Towards Accurate and Efficient Image Quality Assessment with Interest Points

Xiang Zhang, Shiqi Wang, Siwei Ma and Wen Gao
Institute of Digital Media, Peking University, Beijing 100871, China
Email: {x_zhang, sqwang, swma, wgao}@pku.edu.cn

***Abstract*—In recent years, the dramatic development of cloud computing, referring to as the applications and services implemented over Internet, has been witnessed and draws many attentions from both academia and industry. Objective image quality assessment (IQA) is fundamental to a broad range of applications throughout the fields of image processing and computer vision. There is a huge desire in exploiting new design of IQA model which is not only accurate but also efficient for fitting the requirement under the background of big data. Many successful models have been built for accurate prediction of the perceptual visual quality, where some typical characteristics of human visual system (HVS) are utilized and incorporated in IQA systems. The well-known *foveation effect* assumes that the regions around the fixation points are much more attractive to human eyes, thus the quality of these regions would significantly influence the overall visual quality. In this paper, we analyze the correlations between the fixation point and quality assessment by integrating several state-of-the-art interest point detection algorithms into IQAs. Experimental results on public database demonstrate that the addictive information of interest point is helpful for improving accuracy of popular IQA models, and meanwhile dramatically reducing the computational complexity. Furthermore, the parameter impacts on IQA performance are thoroughly analyzed showing that the parameters should be carefully designed for different IQA models as well as viewing conditions.**

I. Introduction

Objective image quality assessment (IQA) aiming at accurately estimating the subjective feeling of natural scenes has been studied for decades. It is an essential controller in many image/video processing applications such as transmission, coding, restoration, denoising, detection and recognition. The simplest IQA method is the mean square error (MSE) or peak signal-to-noise ratio (PSNR) and is successfully applied in many real-time applications due to its efficiency. MSE is fast and effective in measuring the signal difference, but cannot correlate well with the perceptual experience [1]. Thus modern IQA models always take advantages of the HVS features for obtaining more accurate metrics, including the contrast sensitivity function, contrast masking, luminance masking etc. Structure similarity (SSIM) [2] and its variant Multi-scale SSIM [3] schemes have shown better correlation with visual perception and demonstrate its power in many applications such as perceptual video coding [4], [5]. It is assumed that the natural scene is highly structured and human eyes are much adaptive with the structure content. For the arrival of the age of big data, the IQA model meets the big challenge of achieving a good tradeoff between accuracy and efficiency.

(a) Original image (b) Foveated image

Fig. 1. Illustration of the *foveation effect* in visual perception. The white area in green lawn is assumed as the fixation point. The pixels around fixation point are perceived by high resolution while others are blurred.

In this paper, we aim at killing two birds with one stone: high accuracy prediction with low complexity calculation. The feature descriptors are usually stored off-line in the cloud, which facilitate the application of the proposed algorithm in cloud based image processing such as transcoding. The localized quality from the feature descriptors can be directly employed to reflect the quality of the overall image. This is also in accordance with the philosophy when HVS perceives the image quality. Previous works [6]–[8] have attempted to utilize the performance of feature extraction to estimate the subjective perception.

The motivation of this paper is from the fact that the response (or sensitivity of HVS) on different image regions is absolutely distinct, as the human eyes always focus on some interest regions. It can be accounted by the well-known *foveation effect* in visual perception, where the resolution in the retina rapidly decreases with the increasing distance to the fixation centra. When one fixates at a point in an image, the region around it is sampled with the highest resolution, while the remote region would be felt like blurred as shown in Fig. 1. It is because of the nonuniform distributions of cone receptors and ganglion cells in the retina [9]. Therefore it can be inferred that regions should be associated with different weighting factors according to its visual importance.

The weighting process in IQA is also referred to as pooling, where the local quality map is merged as one single score that represents the overall quality of the image. It is known that the pooling strategy is crucial to the IQA performance. A general way assumes that the regions with severe distortion contribute a lot to the ultimate quality. It is called distortion-guided pooling strategy. In [10], the local scores are weighted by the information content and the IQA performance has been

improved. In this paper, we propose a weighting strategy that mimics the *foveation effect* of HVS. First a visual saliency map is generated by the extracted interest points, then this map performs as the weighting factors in pooling stage. Several state-of-the-art algorithms of interest point detection are integrated in IQA models to verify the effectiveness of this method. The interest point detection techniques can be categorized into corner based methods and blob based methods. A corner can be defined as the intersection of two edges and always considered as an fixation point that will attract human attention. A blob region in an image typically contains different properties, such as brightness or color, compared to the surrounding areas.

The contribution of this paper is twofold. First we build the connection between the interest point detection and the image quality assessment, and several state-of-the-art algorithms of interest point detection are integrated and compared in this work. We have found that both the accuracy and efficiency of IQAs can be improved by the guidance of the interest point detection. Second the parameter impacts on IQA performance are deeply analyzed showing that the optimal parameter settings of different IQA models are distinct.

The remaining of this paper is organized as follows. In Section II we introduce some state-of-the-art techniques of interest point detection, then the quality assessment algorithms are modified by incorporating the interest points. The experimental results of IQA performance improvements and parameter impacts are given in Section III. Section IV concludes this paper.

II. Interest Point Guided Image Quality Assessment

A. Interest Point Detection

In this subsection, we will give an overview of the interest point detection technique and several typical methods will be briefly introduced.

An interest point is a clear, well-defined, mathematically well-founded position in an image space and can be detected robustly with illuminance variations as well as geometrical changes including translation, rotation, scaling etc. We divide the interest point detection algorithms into two categories, corner based methods and blob based methods.

1) Corner Methods: A corner can be defined as the intersection of two edges or a point for which there are two dominant and different edge directions in a local neighbourhood of the point. In practice, most so-called corner detection methods detect interest points in general, rather than corners in particular. Without loss of generality, let I denote the given image. Considering a local patch (u, v) and a corresponding patch with a small shifting by (x, y). The weighted sum of squared difference between the two patches can be written as follows,

$$S(x,y) = \sum_u \sum_v w(u,v)(I(u+x, v+y) - I(u,v))^2. \quad (1)$$

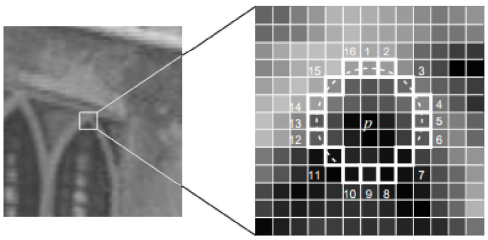


Fig. 2. Illustration of the 16 pixels around a candidate point p in FAST [13].

$I(u+x, v+y) - I(u,v)$ can be approximated by a Taylor expansion. Such that Eqn. (1) can be rewritten as,

$$S(x,y) \approx \sum_u \sum_v w(u,v)(I_x(u,v)x + I_y(u,v)y)^2, \quad (2)$$

where I_x and I_y are the partial derivatives of image I. Alternatively Eqn. (2) can be represented in matrix form,

$$S(x,y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}. \quad (3)$$

A is the structure tensor defined by,

$$A = \sum_u \sum_v w(u,v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}, \quad (4)$$

where the notation $\langle \bullet \rangle$ denotes the summation operator over u and v. A corner could be characterized by a large variation of S in all directions of the vector. By analyzing the eigenvalues of A, this characterization can be expressed in a mathematical way that A should have two "large" eigenvalues for an interest point. In [11], corner is detected by computing the $min(\lambda_1, \lambda_2)$, where λ_1 and λ_2 are two eigenvalues of A. To avoid the computational expenses in solving the eigenvalue, famous Harris Corner method [12] is proposed by using the following function M_c for simplification,

$$M_c = \lambda_1 \lambda_2 - \kappa(\lambda_1 + \lambda_2)^2 = \det(A) - \kappa \cdot trace^2(A), \quad (5)$$

where the computing of eigenvalue is replaced by evaluating the determinant and trace of the matrix A.

Features from accelerated segment test (FAST) [13] is a high-efficiency corner detection method, which is faster than many other well-known feature extraction methods and suitable for real-time image/video processing applications. FAST corner detector uses a circle of 16 pixels (a Bresenham circle of radius 3) to classify whether a candidate point p is actually a corner as shown in Fig. 2. If a set of N contiguous pixels in the circle are all brighter or darker than the intensity of candidate pixel p (denoted by I_p) plus a threshold value t, then p is classified as corner. More recently, the robust BRIEF [14], BRISK [15] and ORB [16] descriptors are proposed as the extensions of the FAST version for faster and more accurate performance in feature detection.

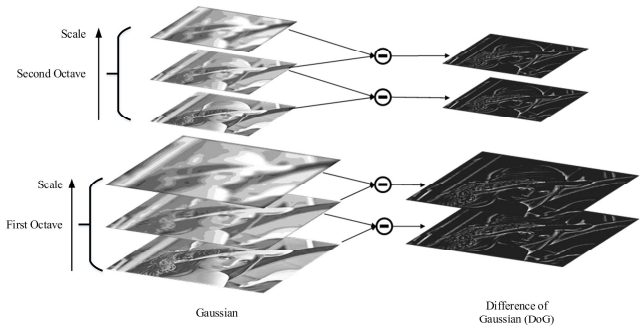


Fig. 3. Illustration of the Difference of Gaussian (DoG) [17].

2) Blob Methods: In the field of computer vision, blob detection refers to mathematical methods aiming at detecting regions in an image that differ in properties, such as brightness or color, compared to areas surrounding those regions. And blob detection has been applied in many feature detection techniques including Scale Invariant Feature Transform (SIFT) [17] and Speeded Up Robust Features (SURF) [18].

Laplacian of Gaussian (LoG) is one of the most common blob detectors. Given an input image $f(x, y)$, this image is convolved by a Gaussian kernel as follows,

$$g(x, y, t) = \frac{1}{\sqrt{2\pi t^2}} e^{-\frac{x^2+y^2}{2t^2}}, \tag{6}$$

where t is the scale parameter. Thus the Gaussian blurred image can be obtained by,

$$L(x, y, t) = g(x, y, t) * f(x, y). \tag{7}$$

Then the Laplacian operator is applied to the blurred image, and the extreme points with maximum or minimum values are detected in the multi-scale LoG space. The LoG operator is effective and efficient in extracting the feature points because the Laplacian method is helpful for detecting corner and boundary points, and the Gaussian blurring can reduce the errors caused by noises.

However the LoG has a disadvantage of big computational complexity induced by the second derivative in Laplacian operator. To overcome this, an approximation method called Difference of Gaussian (DoG) is proposed by computing the difference between two adjacent Gaussian smoothed images in scale space as follows,

$$\begin{aligned} D(x, y, t) &= L(x, y, t + 1) - L(x, y, t) \\ &= (g(x, y, t + 1) - g(x, y, t)) * f(x, y). \end{aligned} \tag{8}$$

This process, which is applied in the famous SIFT descriptor, is schematically illustrated in Fig. 3.

For SURF descriptor, the Determinant of Hessian (DoH) operator is used for extracting the feature points. The Hessian Matrix of a given image is defined as,

$$HL = \begin{bmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{bmatrix}, \tag{9}$$

thus the DoH can be calculated as follows,

$$DoH(x, y, t) = \det HL(x, y, t) = L_{xx}L_{yy} - L_{xy}^2. \tag{10}$$

Maximally Stable Extremal Regions (MSER) [19] is another method of blob detection in images by finding correspondences between image elements from two images with different viewpoints. This method of extracting a comprehensive number of corresponding image elements contributes to the wide-baseline matching, and it has led to better stereo matching and object recognition algorithms.

B. Interest Point Guided Image Quality Assessment

The basic motivation of this paper is to gain some insights into the relationship between the interest points and the quality assessment metrics. It is known that the *foveation effect* will naturally happens behind human physiological system when perceiving visual elements from outside world, and it has significant impact on the ultimate visual experience. Intuitively, we assume that the regions around detected feature points will attract most of the attentions from human eyes and these regions are very important to the image quality assessment algorithms. This approach can be regarded as a kind of low-level "saliency" based pooling method where the "saliency" is a general term indicating the local image feature points are of perceptual significance. The high-level saliency method, also called object-based pooling method, should be associated with some semantic information such as face recognition etc.

Generally, let point p be a marked interest point, the saliency value of another common point p' is dependent to the distance between them. The farther the distance is, the saliency will be weaker. We utilize the normalized Gaussian function to build this model as follows,

$$vi(p'|p) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{d(p,p')^2}{2\sigma^2}}, \tag{11}$$

where the $vi(p'|p)$ indicates the visual saliency value of the point p' given point p. The distance $d(p, p')$ represents the Euclidean distance between two points. The parameter σ is the variance of the Gaussian distribution that controls the change rate of visual saliency around the fixation point. Based on Eqn. (11), the visual saliency map (VSM) can be obtained given the position of the interest points. To alleviate the mistakes caused by some wrong detected noise points, the VSM is calculated in a simpler but more robust way. Let $P_{x,y}$ be a map indicating whether a point at position (x, y) is an interest point or not,

$$P_{x,y} = \begin{cases} 1, & \text{point at } (x, y) \text{ is an interest point;} \\ 0, & \text{otherwise.} \end{cases} . \tag{12}$$

Then a Gaussian kernel with the variance σ and window size w is defined and convolved with the map $P_{x,y}$, which finally create the VSM $V_{x,y}$. Examples of original image with interest points and generated VSM are demonstrated in Fig. 4. The darker regions in the VSM indicate lower significance and vice versa. It can be seen from the figure that the regions containing more interest points have higher importance values.

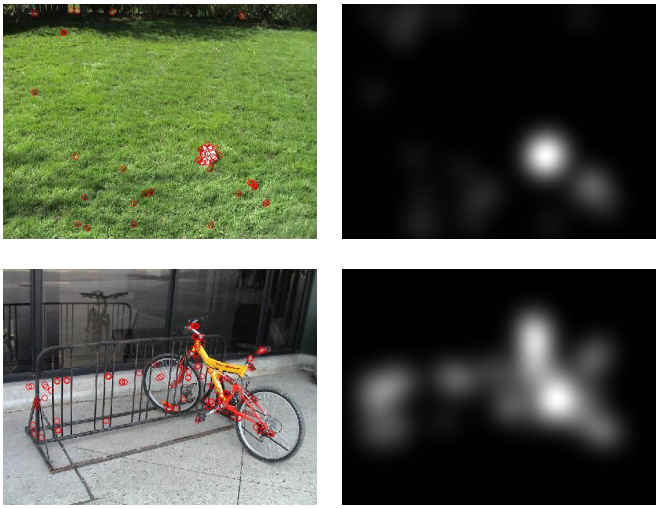


Fig. 4. Generation of visual saliency map by the extracted interest points. The original images with red feature points are shown at the left column, and the corresponding visual saliency maps are shown at the right column. The feature points are detected by SURF algorithm [18]. The parameters σ and w of the Gaussian kernel are set to be 30 and 200 respectively.

It accords well with the perceptual feelings as these areas are more attractive to human eyes. Meanwhile the noise points in the uninterested regions such as the lawn are eliminated by the Gaussian operator.

Finally, we utilize the generated VSM to guide the image quality assessment algorithm. The general IQA scheme consists of two-stage process: local quality measurement and pooling. Let $Q_{x,y}$ be the calculated local quality map using different IQA methods. The simplest and typical way of pooling for getting the ultimate quality score is by averaging the quality map without using any prior knowledge. VSM is such a good indicator which represents the relative significance of each pixel. Thus incorporating the VSM, the average pooling can be replaced by weighted average of the local quality map as follows,

$$Score_W = \frac{\sum_{x,y} Q_{x,y} V_{x,y}}{\sum_{x,y} V_{x,y}}, \tag{13}$$

where the $Score_W$ is the quality score weighted by VSM.

III. Experiments and Analysis

Extensive simulations are conducted to evaluate the performance gain in IQAs guided by interest points. The interest points are detected using different algorithms including Harris [12], FAST [13], ORB [16], BRISK [15], SIFT [17], SURF [18] and MSER [19]. Note that the first four methods are corner based and the last three methods are blob based.

Each algorithm of interest point detection has its own parameters that control the number of detected points. It is unfair if different methods generate distinct number of point when comparing the performance. To solve this problem, we introduce a ranking strategy to make sure every methods are with the same ability of extracting feature points from image. All the detected points are reordered by response value, which indicates the significance level of a point. And the first K points with maximum response value are drawn out for further evaluation. We demonstrate some examples of visual saliency map (VSM) generated by different interest point models in Fig. 5, where the images are chosen from the Bruce' database [20]. The number of extracted points K is fixed to 100 for illustration.

A. Accuracy Improvement of Image Quality Assessment

The simulations for IQA are conducted in public LIVE database [21]. The LIVE database contains 29 reference images and 779 distorted images, undergone 5 different distortion types including JPEG2000, JPEG, white noise, Gaussian blur, and fast fading channel distortion. Three criteria are employed in evaluating the performance of IQAs, which are Pearson linear correlation coefficient (PLCC), Spearman rank-order correlation coefficient (SROCC) and Kendall rank-order correlation coefficient (KROCC).

In order to quantify the accuracy of predicted quality, we use a five-parameter logistic function for non-linear mapping before calculating the PLCC criterion:

$$p(o) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2 (o - \beta_3)}} \right) + \beta_4 o + \beta_5, \tag{14}$$

where o and p represent the objective score and predicted subjective score respectively.

Suppose the original IQA model is X, and the corresponding correlation coefficients between subjective and objective scores are $PLCC_X$, $SROCC_X$ and $KROCC_X$ respectively. The weighted IQA model by interest points is denoted as X_w, and the correlation coefficients are $PLCC_{X_W}$, $SROCC_{X_W}$ and $KROCC_{X_W}$ respectively. Such that the performance improvements $\Delta_{PLCC\,X}$, $\Delta_{SROCC\,X}$ and $\Delta_{KROCC\,X}$ are given by,

$$\begin{aligned} \Delta_{PLCC\,X} &= PLCC_{X_W} - PLCC_X, \\ \Delta_{SROCC\,X} &= SROCC_{X_W} - SROCC_X, \\ \Delta_{KROCC\,X} &= KROCC_{X_W} - KROCC_X, \end{aligned} \tag{15}$$

Three popular IQAs, including PSNR, SSIM [2] and Multi-scale SSIM (MS-SSIM) [3], are involved as evaluation methods. Note that the VSM should be downsampled for fitting different scales in MS-SSIM algorithm. The performance results in terms of PLCC, SROCC and KROCC metrics under different interest point models are listed in Tab. I, where the number of extracted points K is set to be 500, and the parameters of Gaussian blur σ and w are 60 and 400 respectively. It has been shown that all the methods of interest point detection have positive influence to the IQA performance and verifies the hypothesis that the regions around the interest points are with stronger correlation to visual quality. Comparing the average performance gain over different methods, it can be concluded that the blob based methods outperform the corner based methods in the IQA improvement.

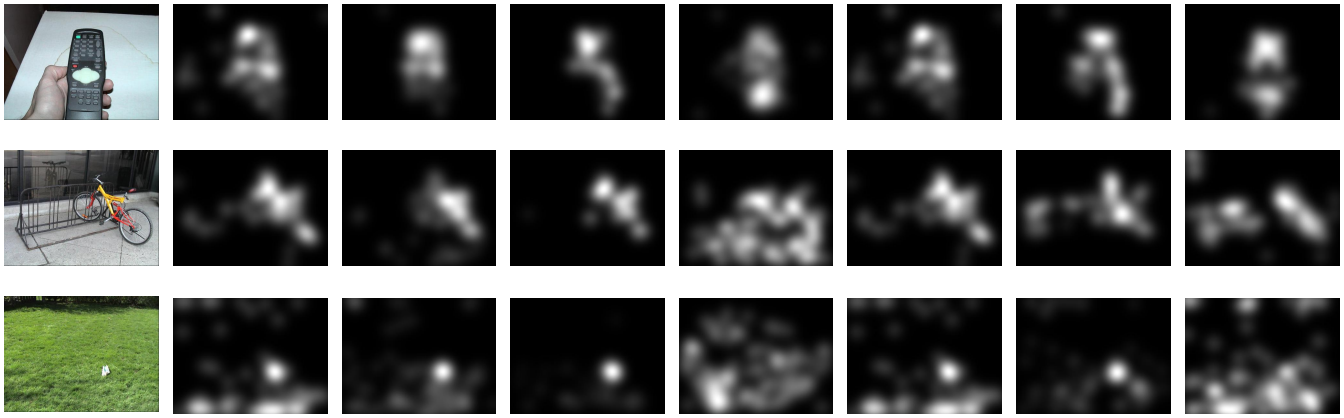


Fig. 5. Results of visual saliency map generated by different interest point models. The original images are given in the leftmost column, and the visual saliency maps generated by Harris, FAST, ORB, BRISK, SIFT, SURF and MSER algorithms are shown subsequently from left to right column respectively.

TABLE I
PERFORMANCE IMPROVEMENTS OF IQAS IN TERMS OF PLCC, SROCC AND KROCC METRICS UNDER DIFFERENT INTEREST POINT MODELS.

	Δ_{PLCC}			Δ_{SROCC}			Δ_{KROCC}			Average		
	PSNR	SSIM	MS-SSIM	PSNR	SSIM	MS-SSIM	PSNR	SSIM	MS-SSIM	PSNR	SSIM	MS-SSIM
Harris	0.01374	0.00716	0.00492	0.01526	0.00844	0.00727	0.01989	0.01883	0.01733	0.01630	0.01148	0.00984
FAST	0.01346	0.00710	0.00507	0.01520	0.00869	0.00772	0.01991	0.01919	0.01851	0.01619	**0.01166**	**0.01043**
ORB	0.01321	0.00686	0.00484	0.01489	0.00847	0.00744	0.01957	0.01870	0.01787	0.01589	0.01134	0.01005
BRISK	0.01483	0.00459	0.00262	0.01679	0.00599	0.00470	0.02152	0.01389	0.01172	**0.01772**	0.00815	0.00635
SIFT	0.01418	0.00795	0.00571	0.01633	0.00954	0.00833	0.02096	0.02025	0.01906	0.01716	**0.01258**	**0.01104**
SURF	0.01566	0.00857	0.00642	0.01789	0.01032	0.00903	0.02201	0.02215	0.02061	**0.01852**	**0.01368**	**0.01202**
MSER	0.01662	0.00677	0.00503	0.01871	0.00826	0.00736	0.02389	0.01807	0.01711	**0.01974**	0.01103	0.00983
Average	0.01453	0.00700	0.00494	0.01644	0.00853	0.00741	0.02111	0.01873	0.01746	0.01736	0.01142	0.00994

B. Computation Reduction of Image Quality Assessment

The proposed scheme not only improve the IQA accuracy but also reduce the computational complexity. The weighting technique provides a clue that the pixels with low significance have almost no impact on visual quality, thus the computation over these regions is redundant and can be further eliminated. In this subsection, we provide the results of the relationship between the computational reduction and the IQA performance changes.

In this simulation, given the visual saliency map $V_{x,y}$, the $r \times 100\%$ data with least saliency value is dropped out by setting the corresponding $V_{x,y}$ to zero as follows,

$$V'_{x,y} = \begin{cases} 0, & V_{x,y} < th \\ V_{x,y}, & otherwise \end{cases}, \tag{16}$$

where th is a threshold associated with the ratio r satisfying that the $r \times 100\%$ data in $V_{x,y}$ is smaller than th. The $V'_{x,y}$ is the modified VSM and is applied in following pooling stage as aforementioned in Eqn. (13).

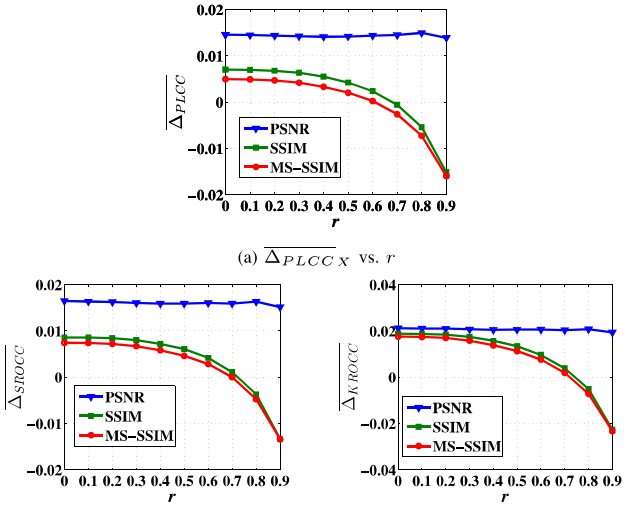
Subsequently, we evaluate the average performance gain of IQAs with changing values of r. Given an IQA model X, the average performance gain is denoted by $\overline{\Delta_{PLCC}}_X$, $\overline{\Delta_{SROCC}}_X$ and $\overline{\Delta_{KROCC}}_X$ respectively as follows,

$$\begin{aligned} \overline{\Delta_{PLCC}}_X &= \frac{\sum_W \Delta_{PLCC_W}}{N}, \\ \overline{\Delta_{SROCC}}_X &= \frac{\sum_W \Delta_{SROCC_W}}{N}, \\ \overline{\Delta_{KROCC}}_X &= \frac{\sum_W \Delta_{KROCC_W}}{N}, \end{aligned} \tag{17}$$

where $N = 7$ is the number of interest point detection method in this paper. These average metrics will be frequently used in the remaining part of the paper for evaluating the parameter impact on the IQA models.

The performance change in terms of r is illustrated in Fig. 6, where r denotes the ratio of computation reduction and the vertical axis represents the performance variance of IQAs. It is noted that the performance gain of PSNR keeps stable even when the 90% data has been discarded ($r = 0.9$), while the SSIM and MS-SSIM are relatively sensitive to the amount of processed data. The performance gain of SSIM and MS-SSIM becomes negative till r reaches about 0.7.

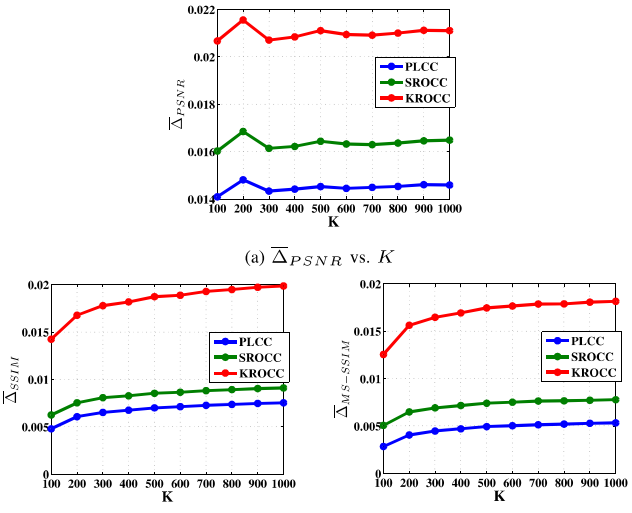
Therefore, we can imply that the computational quantity can be significantly reduced by about 70% while keeping the IQA accuracy unchanged. Even more data can be removed and only a bit amount of data is required for PSNR method. This is because of the fact that most of the significant data has



(a) $\overline{\Delta_{PLCC}}_X$ vs. r

(b) $\overline{\Delta_{SROCC}}_X$ vs. r (c) $\overline{\Delta_{KROCC}}_X$ vs. r

Fig. 6. Curves of the average performance gain (i.e. (a) $\overline{\Delta_{PLCC}}_X$, (b) $\overline{\Delta_{SROCC}}_X$ and (c) $\overline{\Delta_{KROCC}}_X$) in terms of r. The IQA models X are PSNR, SSIM and MS-SSIM respectively.



(a) $\overline{\Delta}_{PSNR}$ vs. K

(b) $\overline{\Delta}_{SSIM}$ vs. K (c) $\overline{\Delta}_{MS-SSIM}$ vs. K

Fig. 7. Curves of the average performance gain in terms of K. The IQA models X are (a) PSNR, (b) SSIM and (c) MS-SSIM respectively.

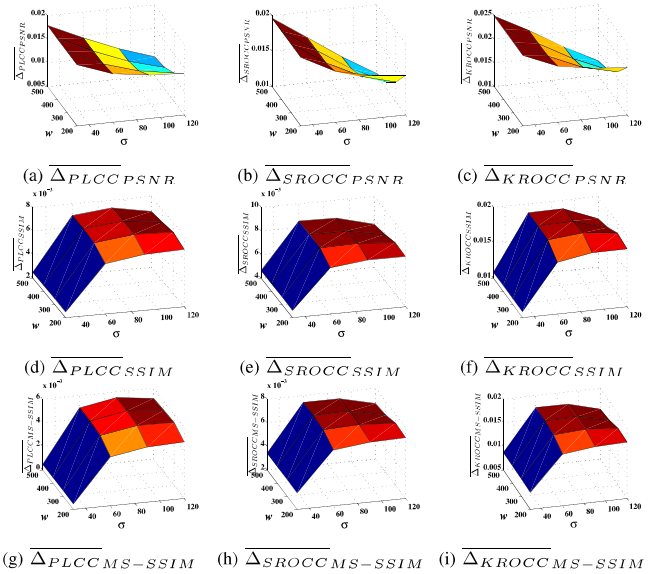
been reserved by the guidance of interest points. The results also indicate that the interest point guided scheme is effective in predicting the perceptual importance of natural images.

C. Parameter Impact on IQA Performance

1) The Number of Interest Point K: The number of extracted points, i.e. K, will influence the VSM and consequently the IQA performance. We evaluate the average performance gain of IQAs under different values of K. Fig. 7 gives the curves of the $\overline{\Delta_{PLCC}}_X$, $\overline{\Delta_{SROCC}}_X$ and $\overline{\Delta_{KROCC}}_X$ in terms of K which ranges from 100 to 1000. The IQA models X are PSNR, SSIM and MS-SSIM respectively. The curves of SSIM and MS-SSIM methods have the similar tendency that gradually increase with the value of K, indicating more interest points involved in the IQA models lead to higher performance. But the performance gain of PSNR method with increasing K is not as remarkable as that of SSIM and MS-SSIM.

2) Parameters of Gaussian Blur – σ and w: The Gaussian blur process is also critical in generating the VSM. In this subsection we analyze the impact of the two parameters in Gaussian blur, i.e. σ and w. In these experiments, the number of feature points K is fixed to be 500 while the σ ranges from 30 to 120 and w from 200 to 500 respectively. The performance improvements on three IQAs (PSNR, SSIM and MS-SSIM) in terms of three measuring metrics (PLCC, SROCC and KROCC) under different parameter sets are illustrated as 3D plots in Fig. 8. It is found that the performance impact on different IQA methods is distinct. The smaller value of σ leads to better performance for PSNR but on the contrary for SSIM and MS-SSIM. The larger value of w has the similar impact on different IQAs that leading to higher performance.

Therefore the parameters should be specifically designed in terms of different IQA models as well as the viewing conditions. For instance good parameters σ and w will account for



(a) $\overline{\Delta_{PLCC}}_{PSNR}$ (b) $\overline{\Delta_{SROCC}}_{PSNR}$ (c) $\overline{\Delta_{KROCC}}_{PSNR}$

(d) $\overline{\Delta_{PLCC}}_{SSIM}$ (e) $\overline{\Delta_{SROCC}}_{SSIM}$ (f) $\overline{\Delta_{KROCC}}_{SSIM}$

(g) $\overline{\Delta_{PLCC}}_{MS-SSIM}$ (h) $\overline{\Delta_{SROCC}}_{MS-SSIM}$ (i) $\overline{\Delta_{KROCC}}_{MS-SSIM}$

Fig. 8. The 3D plots of performance improvements on three IQAs (PSNR, SSIM and MS-SSIM) in terms of three measuring metrics (PLCC, SROCC and KROCC) over the parameter space.

the actual *foveation effect* in visual perception, and this effect would be significantly influenced by the viewing conditions such as viewing distance and image size.

IV. Conclusion

In this paper, we explore the relationship between the interest point and the image quality assessment by mimicking the *foveation effect* of HVS. First the visual saliency map is generated by the extracted interest points and then the importance map performs as the weighting factors in pooling stage of IQAs. Seven popular algorithms of interest point detection are involved and compared in this work. Extensive

experiments have been conducted and the results have shown that the performance of IQAs (including accuracy as well as efficiency) can be improved by introducing the interest point detection techniques. Furthermore parameter impacts on IQA performance are analyzed, which shows that the parameters should be carefully designed for different IQA models as well as viewing conditions.

ACKNOWLEDGMENT

This work was supported in part by the Major State Basic Research Development Program of China (2015CB351800) and in part by the National Science Foundation (61322106, 61390515 and 61210005). This work was also granted by Cooperative Medianet Innovation Center.

REFERENCES

[1] B. Girod, "Psychovisual aspects of image processing: What's wrong with mean squared error?" in *Multidimensional Signal Processing, Proceedings of the Seventh Workshop on*, 1991.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[3] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Signals, Systems and Computers, Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 2, 2003, pp. 1398–1402.

[4] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate-distortion optimization for video coding," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 4, pp. 516–529, April 2012.

[5] ——, "Perceptual video coding based on SSIM-inspired divisive normalization," *Image Processing, IEEE Transactions on*, vol. 22, no. 4, pp. 1418–1429, April 2013.

[6] M. Nauge, M. Larabi, and C. Fernandez, "A reduced-reference metric based on the interest points in color images," in *Picture Coding Symposium (PCS), 2010*. IEEE, 2010, pp. 610–613.

[7] M. Decombas, F. Dufaux, E. Renan, B. Pesquet-Popescu, and F. Capman, "A new object based quality metric based on SIFT and SSIM," in *Image Processing (ICIP), 2012 19th IEEE International Conference on*. IEEE, 2012, pp. 1493–1496.

[8] J. Farah, M.-R. Hojeij, J. Chrabieh, and F. Dufaux, "Full-reference and reduced-reference quality metrics based on SIFT," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 161–165.

[9] Z. Wang and A. C. Bovik, *Modern image quality assessment*. Morgan & Claypool Publishers, 2006.

[10] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *Image Processing, IEEE Transactions on*, vol. 20, no. 5, pp. 1185–1198, 2011.

[11] J. Shi and C. Tomasi, "Good features to track," in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*. IEEE, 1994, pp. 593–600.

[12] C. Harris and M. Stephens, "A combined corner and edge detector." in *Alvey vision conference*, vol. 15. Manchester, UK, 1988, p. 50.

[13] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 1, pp. 105–119, Jan 2010.

[14] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Binary robust independent elementary features," in *Computer Vision–ECCV 2010*. Springer, 2010, pp. 778–792.

[15] S. Leutenegger, M. Chli, and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2548–2555.

[16] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2564–2571.

[17] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2. Ieee, 1999, pp. 1150–1157.

[18] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 404–417.

[19] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.

[20] N. Bruce and J. Tsotsos, "Saliency based on information maximization," in *Advances in neural information processing systems*, 2005, pp. 155–162.

[21] H. Sheikh, Z.Wang, L. Cormack, and A. Bovik, "LIVE image quality assessment database release 2." [Online]. Available: http://live.ece.utexas.edu/research/quality