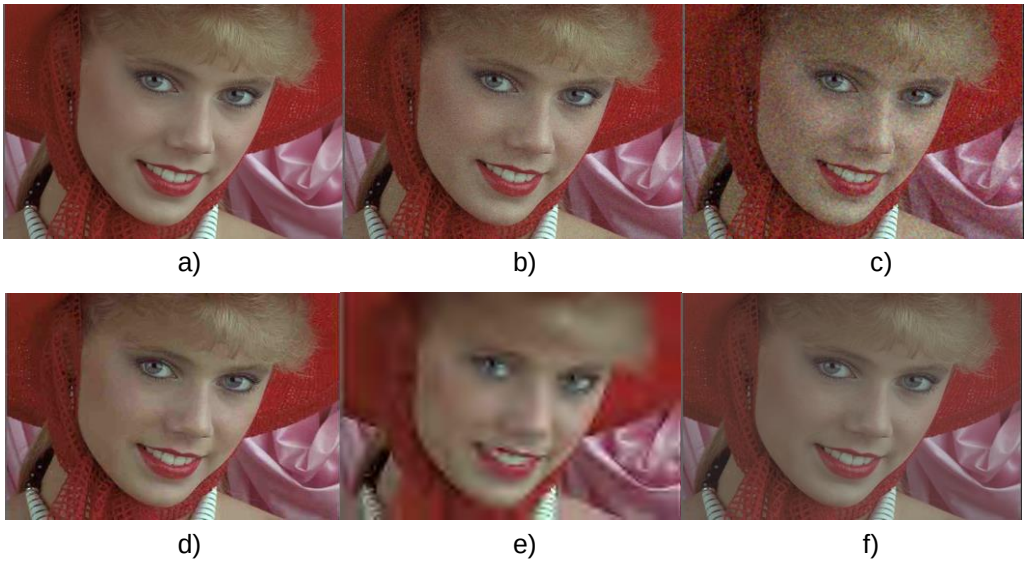


Figure 5. Evaluation of I04 Image Contaminated by Different Distortions

(a) Reference image; (b-f) are the distorted versions of (a) in the TID2008 database. Distortion types of (b-f) are additive Gaussian noise, spatially correlated noise, JPEG compression, JPEG 2000 compression, and contrast change, respectively.

To further examine the robustness of the FR-IQA algorithms, the performance on each distortion type in TID2008 database is shown in Table 4, where the three best FR-IQA algorithms have been highlighted in boldface and underlined numbers represent the best algorithm for each distortion type. The TID2008 database is used since it is the largest database and contains the most distortion types, which are listed in the first column of Table 4. We include only the SROCC values since other performance criteria lead to similar conclusions. From Table 4, we can see that the proposed algorithms, *i.e.*, the FFSIM and MS-FFSIM, perform quite well (*i.e.*, always the best or the second best, especially for Gaussian noise and JPEG compression that are the most common types of distortions). More importantly, MS-FFSIM and FFSIM outperform FSIM, which suggests that they perform much better than the other IQA algorithms, such as VIF, IFC and NQM. We did not list their data due to the space restrictions of the paper. Results in Table 4 corroborate that the Weber contrast visual saliency does affect the perceptible quality and lead to consistent improvement.

Table 2. Quality Evaluation of Images in Figure 4

Figures	MOS	IQA algorithm								
		FSIM	SSIM	MSSIM	IFC	VSNR	NQM	PSNR	FFSIM	MS-FFSIM
Figure 4(a)	4.3600	0.8931	0.7524	0.9506	5.4302	29.5359	26.6275	27.9629	0.7568	0.9562
Figure 4(b)	3.6400	0.9189	0.8498	0.9373	5.1535	25.8149	23.8981	30.9293	0.8252	0.9335
Figure 4(c)	4.8077	0.8937	0.9410	0.9830	6.6103	28.8883	23.6474	24.9560	0.8813	0.9770
Figure 4(d)	6.0769	0.9492	0.9721	0.9964	7.4930	38.9113	30.0237	29.3434	0.9354	0.9927
Figure 4(f)	5.0000	0.9695	0.9390	0.9691	31.8997	19.7930	13.1761	22.9439	0.9813	0.9730
SRCC	—	0.6000	0.8000	0.9000	0.9000	0.3000	0.1000	0.4000	0.8000	0.9000

Table 3. Quality Evaluation of Images in Figure 5

Figures	MOS	IQA algorithm								
		FSIM	SSIM	MSSIM	IFC	VSNR	NQM	PSNR	FFSIM	MS-FFSIM
Figure 5(a)	4.9429	0.8765	0.6866	0.9462	2.7659	25.2771	25.7385	30.4824	0.8535	0.9745
Figure 5(b)	4.6471	0.8444	0.7352	0.8785	1.7336	18.1678	17.4913	30.4551	0.8495	0.9283
Figure 5(c)	5.2857	0.9202	0.8472	0.9659	2.3451	25.6631	25.1103	33.7854	0.9101	0.9803
Figure 5(d)	2.8286	0.7484	0.7183	0.8677	0.5156	16.2331	13.7279	28.2144	0.8278	0.9190
Figure 5(f)	5.1765	0.9824	0.9748	0.9743	13.5516	17.4520	12.5297	29.4388	0.9856	0.9801
SRCC	—	0.9000	0.6000	0.9000	0.7000	0.7000	0.2000	0.7000	0.9000	1.0

Table 4. SRCC Comparisons for Individual Distortion Type on TID2008 Database

Distortion Number	IQA algorithm						
	SSIM	MSSIM	VSNR	IFC	FSIM	FFSIM	MS-FFSIM
1	0.7965	0.8087	0.7734	0.6166	**0.8570**	**0.8866**	**0.8885**
2	0.8075	0.8045	0.7798	0.5460	**0.8515**	**0.8697**	**0.8691**
3	0.8261	0.8198	0.7673	0.6000	**0.9272**	**0.8796**	**0.886**
4	**0.8120**	**0.8151**	0.7348	0.6737	**0.8023**	0.7875	0.7856
5	0.8438	0.8722	0.8822	0.7317	**0.9093**	**0.9126**	**0.9102**
6	**0.7470**	0.6868	0.6438	0.5395	0.7456	**0.7804**	**0.7806**
7	0.8016	0.8530	0.8255	0.6871	**0.8555**	**0.8614**	**0.8819**
8	0.9386	**0.9606**	0.9323	0.8559	**0.9472**	**0.9557**	0.9471
9	0.9272	0.9566	0.9302	0.7973	**0.9604**	**0.9601**	**0.9708**
10	0.8989	0.9347	0.9181	0.8180	**0.9372**	**0.9382**	**0.9406**
11	0.8875	**0.9734**	0.9444	0.9437	**0.9775**	0.9705	**0.9807**
12	0.8194	0.8737	**0.8079**	0.7912	0.8708	**0.8861**	**0.8796**
13	0.8461	0.8521	0.7914	0.7301	**0.8542**	**0.8884**	**0.8869**
14	0.6948	**0.7335**	0.5723	**0.8420**	0.7495	0.7312	**0.7592**
15	**0.8851**	0.7602	0.1826	0.6770	0.8488	**0.9021**	**0.8513**
16	**0.7177**	**0.7347**	**0.7475**	0.4250	0.6695	0.5859	0.5718
17	0.4873	0.6394	0.4774	0.1713	**0.6480**	**0.6606**	**0.6450**

Note 1: Additive Gaussian noise; 2: Additive noise in color components 3: Spatially correlated noise 4: Masked noise 5: High frequency noise 6: Impulse noise 7: Quantization noise 8: Gaussian blur 9: Image denoising 10: JPEG compression 11: JPEG2000 compression 12: JPEG transmission errors 13: JPEG2000 transmission errors 14: Non eccentricity pattern noise 15: Local block-wise distortions of different intensity 16: Mean shift (intensity shift) 17: Contrast change.

To examine the robustness and effects of FFSIM and MS-FFSIM, we carried out a comprehensive test on the LIVE, TID2008, CSIQ, IVC, A57, and MICT databases. The results are shown in Table 5 in terms of the SRCC, KRCC and PLCC metrics, respectively. From the experimental results summarized in Table 5, we can see that our

algorithms achieve the best robust results on all the databases. Specifically, it can be observed that FSIM, FFSIM and MS-FFSIM significantly have higher performance than the others for nearly all the six databases. The proposed FFSIM and MS-FFSIM are worse than the FSIM sometimes, however, considering the scales of the databases including the number of images and the number of distortion types, we think that the results obtained on TID2008, CSIQ and LIVE are much more convincing than those obtained on IVC, MICT and A57. Overall speaking, FFSIM and MS-FFSIM achieve the most consistent and robust performance across all the six databases. By contrast, for the other algorithms, they may work well on some databases but fail to provide good results on other databases.

Table 5. Performance Comparison for IQA Algorithms on Six Databases

Database	Index	IQA Algorithm						
		SSIM	MSSIM	VSNR	IFC	FSIM	FFSIM	MS-FFSIM
LIVE	PLCC	0.9593	0.9390	0.9587	0.9620	**0.9463**	**0.9644**	0.9648
	SRCC	0.9536	**0.9677**	0.9607	0.9474	**0.9708**	**0.9684**	0.9670
	KRCC	0.8179	**0.8561**	0.8335	0.8085	**0.8639**	0.8538	**0.8554**
TID2008	PLCC	0.8111	0.8315	0.7552	0.6819	**0.8511**	**0.8590**	**0.8675**
	SRCC	0.8081	0.8282	0.7477	0.6733	**0.8477**	**0.8504**	**0.8491**
	KRCC	0.6199	0.6447	0.5707	0.4894	**0.6674**	**0.6743**	**0.6726**
CSIQ	PLCC	0.8802	0.9002	**0.9185**	0.8315	0.8759	**0.9102**	**0.9048**
	SRCC	0.8875	**0.9558**	0.9169	0.8293	**0.9496**	0.9468	**0.9552**
	KRCC	0.7115	**0.8180**	0.7461	0.6705	**0.8051**	0.7980	**0.8157**
IVC	PLCC	0.7958	0.9264	0.8361	**0.9504**	**0.9554**	**0.9468**	0.9428
	SRCC	0.8007	0.9047	0.8098	**0.9342**	**0.9365**	0.9240	**0.9258**
	KRCC	0.6357	0.7580	0.6497	**0.7948**	**0.8076**	**0.7838**	0.7821
A57	PLCC	0.9182	0.9616	**0.9630**	0.8259	0.9505	**0.9745**	**0.9647**
	SRCC	0.8111	**0.9139**	**0.9528**	0.7445	**0.9167**	0.9056	0.8722
	KRCC	0.6759	0.8055	**0.8889**	0.5926	**0.8148**	**0.8056**	0.7593
MICT	PLCC	0.9211	**0.9507**	0.9302	0.8192	**0.9623**	0.9352	**0.9423**
	SRCC	0.9148	**0.9470**	0.9246	0.8103	**0.9561**	0.9278	**0.9375**
	KRCC	0.7506	**0.8091**	0.7593	0.6007	**0.8189**	0.7709	**0.7894**

To provide an overall indication of the comparative performance of the different algorithms, Table 6 gives the average SROCC, KROCC, and PLCC results over six databases, where the average values are computed, different weights are assigned to the databases depending on the number of distorted images in each database. In Table 6, we can see that the proposed algorithm performs the best or close to the best on average.

Table 6. Average Performance over Six Databases

Index	IQA algorithm						
	SSIM	MSSIM	VSNR	IFC	FSIM	FFSIM	MS-FFSIM
PLCC	0.8638	0.8805	0.8487	0.7952	**0.8872**	**0.9010**	**0.9039**
SRCC	0.8530	0.8897	0.8370	0.7794	**0.9006**	**0.8981**	**0.8991**
KRCC	0.7035	0.7539	0.6981	0.6205	**0.7628**	**0.7611**	**0.7645**

Finally, to compare the computational complexity of different algorithms, we measured the average computation time required to assess an image of size 512*512 (using a computer with Intel(R) Core(TM) i5-2450M processor at 2.5 GHz). Table 7

reports the required time in seconds per image, with all the codes being implemented with MATLAB. From Table 7, we can see that SSIM has the lowest computational cost, whereas FSIM has the highest computational cost among all compared algorithms. It can be seen that MS-FFSIM is slightly slower than MSSIM, but faster than FSIM significantly. To be more precise, the proposed method takes only about 20% of the time taken by FSIM. The savings over FSIM are due to the use of the Weber contrast visual salience rather than the phase congruency that has higher computational complexity. The proposed MS-FFSIM has quite high speed (less than one second per image), that is to say, computational complexity may not be a major concern in most real-world applications.

Table 7. Execution Time (in Seconds Per Image) for Different Algorithms

	IQA algorithm						
	SSIM	MSSIM	VSNR	IFC	FSIM	FFSIM	MS-FFSIM
Time (s)	0.0581	0.1530	0.0396	1.3111	2.0127	0.0651	0.3902

It should be noted that only the luminance components of the images were employed when performing comparative experiments, although the images of the six test databases are color images. Similar to the FSIM, better performance can be achieved if the chrominance information is incorporated in FFSIM and MS-FFSIM. We did not list their data due to the space restrictions of the paper.

5. Conclusions

In this paper we proposed a fast and multi-scale version of feature similarity IQA algorithm, namely FFSIM and MS-FFSIM. The underlying principle of the FFSIM and MS-FFSIM is that HVS perceives an image mainly based on its saliency low-level image feature. Specifically, the local image quality measurement based on the perception nonlinearity (Weber contrast visual saliency) of image gradient magnitude is implemented, which can better characterize local quality and has much lower complexity than phase congruency-based visual salience. Our extensive tests across six publicly-available independent image databases verified that performances of FFSIM and MS-FFSIM are often superior otherwise similar when compared to the other representative or prominent IQA algorithms in terms of correlation between objective measured quality values and subjective observations, validating that it is a very robust IQA algorithm. We believe that our results support the general principle underlying our approach, *i.e.* some HVS properties on image gradient should been explored and incorporated into designing the new IQA algorithm. We expect that our work could give new insights to people who are interested in image quality assessment.

There are a number of issues that deserve further investigation. Specifically, the low-level image feature currently being employed to capture local characteristics of natural images is based upon local gradient magnitude only. Advanced models that consider both gradient magnitude and gradient direction may lead to more accurate local image quality measurement. Moreover, apart from perception nonlinearity, more HVS visual properties, *e.g.*, just noticeable differences, should be considered to make local image quality measurement correlate well with the human perception in the future.

Acknowledgments

This work was supported by the Chinese National 863 Program under Grant 2013AA013804, the National Natural Science Foundation of China under Grant

61163023, 61379018 and 41261091, the Education Department of Jiangxi Province of China under Grant GJJ14141.

References

[1] W. Lin and C. C. J. Kuo, "Perceptual visual quality metrics: A survey", Journal of Visual Communication and Image Representation, vol. 22, no. 4, **(2011)**, pp. 297-312.
[2] H. R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment", IEEE Transactions on Image Processing, vol. 15, no. 11, **(2006)**, pp. 3440-3451.
[3] M. A Saad, A. C. Bovik, and Charrier C., "A DCT statistics-based blind image quality index", IEEE Signal Processing Letters, vol. 17 , no. 6, **(2010)**, pp. 583-586 .
[4] C. Li, A. C. Bovik, and X. Wu, "Blind image quality assessment using a general regression neural network", IEEE Transactions On Neural Networks, vol. 22 , no. 5, **(2011),** pp. 793-799.
[5] X. H. Zhang, W. S. Lin, and P. Xu, "Improved estimation for just-noticeable visual distortion", Signal Processing, vol. 85 , no. 4, **(2005)**, pp.795-808.
[6] X. Fei, L. Xiao, and Y. Sun, "Perceptual image quality assessment based on structural similarity and visual masking", Signal Processing: Image Communication, vol. 27, no. 7, **(2012)**, pp. 772-783.
[7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions On Image Processing, vol. 13, no. 4, **(2004)**, pp. 600-612.
[8] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment", Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computers, Pacific Grove, Canada, **(2003)** November 9-12.
[9] J. Li, K. Wu, X. Zhang, and M. Ding, "Image quality assessment based on multi-channel regional mutual information", International Journal of Electronics and Communications, vol. 66, no. 9, **(2012)**, pp.784-787.
[10] L. Zhang, L. Zhang, X. Mou and D. Zhang, "Feature similarity index for image quality assessment", IEEE Transactions On Image Processing, vol. 20 , no. 8, **(2011)**, pp. 2378-2386 .
[11] Zheng Liu and R. Laganiere, "Phase congruence measurement for image similarity assessment", Pattern Recognition Letters, vol. 28, no. 1, **(2007)**, pp. 166-172.
[12] Shnayderman, A. Gusev and A. M. Eskicioglu, "An SVD-based grayscale image quality measure for local and global assessment", IEEE Transactions On Image Processing, vol. 15, no. 2, **(2006)**, pp. 422-429.
[13] D. O. Kim, H. S. Han, and R. H. Park, "Gradient information-based image quality metric", IEEE Transactions on Consumer Electronics, vol. 56 , no. 2 , **(2010)**, pp. 930-936.
[14] G. H Chen, C. L. Yang, and S. Xie, "Gradient-based structural similarity for image quality assessment", IEEE International Conference on Image Processing, Georgia, U.S.A, 2929-2932, **(2006)** October 8-11.
[15] J. Y Zhu and N. C. Wang, "Image quality assessment by visual gradient similarity", IEEE Transactions On Image Processing, vol. 21, no. 3, **(2012)**, pp. 919-933.
[16] Liu, W. Lin and M. Narwaria, "Image quality assessment based on gradient similarity", IEEE Transactions On Image Processing, vol. 21 , no. 4, **(2012)**, pp. 1500-1512.
[17] H. R. Sheikh and A. C. Bovik, "Image information and visual quality", IEEE Transactions On Image Processing, vol. 15, no. 2, **(2006)**, pp. 430–444.
[18] Z. Wang and Q. Li, "Information Content Weighting for Perceptual Image Quality Assessment", IEEE Transactions On Image Processing, vol. 20, no. 5, **(2011)**, pp. 1185-1198.
[19] C. F. Li and A. C. Bovik, "Content-partitioned structural similarity index for image quality assessment", Signal Processing: Image Communication, vol. 25, no. 7, **(2010)**, pp. 517-526.
[20] J. H. Shen, "On the foundations of vision modeling: I.Weber's law and Weberized TV restoration", Physica Nonlinear Phenomena, vol. 175, no. 3-4, **(2003)**, pp.241-251.
[21] H. Chen and P. K. Varshney, "A perceptual quality metric for image fusion based on regional information", Proceedings of the International Society for Optical Engineering, Orlando, FL, United states, vol. 5831, no. 4, **(2005)** March 30-31, pp. 34-55.

Authors

Shaoping Xu received the M.S. degree in Computer Application from the China University of Geosciences, Wuhan, China, in 2004 and the Ph.D. degree in Mechatronics Engineering from the Nanchang University, Nanchang, China in 2010. He is currently a Professor in the Department of Computer Science and Technology, School of Information Engineering, at the Nanchang University. As the applicant and key investigator of projects, his applications were

approved by the National Natural Science Foundation of China under Grant 61163023 and 61175072 in the year 2011, respectively. As the key investigator of project, his application was also approved by Chinese National 863 Program under Grant 2013AA013804 in the year 2012. Dr. Xu has published more than 30 research articles and serves as a reviewer for several journals including IEEE Transactions on Instrumentation and Measurement, Signal, Image and Video Processing, and International Journal of Robotics and Automation. His current research interests include digital image processing and analysis, computer graphics, virtual reality, surgery simulation, in many practical scenarios *etc.*

Xiaoping Liu received his Ph.D. degree from the University of Alberta in 2002. He has been with the Department of Systems and Computer Engineering, Carleton University, Canada since July 2002 and he is currently a Canada Research Chair. He is also with the School of Information Engineering, Nanchang University as an Adjunct Professor. Dr. Liu has published more than 150 research articles and serves as an Associate Editor for several journals including IEEE/ASME Transactions on Mechatronics and IEEE Transactions on Automation Science and Engineering. He received a 2007 Carleton Research Achievement Award, a 2006 Province of Ontario Early Researcher Award, a 2006 Carty Research Fellowship, the Best Conference Paper Award of the 2006 IEEE International Conference on Mechatronics and Automation, and a 2003 Province of Ontario Distinguished Researcher Award. Dr. Liu is a licensed member of the Professional Engineers of Ontario (P.Eng) and a senior member of IEEE.

Shunliang Jiang is Professor of Computer Science & Technology at Nanchang University, china. His major interest is in the numerical methods with emphasis on the optimization and simulation of industry processing. His recent interests are in the areas of concurrent data structures and heuristic optimization techniques with the application to image processing, analysis, and machine vision.