

Goal oriented image quality assessment

Kiruthika S. | Dr. Masilamani V.

Department of Computer Science and Engineering,
Indian Institute of Information Technology, Design
and Manufacturing, Kancheepuram, Chennai, India

Correspondence

Kiruthika S., Department of Computer Science and
Engineering, Indian Institute of Information Tech-
nology, Design and Manufacturing, Kancheepuram,
Chennai, India.

Email: coe18d003@iiitdm.ac.in

Abstract

The area of image quality assessment(IQA) is an active research area in image processing and computer vision. All IQA algorithms reported in literature are attempting to quantify only the visual quality of the images/videos. An interesting question to be answered is that given a goal (task) and an image (with good visual quality), is the image good to achieve the goal by the best possible algorithm. In an attempt to spur the research community to answer this question, a new paradigm of IQA, called goal oriented IQA (GO-IQA), is introduced. GO-IQA is defined as given an image and a goal, predicting how good the image is to achieve that goal by the best possible algorithm. The need for GO-IQA is that if GO-IQA score is less, then any arbitrary algorithm attempting to achieve the goal will eventually fail. In this paper, considering segmentation as goal, a GO-IQA algorithm is proposed to predict how good the image is to do accurate segmentation by the best possible segmentation algorithm. A support vector regression model has been trained with features related to image segmentation along with the accuracies of a class of well-known image segmentation algorithms. In the process, we have obtained encouraging results in terms of the Pearson linear correlation coefficient and mean squared error. The predictions given by the proposed model have been validated with the scores of state of the art segmentation algorithm.

1 | INTRODUCTION

IQA is one of the challenging research fields in image processing. The term IQA refers to quantifying the perceptual quality of an image. The IQA has been first used for assessing the quality of television images [1]. This field emerges because of its numerous applications like security and surveillance [2], automatic quality control [3], digital watermarking [4], medical imaging [5], entertainment industry [1] etc. In terms of quantifying the quality, subjective and objective quality assessment methods are used. In the initial days, the quality of an image used to be completely assessed by humans. The score given by human experts is named as the subjective score [1]. This method of assessing an image is said to be a subjective quality assessment. Later due to the growth of digital imaging, the subjective assessment has become a time-consuming process, and hiring the human experts for assessing the quality of an image to the real-time applications has become very difficult. This made the researchers assess the quality of an image with the help of algorithms by mimicking the human visual system (HVS) [6]. The

image quality score given by the IQA algorithm is named as the objective score. Such type of assessing the quality is said to be an objective quality assessment.

Some information may be useful to build such an autonomous algorithm to compute the quality of an image. They are, (1) information about original image (referred as reference image) or the distorted version, (2) the information about distortions that have affected the quality of an image, (3) information about the application where this assessment will be useful (e.g., if the application is a visual display, then the information about human visual physiology can be used for effective prediction of the image quality) [7]. Based on the requirement of additional information to measure the quality of a given image, the IQA algorithms are broadly classified into full-reference IQA [8], reduced reference IQA [9] and no reference IQA methods [10]. As the name implies the full reference IQA algorithm uses a reference image to quantify the visual quality of the given images. The drawback of this method is that it is difficult to get the reference image in some domains such as satellite images, underwater images etc. The reduced reference IQA quantifies

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Image Processing* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology

the quality of the image using the information or attributes of reference images instead of using the reference image itself. This dependency on the information will also restrict the application range, and hence these two methods will not be useful in the situations where the reference image and information about the reference image are unavailable. The researchers started to design a class of algorithms called no reference IQA (NR-IQA) which does not require any information about the reference image to assess the quality of an image. NR-IQA plays a key role in IQA because of its independence on the reference image. It does not require reference or information about the reference image; it blindly assesses the quality of the images.

Most of the NR-IQA algorithms are machine learning algorithms. The machine learning models are trained with image features along with the subjective human scores. In literature, many quality aware features both in spatial domain and transformed domain have been considered. Some of the efficient features are based on information from discrete cosine transform (DCT) statistics [11], natural scene statistics (NSS) [12], hybrid NSS [13], free energy principle and HVS inspired features [14], color space distribution [15] and a lot more. The NR-IQA algorithms that do not consider any additional information are able to quantify image quality when images are affected by some distortions, but not for all. Moreover performance of such algorithm goes down when multiple distortions are there in the image. Hence researchers attempted to design NR-IQA algorithms to assess the quality of an image when it is distorted by a specific distortion or by multiple distortions. Such NR-IQA algorithms are called as distortion specific NR-IQA algorithms. Some of the NR-IQA algorithms for handling multiple distortions are presented in [16], and some specific distortions considered for designing distortion specific NR-IQA algorithms are contrast [17], blur [18], transmission error, and quantization noise [19]. For these types of algorithms, the type of distortion of the image should be known to the researcher.

It has been observed that the NR-IQA algorithm performs well for images of some domains, but not in other domains. This motivated the researchers to consider designing NR-IQA dedicated for images belong to one domain. Such algorithms are called as domain specific NR-IQA algorithms. Some of the domains are aerial images [20], stereoscopic images [21], fetal [22] and liver ultrasound [23] medical images and so on. Recently, efficient NR-IQA have been reported for stereoscopic video [21], and also for synthesized images [24]. The domains in computer graphics such as AR and VR are currently considered for NR-IQA. Domain-specific NR-IQA algorithms are designed and best suited for that specific domain alone. As far as we know, all IQA algorithms including domain-specific algorithms reported in the literature attempt to quantify only the visual quality of the images. In this paper we pose a theoretical question: Given an image and a goal, how good the image is to achieve the goal? If we attempt to answer this question using existing IQA algorithms, the answer is not satisfactory because of the fact that the existing IQA predicts only the visual quality, and image with good visual quality need not be good to achieve the given goal. For instance, segmentation is considered as goal,

some times the visually good quality image cannot be well segmented by the best known segmentation algorithm[25]. Sometimes almost all good segmentation algorithms may fail to segment the visually good quality images. This implies that there is something, apart from visual quality, available in the image which is causing the good segmentation. We attempt to quantify that. Segmentation is considered only as an example goal. Goals such as reconstruction, depth computation, object detection, etc., can be considered. Hence this paper introduces a new paradigm in IQA called goal-oriented IQA (GO-IQA).

The rest of the paper is organized as follows. The need for GO-IQA and GO-IQA with a specific goal, namely segmentation, are discussed in the next section. In Section 3, feature extraction and generation of the label are explained in detail along with the training and testing of the proposed method. In Section 4, the detailed discussion done on the experimental study is presented. The theoretical time complexity of the proposed algorithm is presented in Section 5. The summary, challenges and direction for future work are presented in section 6.

2 | GO-IQA

Although the IQA has been one of the main focus of research in image processing for a long time, all the existing IQA are predicting the image quality measure that are aligned with quality of the image perceived by HVS. But to achieve the specific goals like computation of depth image, segmenting the image, detecting and tracking an object etc., the visual quality alone is not sufficient. To check if an image is good to achieve a goal related to the image, the quality of the image is to be computed based on features related to the specific goal instead of quantifying the image quality correlated well with HVS. As a case study, segmentation is considered as the goal. The observation made through Figures 1 and 2 motivates us to look for a different approach (other than HVS based IQA) to quantify the information available in the image for a good segmentation. The images shown in Figures 1 and 2 are taken from the Berkeley segmentation dataset (BSD). In Figure 1, three images are shown in the descending order of image quality, where the image quality is measured using one of best known NR-IQA method, called natural image quality evaluator (NIQE). But F1 score (segmentation score) for those three images are not in the same descending order. The second image (tiger) in Figure 1 has better image quality, but poorer segmentation score than the third (note that lesser NIQE score means better quality). In BSD dataset, these three images (in the same order of display) have obtained the ranks of 56, 100 and 54, respectively, in terms of their F1 score, but in terms of quality their corresponding ranks are 1, 2 and 4. The images shown in Figure 2 obtained the top three ranks among all the images in BSD in terms of F1 score, but those same images have obtained the ranks of 94, 100 and 78 (in the order of display), respectively, in terms of quality (NIQE score). Therefore, the observation made from Figures 1 and 2 is that although visually good quality images may be well segmented by arbitrary segmentation algorithms in some cases,



FIGURE 1 The images are ordered as per NIQE score from left to right. The NIQE scores are 1.867, 2.135 and 2.160, respectively. The given segmentation F1 scores are 0.77, 0.48 and 0.75, respectively



FIGURE 2 The images are ordered as per the given segmentation F1 score from left to right. The given segmentation F1 scores are 0.93, 0.92 and 0.90, respectively. The NIQE scores are 4.795, 8.315 and 3.462, respectively

they do not get well segmented in other cases. Hence there is a need to look for an alternative way to quantify the amount of information an image has to achieve the given goal in general, segmentation in particular. We call algorithms that assess the information an image has to achieve the given goal as GO-IQA algorithms

Face recognition is an another example given here to show the importance of GO-IQA. One of the preprocessing steps for face recognition when the face is captured at a distance is to convert the low-resolution image into the high-resolution image. This conversion process is also focused to improve the quality of the face image in terms of visual but not for recognition process. The same is observed in the literature that high-resolution images are good for the human visual system, but not all high-resolution images are suitable for the system to recognize the face [26]. Thus GO-IQA algorithm for face recognition may be developed to predict whether the image is best suited for recognition purposes or not without executing the face recognition algorithms. In such case, the goal is face recognition, and the features for predicting the quality of an image will be formed by considering the features needed to recognize the face. Mean scores of class of well known face recognition algorithm can be used to train the model. Therefore, it is required to answer the question that given an image and a goal (task), is the image good enough to achieve the goal by any arbitrary algorithm which attempts to achieve the goal. A special case of the problem is that given an image and the goal as segmentation, can we predict whether the image can be segmented well by an arbitrary segmentation algorithm.

With the strong motivation discussed above, this paper proposes a new paradigm of IQA called GO-IQA. We define

GO-IQA as follows: Given an image and a goal, find out how good the image is to meet the goal by the best possible algorithm. For example, given the stereoscopic image (left and right camera images) find out the quality of images in achieving a very good depth image without computing the depth image. If left and right images are taken with a good quality cameras and the cameras are in the same plane with the same focal length, then it is possible to obtain a good quality depth image, whereas if the input stereo images themselves are taken with low contrast, noisy and also cameras are with different focal length and not co-planar, then even the best-known depth computation algorithm may not produce good quality depth image. In this case the goal is to predict the quality of depth map that has to be computed from the information available in the input image (stereo image). In this attempt, we should not run any depth map computation algorithm to find out the quality of information available in a stereo input image.

Figure 3 shows the binary version of GO-IQA. In this paper, we consider our goal as segmenting the image. We consider the measure for the segmentation as the F1 score. F1 score is the harmonic mean of precision and recall. These scores are calculated between the ground truth (segments drawn by the human) and the output of the segmentation algorithm. The predicted score of a GO-IQA for image segmentation will tell how well this image can be segmented by an arbitrary segmentation algorithm. The prediction process does not involve segmentation. If the GO-IQA score is high then the image can be well segmented by an arbitrary efficient image segmentation algorithm.

The need for GO-IQA for image segmentation is to predict how well the given image will be segmented by any arbitrary segmentation algorithm. In simple words, predicting the F1

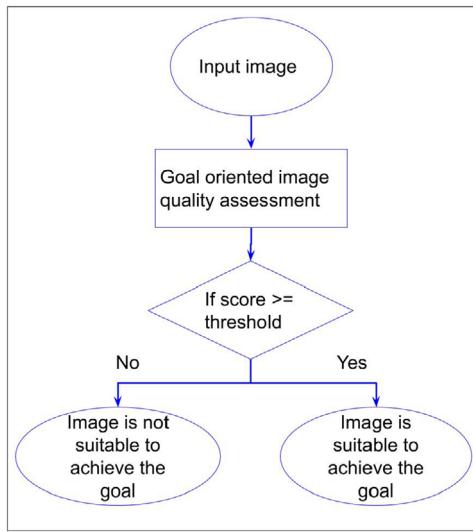


FIGURE 3 Objective of GO-IQA

score for the best possible segmentation algorithm that could ever exist, for the given image without executing the image segmentation algorithm. As far as we know, there is no IQA designed for a specific goal in literature. Thus we develop a GO-IQA for image segmentation. The score given by our method can be the benchmark segmentation score for this image, which can be segmented using any image segmentation algorithm. An interpretation of GO-IQA score for a given image is that if GO-IQA score for an image is less, then all image segmentation algorithms, however complex they may be, will not segment the image with good accuracy. The advantage of our method lies in testing the image without executing any segmentation algorithm. The case of less GO-IQA score of an image can be interpreted as the segmentation algorithm performance will be poor for that image despite the algorithm being faultless. One of the applications of GO-IQA with segmentation as goal is to decide whether or not complex preprocessing is required before segmentation process. GO-IQA score can be obtained, and if it is found to be good, such complex preprocessing can be avoided. Another application is to check for termination of iterative image reconstruction algorithms (CT Scan). All the iterative algorithms that are used in CT image reconstruction aim to provide good visual quality, assuming that the diagnosis is done by medical experts or reconstructed image with good visual quality will lead to good segmentation (which will be a preprocessing step for automatic diagnosis). But both the assumptions need not be true. GO-IQA score can be used to terminate the iteration. If GO-IQA score is less, the reconstruction algorithm may take X-ray images in a few more directions and obtain slice image using X-ray images taken in the previous iterations along with these new X-ray images. If GO-IQA score is high then the reconstruction algorithm will stop. The reconstructed image obtained using this procedure will be very good for automatic segmentation. Since good quality segmentation is required for many of the automatic diagnosis approaches and high GO-IQA ensures efficient automatic segmentation, GO-IQA is a very useful measure.

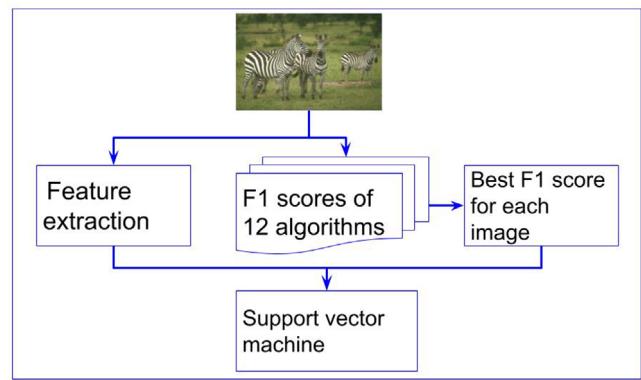


FIGURE 4 Training phase of the proposed method. Feature extraction phase is explained in the Figure 6

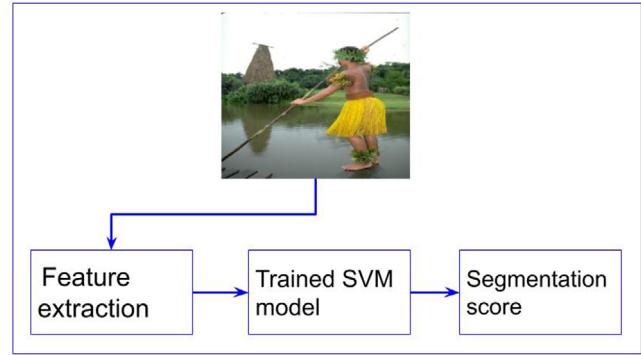


FIGURE 5 Testing phase of the proposed method. Feature extraction phase is explained in the Figure 6

3 | PROPOSED METHOD

Overview of the proposed GO-IQA for image segmentation is given in Figures 4 and 5. As shown in the figure, the support vector regressor is used to predict the quality of an image in terms of segmentation. NIQE [27], asynchronous generalized Gaussian distribution (AGGD) [28], non-subsampled contourlet transform (NSCT) coefficients [29] features from the image and speeded up robust features (SURF) [30] are obtained from the rotated versions and the noisy versions of the image. For training the model, along with these features, the images are labelled with the best segmentation score for an image. We consider the class of 12 segmentation algorithms and each of the algorithms give segmentation score (F1 score) for every image [31]. For each image, the best segmentation score among all the segmentation scores given by the class of algorithms is considered as the best possible segmentation score (best possible score for image I may be the score of an algorithm A_i , the best score for the image J by the algorithm A_j and so on). Thus the best F1 score for each image which is obtained from the 12 best-known segmentation algorithms score is used as the label to train the model. Then the trained model will be used to predict the GO-IQA score of the test images in terms of image segmentation.

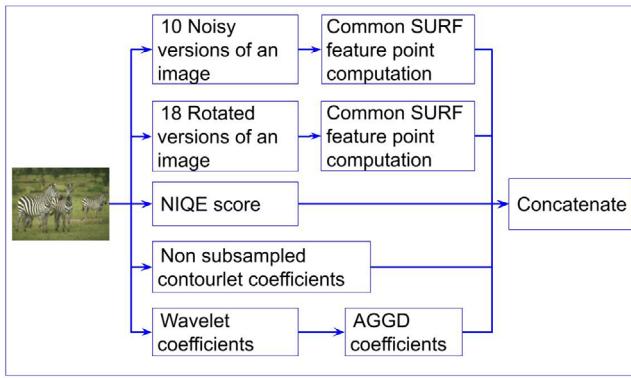


FIGURE 6 Feature extraction for the proposed method

3.1 | Feature extraction

There are five different sets of features considered as shown in Figure 6. The first set of features is from the Haar wavelet coefficients of an image. Discrete Haar wavelet transformation of an image gives a multiresolution representation of an image. The Haar wavelet decomposition [32] of an image f (with the size (M_1, M_2) , where M_1 and M_2 are even numbers) at the different scale factors can be obtained using the equation (1) and (2).

$$W_s(j_0, n_1, n_2) = \frac{1}{\sqrt{M_1, M_2}} \sum_{a=0}^{M_1-1} \sum_{b=0}^{M_2-1} f(a, b) \varphi_{j_0, n_1, n_2}(a, b), \quad (1)$$

$$W_\Psi^D(j_0, n_1, n_2) = \frac{1}{\sqrt{M_1, M_2}} \sum_{a=0}^{M_1-1} \sum_{b=0}^{M_2-1} f(a, b) \Psi_{j_0, n_1, n_2}^D(a, b), \quad (2)$$

where, j_0 is the scale factor $j_0 \geq 1$, $D \in \{b, v, d\}$, $0 \leq n_1 \leq (M_1/2)$, $0 \leq n_2 \leq (M_2/2)$, and the scaling and wavelet functions used in 1 and 2 are,

$$\begin{aligned} \varphi_{j, n_1, n_2}(a, b) &= \varphi_{j, n_1}(a) \varphi_{j, n_2}(b), \\ \Psi_{j, n_1, n_2}^b(a, b) &= \varphi_{j, n_1}(a) \Psi_{j, n_2}(b), \\ \Psi_{j, n_1, n_2}^v(a, b) &= \Psi_{j, n_1}(a) \varphi_{j, n_2}(b), \\ \Psi_{j, n_1, n_2}^d(a, b) &= \Psi_{j, n_1}(a) \Psi_{j, n_2}(b), \end{aligned}$$

$\Psi(a, b)$ is the 2D wavelet function and $\varphi(a, b)$ is the corresponding scaling function of the wavelet transform. $\Psi(a)$ and $\varphi(a)$ are the corresponding 1D functions. In wavelet transform, the image will be decomposed into $2 * 2$ submatrices where each submatrix is of size $(\frac{M_1}{2}, \frac{M_2}{2})$ at each level. These four distinct submatrices will contain the detailed information about the sub-scaling approximation W_s , horizontal W_Ψ^b , vertical W_Ψ^v and diagonal W_Ψ^d directions of an image. Horizontal, vertical and diagonal matrix wavelet coefficients are required for edge detection

in an image. The information extracted for GO-IQA is

$$W = W_\Psi^b + W_\Psi^v + W_\Psi^d,$$

In our experiment, we uniformly resized all the images to (256,256). Thus our submatrices and W are of size (128,128) with the single level of wavelet decomposition. We observed that the coefficients of Haar wavelet are univariate and can be modeled using AGGD to capture the behaviour of the coefficients. And experimentally we found that these coefficients are best fitted with AGGD model. The AGGD parameters, namely $\mu, \gamma_1, \gamma_2, \rho$ are obtained using the following AGGD equation

$$f(w|\mu, \gamma_1, \gamma_2, \rho)$$

$$= \begin{cases} \frac{\rho}{(\gamma_1 + \gamma_2)\Gamma(1/\rho)} e^{-[(w-\mu)/\gamma_1]\rho} & \text{if } w < \mu \\ \frac{\rho}{(\gamma_1 + \gamma_2)\Gamma(1/\rho)} e^{-[(w-\mu)/\gamma_2]\rho} & \text{if } w \geq \mu \end{cases},$$

$\forall w \in W$, where γ_1, γ_2 are left and right shape parameters, μ is the mean value and $\rho = 2$ for asymmetric Gaussian distribution. $\Gamma(\cdot)$ is the gamma function. The parameters $\mu, \gamma_1, \gamma_2, \rho$ are the first set of features considered for proposed method to predict the segmentation score of an image. The reason for choosing the Haar wavelet transform is that it is best suited for edge detection [33] because the Haar wavelet provides more information about a spatial domain when compared to all other wavelets [34]. For finding the quality of an image the discrete wavelet transform will give more information about the geometric distortion and global sharpness of an image [35]. It provides a multiscale analysis of the image in both spatial and frequency domains. The main advantage of this transform is that it will provide multiresolution information, and it helps to model the HVS [36]. One of the best applications of wavelet transform is to detect the objects [37] in the image. As shown in the graph (Figure 7), the objects are represented by the corresponding peaks of the histogram, so the capability of segmenting the image is evident in the graph of wavelet coefficients, and these coefficients are fitted with AGGD model. Thus the AGGD parameters will bring wide information possessed by the image in terms of segmentation capability.

The second set of features is obtained from SURF [30]. The reason for choosing the SURF feature descriptors is that it is faster than scale-invariant feature transform (SIFT). SURF keypoints of the noisy and geometrically distorted versions of the same object do not change much. If the SURF keypoints are different in the geometrically distorted one from the original one then it can be observed that such an image cannot be segmented well. As many of the segmentation algorithms are looking at the boundary points and key points that will not change across many geometric distortions, SURF key points are considered as one set of features in our model. We extract the SURF features not only from each image but also from the rotated and noisy versions of the image.

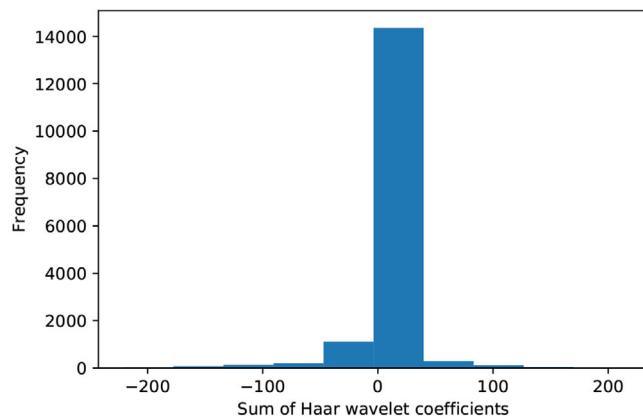


FIGURE 7 An image and its corresponding histogram of sum of wavelet coefficients

The common SURF feature points of rotated and noisy versions of each image I are considered. For generating the noisy images, we added the Gaussian noise with the zero mean and the variances from 0.1 to 0.55 (with the step size of 0.05) to generate 10 noisy images. Repeatability, precision, recall, and F1 score are the features of SURF keypoint prediction generated for each noisy image as stated in the Algorithm 1. Then the mean, median, mode and standard deviation were found for the above features. The feature vector for the SURF noisy version of an image is $(1 \times 16), (4 \text{ (Repeatability, precision, recall, and F1 score)} \times 4 \text{ (mean, median, mode and standard deviation)}) =$

ALGORITHM 1 Common SURF feature point computation

```

Input: Image -  $I$ , Rotated version of  $I$  -  $I_1$ 
Output: Repeatability, Precision -  $Pr$ , Recall -  $Re$ , F1 score -  $F1$ 
1 Let  $f_I$  be the detected SURF feature points of  $I$ 
2  $f_{I_1}$  be the detected SURF feature points of  $I_1$ 
3  $commonfvcnt = 0$ 
4 for  $i = 1$  to  $size(f_I)$  do
5   for  $j = 1$  to  $size(f_{I_1})$  do
6     | Find the absolute difference between that feature
      | points
7   end
8   if  $difference \leq 2$  then
9     | increment  $commonfvcnt$  by 1
10  end
11 end
12  $Repeatability =$ 
     $commonfvcnt/min(size(f_I), size(f_{I_1}))$ 
13  $countoftp = commonfvcnt$ 
14  $countoffp = size(f_{I_1}) - commonfvcnt$ 
15  $countoffn = size(f_I) - commonfvcnt$ 
16  $Pr = countoftp/(countoftp + countoffp)$ 
17  $Re = countoftp/(countoftp + countoffn)$ 
18  $F1 = (2 * (Re * Pr))/(Re + Pr)$ 

```

16). Figure 8 shows the observation that if the amount of noise increases then the value of these features decreases. So it made us stop at the variance of noise 0.55. In the graph, X -axis is the amount of noise added in the image and Y -axis is the value of that feature predicted by Algorithm 1.

Algorithm 1 is also used to find the features in the rotated versions of an image. The degree of rotation varies from 10 to 180, a total of 18 different images were considered. Features of SURF keypoint prediction for 18 rotated images were found similar to the generation of SURF keypoint prediction in noisy images. In Figure 9, we found that in all the features, values are repeated from 180 to 360 degrees of the rotated images, so we considered the features of 10 to 180 degree rotated images. In the graph, X -axis is the degrees of rotation in the images and Y -axis is feature values of SURF keypoint prediction, respectively. The size of the feature vector is (1×16) for rotation.

The statistical features of an image in the spatial domain are important for image segmentation. The next feature used in the process is NIQE score [27], which is used to obtain the quality measure of an image by analyzing its statistical features. Statistical NSS features of pristine and test images were modeled by MVG model. Image I is normalized and partitioned into multiple patches. For each patch, the sharpness is calculated. The patches with the sharpness greater than the threshold (>0.75)

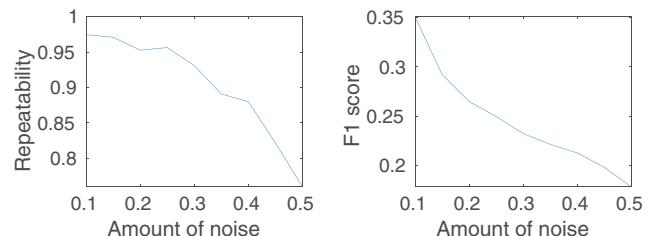
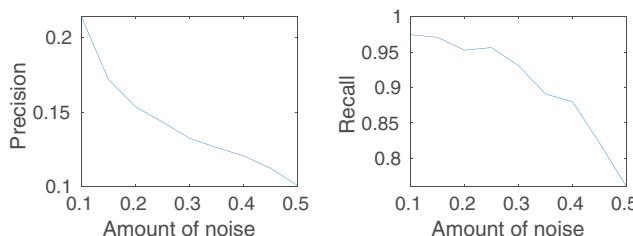


FIGURE 8 Precision, recall, repeatability and F1 score obtained from SURF for noisy version of an image. X -axis are numbered with the amount of noise added in the image and Y -axis are numbered with respective feature score of the noisy image

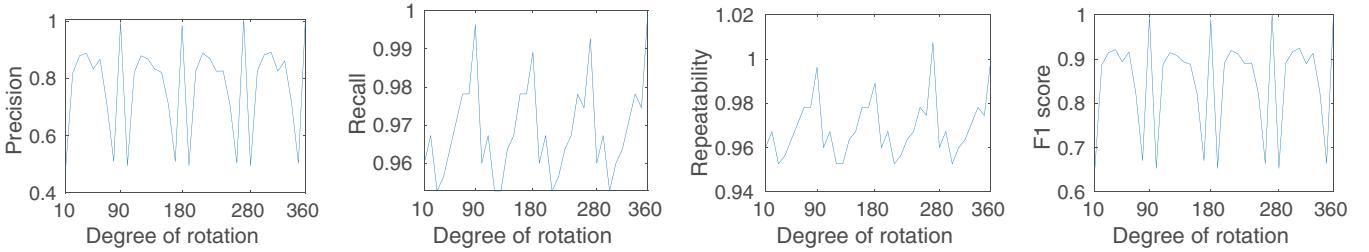


FIGURE 9 Precision, recall, repeatability and F1 score obtained from SURF for rotated version of an image. X-axis is the degrees of rotation in the image and Y-axis is feature score of the rotated image

were considered for feature extraction. 125 high quality natural images were chosen as the pristine images. Mean statistical NSS features for those images were extracted and stored. Now, for the given test image the statistical NSS features are obtained in a similar way. At the end, the NIQE score is calculated as the difference between NSS features of pristine and test image using Equation (3).

$$\text{Difference}(V_{pri}, V_{te}, \Sigma_{pri}, \Sigma_{te}) = \sqrt{(V_{pri} - V_{te})^T \left(\frac{\Sigma_{pri} + \Sigma_{te}}{2} \right)^{-1} (V_{pri} - V_{te})}, \quad (3)$$

where V_{pri} , V_{te} and Σ_{pri} , Σ_{te} are the mean vectors and covariance matrices of multivariate Gaussian (MVG) model for the pristine and test images, respectively. Good quality images will have very less difference whereas quality degraded images will have huge amount of difference. The main advantage of NIQE is, it is completely blind to compute the quality score of an image and therefore it does not require any reference images. We generated the NIQE score for the images from the MATLAB package. For our experiments, the features of pristine images are obtained from the mat file provided along with the NIQE code.

The next set of features is extracted from NSCT. The NSCT is a shift-invariant contourlet transform. It is the combination of non-subsampled pyramids (NSP) and non-subsampled directional filter bank (NSDFB). With the properties of shift-invariance and multidirectional, the NSCT is used in the image segmentation algorithm [38]. The segmentation exploits the spatial correlation between the neighboring pixels. This correlation is obtained from NSP and NSDFB of NSCT [39]. Features from the high and low bandpass filters in the pyramid help to obtain the frequency decomposition of an image at a different scale. NSCT decomposition of an image at level L gives L bandpass and a lowpass subband filter informations. Each L bandpass subband can be further decomposed to any scale between 1 to L. The decomposition of an image [40] is given in the equation (4).

$$SB(a, b) = \left\{ SB_i(a, b) \mid i = 1, 2, \dots, \left(\sum_{level=1}^L 2^{N_{level}} + 1 \right) \right\}, \quad (4)$$

where L is the total number of levels in the decomposition. $1 \leq a \leq H$, $1 \leq b \leq W$, H and W are the height and width of an image. Further decomposition of the bandpass filter at each level is $\sum_{level=1}^L (2^{N_{level}} + 1)$. The value of N ranges from 1 to L (different values of N at each level is accepted). Multi-scale and the directional decomposition are independent at each level (level of decomposition to be any arbitrary power of two's). With the help of MATLAB package, we obtained the NSCT coefficients of an image at the scale factor 5. We computed the mean for each direction coefficient. The size of the feature vector is (1×16).

By concatenating the above features, the feature vector size for the proposed system is (1×53) where 32 are from SURF, 1 from NIQE score, 4 from AGGD and 16 from NSCT.

3.2 | Generation of label

Initially, the IQA model has been trained with a human opinion score as the label. In recent days, the researchers have started to design a model to work on opinion-unaware NRIQA models. These opinion-unaware models are labeling an image with an algorithm instead of a human opinion score [41]. Likewise, here the goal is to predict the quality of an image in terms of image segmentation, we need to label the image by the best score among the 12 best-known segmentation algorithms, not by the human opinion score. If we train the model with the score given by any random image segmentation algorithm then the proposed method will be proportional to the performance of that segmentation algorithm. University of California, Berkeley has created BSD, which contains the dataset and 12 algorithms for image segmentation. Dataset is provided with the rankings for the test images based on how well the algorithms can find the boundaries in that images [31]. The rank is based on the F1 score of an image where the highest score has a rank one and so on. Higher the F1 score better the segmentation of an image. Thus BSD has the 12 scores for each test image generated by 12 different algorithms. Thus the best F1 score of an image is chosen from the 12 scores given by 12 algorithms, choosing the best score. Hence we consider this best score as a label for the corresponding image to train the proposed model.

3.3 | Training and testing

In the proposed method, support vector regression (SVR) [42] is trained with features that are computed from the image (as discussed in the feature extraction phase) along with its corresponding best F1 score. Once the training is done, we can use the model for testing. As shown in Figure 5 the testing can be done by computing the feature vector for the test image and passed as the input to the trained SVR to obtain the score for an image. The reason for choosing the model is the symmetrical loss function and it is efficient to predict the values in real-time. SVR has been predominantly used in IQA [11, 43, 44].

4 | EXPERIMENTAL RESULTS

The performance of the proposed method with a BSD dataset is shown in the tables. To show the proposed method is not distortion specific we experimented with four different types of distortions. The performance is evaluated in terms of consistency in efficiency of the system and in algorithm independency. We correlated the results of the proposed method along with results of some of the existing segmentation algorithms. Pearson correlation coefficient (PLCC) and mean squared error (MSE) are the metrics used to measure the performance of the proposed method. Correlation between the predicted GO-IQA score and the segmentation score of several segmentation algorithms are considered. We ideally expect that the predicted GO-IQA score of the proposed model should correlate well with the best segmentation algorithm, slightly low with second best algorithm and even more low with third best algorithm and so on. The results obtained by the proposed model meets our expectation.

4.1 | Database description

BSD 300 [45] dataset contains 200 train images and 100 test images along with the human segmentation score as a ground truth. BSD images are natural images including birds, forests, animals, humans, temples, buildings etc. For experimental purposes, grayscale images are chosen and the corresponding segmentation scores are considered as the label for the images. Features for GO-IQA are extracted from the grayscale images. BSD dataset has color images, these color images are converted to grayscale images. Segmentation scores for the color and grayscale images are available in their website [31]. (For better visualization purpose, color images are used as examples in this paper)

4.2 | Evaluation on BSD300 dataset

The performance of proposed method on the BSD300 dataset is shown in Table 1. We trained the method with 75% of images and tested with remaining 25% of images in BSD300. We create a feature vector for each image as discussed in Section 3.1. For each image, the label considered is the maximum score

TABLE 1 Performance of GO-IQA with BSD300 dataset

Dataset/correlation	PLCC	MSE
BSD300	0.837	0.003

among the scores given by those 12 segmentation algorithms reported in BSD [31]. While testing the model, the correlation is calculated between the GO-IQA predicted segmentation score of the test images and the best score among 12 segmentation scores provided by respective 12 algorithms for those test images. Generally, in IQA, the human opinion score is used for labeling the images. But in GO-IQA for image segmentation, image segmentation score is used for labeling the image. This is because the motivation for IQA is to measure the visual quality and hence IQA models will be trained with the human opinion score. But in GO-IQA, score is measured for the specific goal, here we considered image segmentation as goal, thus it is trained with segmentation score.

Table 1 shows the performance of GO-IQA. The 10-fold cross validation has been performed and the average of the 10-fold metrics have been reported in the table. In 10-fold cross validation, for each fold (run) training and testing datasets (75% and 25% of images) are randomly chosen disjoint sets from the database. Correlation between the predicted GO-IQA scores and best scores of the segmentation algorithms from BSD benchmark dataset is 0.837 in terms of PLCC score and obtained the MSE of 0.003 for the test images.

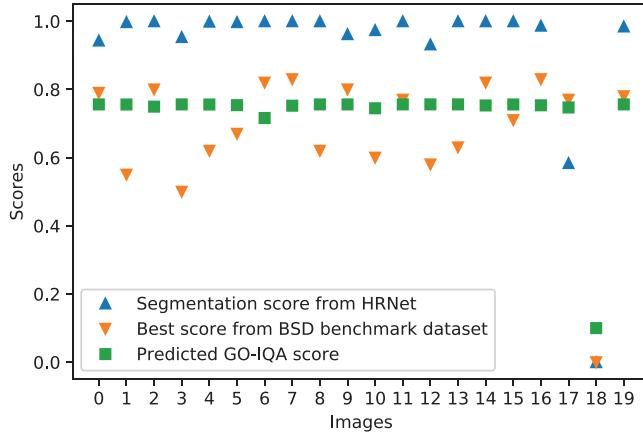
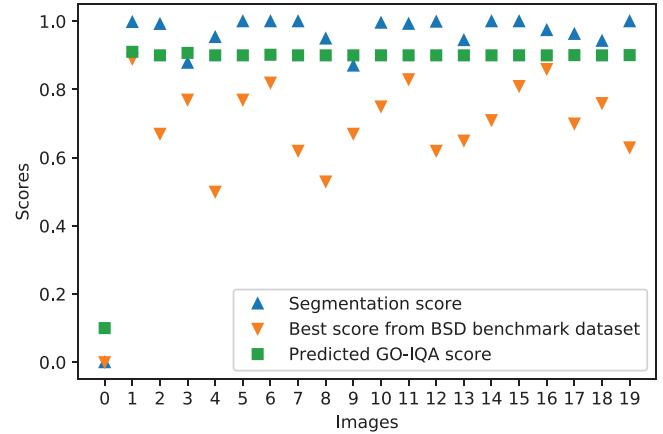
4.3 | Algorithm independency test

To ensure GO-IQA is not skewed to the algorithm scores which are used to train the model, the algorithm independency test has been performed on state-of-the-art HRNet segmentation algorithm [46] on BSD dataset. Two different sets of experiments have been conducted to validate the GO-IQA. In experiment 1, GO-IQA has been trained with the segmentation scores obtained by HRNet segmentation algorithm. In the testing phase, predicted GO-IQA scores are correlated with both the best segmentation scores from BSD benchmark dataset as well as with segmentation scores obtained by HRNet segmentation algorithm. In experiment 2, GO-IQA is trained with the best algorithm score for those images among 12 best-known segmentation algorithms as listed in the website [31]. Similar to the experiment 1, in the testing phase, predicted GO-IQA scores are correlated with both the best segmentation scores from BSD benchmark dataset as well as with segmentation scores obtained by HRNet segmentation algorithm.

Table 2 shows the results of the proposed method on both experiments. In segmentation, performance of HRNet segmentation algorithm is better than the algorithms considered in the website [31]. So, GO-IQA score is ideally expected to be well correlated with HRNet when compared to the other algorithms considered in [31]. The results meet the expectation. When GO-IQA is trained with best segmentation scores of BSD, PLCC value between the prediction scores of

TABLE 2 Performance of GO-IQA in algorithm independency test

Segmentation score used for training	Segmentation score used for correlation	PLCC	MSE
HRNet segmentation score [46]	HRNet segmentation score	0.9839	0.0067
	Best segmentation score from BSD benchmark dataset	0.8599	0.0365
Best segmentation score from BSD benchmark dataset [31]	Best segmentation score from BSD benchmark dataset	0.8745	0.0090
	HRNet segmentation score	0.9378	0.0629

**FIGURE 10** Scatter plot for GO-IQA model trained with best score from BSD benchmark dataset**FIGURE 11** Scatter plot for GO-IQA model trained with HRNet segmentation score

GO-IQA and the scores of HRNet is 0.93, whereas PLCC value between the predictions scores of GO-IQA and best scores reported in BSD is 0.87. It is also observed that the prediction scores of GO-IQA trained with HRNet segmentation scores is very well correlated with the scores obtained by HRNet segmentation algorithm for test images(PLCC values 0.98). Whereas the PLCC value between prediction scores of GO-IQA trained with HRNet segmentation scores and best segmentation scores from BSD is 0.85. Top 20 scores are obtained from the 50 folds and the average of those 20 scores are reported in Table 2.

Figures 10 and 11 show the scatter plot for predicted GO-IQA scores, HRNet segmentation score and best segmentation score from BSD benchmark dataset for a single run. For the same run, the PLCC value between GO-IQA score and best score obtained from BSD is 0.82, and 0.91 is the PLCC value between GO-IQA prediction scores and HRNet segmentation score. In X axis, indices for the test images are given and in Y axis, the scores for the corresponding images are given.

The same can be observed from Figure 10, where the best scores from BSD benchmark dataset are scattered in the plot whereas predicted GO-IQA scores and HRNet segmentation scores are aligned. Similarly, the same trend is followed in the Figure 11 when the GO-IQA model is trained with HRNet segmentation score. In this case, correlation between the HRNet segmentation scores and predicted GO-IQA score is 0.98 in terms of PLCC. PLCC value is 0.83 for correlation between the

TABLE 3 Validating the proposed method with pre existing segmentation algorithm

Score used for training/ correlation	PLCC	MSE
Second best score	0.824	0.003
Third best score	0.822	0.004
Fourth best score	0.816	0.004

predicted GO-IQA score and best score obtained from BSD benchmark dataset.

4.4 | Generalization capability of the model

To further validate GO-IQA, another experiment has been performed to ensure that the GO-IQA will perform well with the newer (better) methods that may be developed for the task of segmentation in future. To validate that, GO-IQA is trained with the second best segmentation score of an image as a label. Predicted scores during the testing phase are correlated well with the best segmentation score. Similarly we trained with the third and fourth best segmentation score of an image, and then correlated the score with the best score of an image. Table 3 gives the experimental results. This experimental study shows that the system will perform well with the new segmentation algorithms in future.



FIGURE 12 Gaussian blur versions of an image created with zero mean and the amount of variance is varied as 3.1, 4.1 and 6.1 from left to right. The segmentation scores obtained from the gPb algorithm (Alg 1) [47] for the images are 0.671, 0.627 and 0.556. The predicted image segmentation scores are 0.63, 0.66 and 0.63, respectively



FIGURE 13 Contrast versions of an image generated by varying the values of maximum luminosity span as 100, 150 and 190 from left to right. The segmentation scores obtained from the gPb algorithm (Alg 1) [47] for the images are 0.758, 0.734 and 0.735. The predicted image segmentation scores are 0.728, 0.720 and 0.723, respectively

4.5 | Distortion specific test

Distortion specific test is conducted to ensure the performance of the method in assessing the segmentation quality of an image even if it is distorted by various noise. Experiment has been conducted for 40 noisy images on GO-IQA model. For generating noisy images, we added four different types of noise to the image, namely white Gaussian noise, salt and pepper noise, speckle noise, and also image with varying contrast have been considered. Ten Gaussian blur versions of an image are created with zero mean and with variance from 0.1 to 9.1 (with the step size of 1). Ten different contrast versions of an image are generated by varying the maximum luminosity span ranges from 100 to 190. The third set of ten images are generated by varying the noise density from 0.05 to 0.45 in the salt and pepper noise. The last set of images are generated by adding the speckle noise to an image by varying the variance of uniformly distributed random noise from 1 to 1.5 (with the step size of 0.05) and keeping the mean as zero. Thus 40 distorted images are created, 10 for each noise. Figures 12–15 give the noisy versions of some images along with their segmentation F1 score and predicted segmentation score. The obtained PLCC and MSE score are listed in the Table 4. From the Table 4, it can be observed that the GO-IQA model is robust to noise.

4.6 | Consistency of efficiency of the proposed method

In this section, we discussed the consistency of efficiency of the proposed method with the comparison of our method with various image segmentation algorithms. Table 5 shows that the proposed method has a high correlation with the good algorithms and less correlation with the algorithm whose performance is not up to the level. For this test, we used the BSD300 dataset. The best segmentation algorithm score of an image is used to label the input image. For testing, we used the scores of the segmentation algorithm for the images. We considered all the benchmark segmentation algorithms on the website [31]. The segmentation algorithms listed in the Table 5 are ranked by their performance.

The coefficients of correlation between the predictions of proposed GO-IQA and predictions of each of those 12 algorithms are presented in the second column. And also the mean squared error between the predictions of the proposed and each of those algorithms is given in column 3. Ideally, the correlation with predictions of algorithm i is expected to be more than that of with algorithm j iff $i < j \forall 1 \leq i, j \leq 12$. The expectation is almost satisfied. The correlation with algorithm 12 is the least as it is the poorest algorithm among those 12 chosen algorithms.



FIGURE 14 Salt and pepper noise versions of an image are created by varying the noise density 0.05, 0.2 and 0.45 from left to right. The segmentation scores obtained from the gPb algorithm (Alg 1) [47] for the images are 0.728, 0.727 and 0.716. The predicted image segmentation scores are 0.637, 0.635 and 0.615, respectively

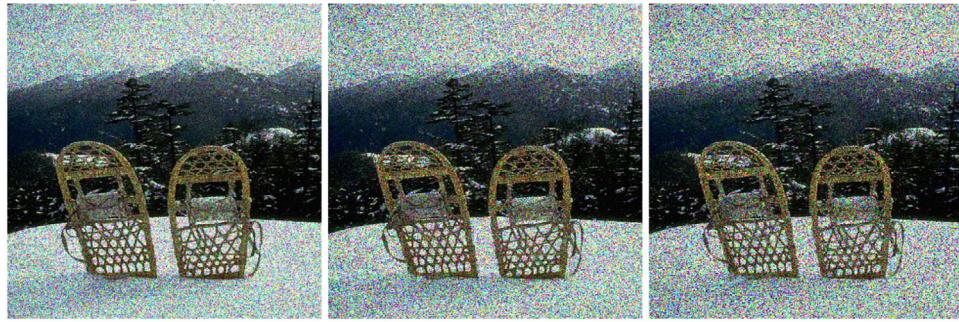


FIGURE 15 Speckle noise versions of an image generated by uniformly distributed random noise with mean zero and variance as 1, 1.3 and 1.45 from left to right. The segmentation scores obtained from the gPb algorithm (Alg 1) [47] for the images are 0.723, 0.708 and 0.706. The predicted image segmentation scores are 0.678, 0.681 and 0.691, respectively

GO-IQA segmentation scores are well correlated with the top three segmentation algorithm scores. An average correlation with the algorithms ranged from rank four to nine and even less correlation with the least ranking algorithms is observed from the experiment and listed the values of performance metric in the Table 5.

5 | TIME COMPLEXITY

The core of SVM is to solve quadratic programming, sequential minimal optimization (SMO) is used to solve the optimization problem. The time complexity of SMO is $O(N^{2.2})$, where N is the number of training samples, assuming the length of the feature vector is constant [53]. Thus the time complexity for

training SVM with K epochs is $O(K * N^{2.2})$. For testing, the time taken for SVM is $O(P)$, considering constant length feature vector , where P is the number of test samples.

The empirical time complexity of the proposed method is discussed here. Time taken for the construction of

TABLE 5 Performance of proposed GO-IQA for segmentation task, considering 12 standard algorithms on BSD300 dataset

List of segmentation algorithms/correlation	PLCC	MSE
Alg 1 [47]	0.801	0.0038
Alg 2 [48]	0.804	0.0038
Alg 3 [49]	0.808	0.0038
Alg 4 [50]	0.777	0.0035
Alg 5 [51]	0.756	0.0038
Alg 6 [31]	0.748	0.0039
Alg 7 [31]	0.708	0.0038
Alg 8 [31]	0.774	0.0038
Alg 9 [31]	0.699	0.0039
Alg 10 [31]	0.693	0.0038
Alg 11 [31]	0.688	0.0038
Alg 12 [52]	0.673	0.0038

TABLE 4 Performance on distortion specific test

Type of noise/correlation	PLCC	MSE
Blur	0.810	0.020
Contrast	0.989	0.003
Salt and pepper noise	0.818	0.026
Speckle noise	0.988	0.002

feature vector for an image of the size 321*481 is 53.626s, time taken for training the model with the feature vector for 100 images is 0.0115s and for testing an image takes 0.0018s. The computing environment is Ubuntu(*Intel(R)Core™ i7 – 7500UCPU@2.70GHz × 4, 8.00GB, RAM*).

6 | CONCLUSION AND FUTURE WORK

In this paper, a new paradigm of image quality assessment, called goal oriented image quality assessment has been proposed. Image segmentation has been considered as the goal here. The goodness of image (GO-IQA score) for segmentation is predicted using SVR. To validate the model, the goodness scores for a sequence of images have been computed, and that sequence of scores has been found to be well correlated with the segmentation performance (F1 Score) scores of the corresponding images by the best known segmentation algorithm. It is also observed that correlations of GO-IQA scores of a sequence of images with the segmentation performance scores of other standard algorithms are proportional to the efficacy of the segmentation algorithm. We obtained the PLCC value 0.98 when GO-IQA prediction scores and the scores of the state of the art segmentation algorithm are correlated, and the MSE of 0.006 between the scores of the proposed model and scores of the state of the art method. This observation helps us to conclude that if a new image segmentation algorithm outperforms all the existing segmentation algorithms, then the correlation of GO-IQA scores of sequence of images with the segmentation performance scores of that algorithm of that sequence of images will be higher. It can be inferred that if GO-IQA score is less than any arbitrary segmentation algorithm can not segment it well. This prediction of goodness of image for performing a particular task (goal) helps the research community to understand the input data for the task. The algorithm performed well in terms of PLCC and MSE but the Spearman rank correlation coefficient has suffered. In the future work, the feature selection can be performed to improve the prediction results. The GO-IQA can also be extended to predict the scores for different goals such as depth map computation, face recognition, and object detection etc.

REFERENCES

- Mertz, P., et al.: Quality rating of television images. *Proc. IRE* 38(11), 1269–1283 (1950)
- Samani, A., et al.: Transform domain measure of enhancement tdme for security imaging applications. In: 2013 IEEE International Conference on Technologies for Homeland Security HST, pp. 265–270. IEEE, Piscataway (2013)
- Anchini, R., et al.: Metrological characterization of a vision-based measurement system for the online inspection of automotive rubber profile. *IEEE Trans. Instrum. Meas.* 58(1), 4–13 (2008)
- Wang, S., et al.: Adaptive watermarking and tree structure based image quality estimation. *IEEE Trans. Multimedia* 16(2), 311–325 (2013)
- Panetta, K., et al.: Choosing the optimal spatial domain measure of enhancement for mammogram images. *Int. J. Biomed. Imaging* 2014, 937849 (2014), <https://doi.org/10.1155/2014/937849>
- Wang, Z., et al.: Why is image quality assessment so difficult? In: Proc 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing 4, pp. IV-3313–IV-3316 (2002)
- Wang, Z., Bovik, A.C.: Reduced-and no-reference image quality assessment. *IEEE Signal Process Mag.* 28(6), 29–40 (2011)
- Wang, Z., et al.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13(4), 600–612 (2004)
- Liang, Y., et al.: Image quality assessment using similar scene as reference. In: Proceedings of European Conference on Computer Vision, pp. 3–18. Springer, Vham (2016)
- Sheikh, H.R., et al.: No-reference quality assessment using natural scene statistics: Jpeg2000. *IEEE Trans. Image Process.* 14(11), 1918–1927 (2005)
- Saad, M.A., et al.: Blind image quality assessment: a natural scene statistics approach in the dDCT domain. *IEEE Trans. Image Process.* 21(8), 3339–3352 (2012)
- Moorthy, A.K., Bovik, A.C.: A two-step framework for constructing blind image quality indices. *IEEE Signal Process Lett.* 17(5), 513–516 (2010)
- Qin, M., et al.: Hybrid NSS features for no-reference image quality assessment. *IET Image Proc.* 11(6), 443–449 (2017)
- Gu, K., et al.: Using free energy principle for blind image quality assessment. *IEEE Trans. Multimedia* 17(1), 50–63 (2014)
- Liu, H., et al.,: Enhanced image no-reference quality assessment based on colour space distribution. *IET Image Proc.* 14(5), 807–817 (2020)
- De, K., Masilamani, V.: No-reference image quality measure for images with multiple distortions using random forests for multi method fusion. *Image Analysis Stereology* 37(2), 105–117 (2018)
- Kanjar, D., Masilamani, V.: No-reference image contrast measure using image statistics and random forest. *Multimedia Tools Appl.* 76(18), 18641–18656 (2017)
- Wang, X., et al.: Blind image quality assessment for measuring image blur. In: Proceedings of 2008 Congress on Image and Signal Processing 1, pp. 467–470. IEEE, Piscataway (2008)
- Samani, A., et al.: Quality assessment of color images affected by transmission error, quantization noise, and noneccentricity pattern noise. In: Proceedings of 2015 IEEE International Symposium on Technologies for Homeland Security (HST), pp. 1–6. IEEE, Piscataway (2015)
- Gui-Feng, Z., et al.: No-reference aerial image quality assessment based on natural scene statisticsand color correlation blur metric. In: Proceedings of 2016 IEEE PES 13th International Conference on Transmission & Distribution Construction, Operation & Live-Line Maintenance (ESMO). pp. 1–4. IEEE, Piscataway (2016)
- Appina, B., et al.: Study of subjective quality and objective blind quality prediction of stereoscopic videos. *IEEE Trans. Image Process.* 28(10), 5027–5040 (2019)
- Wu, L., et al.: Fuiqa: fetal ultrasound image quality assessment with deep convolutional networks. *IEEE Trans. Cybern.* 47(5), 1336–1349 (2017)
- Outtas, M., et al.: Evaluation of no-reference quality metrics for ultrasound liver images. In: Proceedings of 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–3. IEEE, Piscataway (2018)
- Gu, K., et al.: Model-based referenceless quality metric of 3d synthesized images using local image description. *IEEE Trans. Image Process.* 27(1), 394–405 (2018)
- Shaikh, S.H., et al.: Moving object detection approaches, challenges and object tracking. In: Moving Object Detection Using Background Subtraction, pp. 5–14. Springer, Cham (2014)
- Zangeneh, E., et al.: Low resolution face recognition using a two-branch deep convolutional neural network architecture. *Expert Syst. Appl.* 139, 112854 (2020)
- Mittal, A., et al.: Making a completely blind image quality analyzer. *IEEE Signal Process Lett.* 20(3), 209–212 (2012)
- Lee, J.Y., Nandi, A.: Parameter estimation of the asymmetric generalised gaussian family of distributions. In: Proceedings of IEE Colloquium on Statistical Signal Processing (Ref No 1999/002), pp. 9–1. IET, Stevenage (1999)

29. Da Cunha, et al.: The nonsubsampled contourlet transform: theory, design, and applications. *IEEE Trans. Image Process.* 15(10), 3089–3101 (2006)
30. Bay, H., et al.: Surf: speeded up robust features. In: *Proceedings of European Conference on Computer Vision*, pp. 404–417. Springer, Cham (2006)
31. The Berkeley Segmentation Dataset and Benchmark, <https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/bench/html/images.html> (accessed: December 2019)
32. Vijayaraju, P., et al.: Hybrid memetic algorithm with two-dimensional discrete haar wavelet transform for optimal sensor placement. *IEEE Sens. J.* 17(7), 2267–2278 (2017)
33. Rao, K.R., Ben-Arie, J.: Optimal edge detection using expansion matching and restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* 16(12), 1169–1182 (1994)
34. Wang, X.: Moving window-based double haar wavelet transform for image processing. *IEEE Trans. Image Process.* 15(9), 2771–2779 (2006)
35. Wang, G., et al.: Blind quality metric of dibr-synthesized images in the discrete wavelet transform domain. *IEEE Trans. Image Process.* 29, 1802–1814 (2019)
36. Tello, M., et al.: A novel algorithm for ship detection in sar imagery based on the wavelet transform. *IEEE Geosci. Remote Sens. Lett.* 2(2), 201–205 (2005)
37. Strickland, R.N., Hahn, H.I.: Wavelet transform methods for object detection and recovery. *IEEE Trans. Image Process.* 6(5), 724–735 (1997)
38. Xin, F., et al.: Unsupervised image segmentation based on nonsubsampled contourlet hidden markov trees model. In: *Proceedings of 2009 2nd Asian-Pacific Conference on Synthetic Aperture Radar*, pp. 485–488. IEEE, Piscataway (2009)
39. Prakash, R.M., Kumari, R.S.S.: Nonsubsampled contourlet transform based expectation maximization method for segmentation of images. In: *Proceedings of 2012 International Conference on Machine Vision and Image Processing (MVIP)*, pp. 137–140. IEEE, Piscataway (2012)
40. Kanjar, D., Masilamani, V.: Image quality assessment for blurred images using nonsubsampled contourlet transform features. *J. Comput.* 12(2), 156–164 (2017)
41. Gu, K., et al.: No-reference quality assessment of screen content pictures. *IEEE Trans. Image Process.* 26(8), 4005–4018 (2017)
42. Drucker, H., et al.: Support vector regression machines. In: *Proceedings of Advances in Neural Information Processing Systems*, pp. 155–161. MIT Press, Cambridge (1997)
43. Narwaria, M., Lin, W.: Objective image quality assessment based on support vector regression. *IEEE Trans. Neural Networks* 21(3), 515–519 (2010)
44. Kanjar, D., Masilamani, V.: A no-reference image quality measure for blurred and compressed images using sparsity features. *Cognit. Comput.* 10(6), 980–990 (2018)
45. Martin, D., et al.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *Proceedings of 8th International Conference on Computer Vision*, vol. 2, pp. 416–423. IEEE, Piscataway (2001)
46. Wang, J., et al.: Deep high-resolution representation learning for visual recognition. in *IEEE Trans. on Pattern Anal. Mach. Intell.* (2020), <https://doi.org/10.1109/TPAMI.2020.2983686>
47. Arbelaez, P., et al.: Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33(5), 898–916 (2010)
48. Xiaofeng, R., Bo, L.: Discriminatively trained sparse code gradients for contour detection. In: *Proceedings of Advances in Neural Information Processing Systems*, pp. 584–592. MIT Press, Cambridge (2012)
49. Maire, M., et al.: Using contours to detect and localize junctions in natural images. In: *Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8. IEEE, Piscataway (2008)
50. Ren, X.: Multi-scale improves boundary detection in natural images. In: *Proceedings of European Conference on Computer Vision*, pp. 533–545. Springer, Cham (2008)
51. Dollar, P., et al.: Supervised learning of edges and object boundaries. In: *Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 2, pp. 1964–1971. IEEE, Piscataway (2006)
52. Yu, S.X.: Segmentation induced by scale invariance. In: *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 444–451. IEEE, Piscataway (2005)
53. Platt, J.: Sequential minimal optimization: a fast algorithm for training support vector machines, MSR-TR-98-14, 1998

How to cite this article: Kiruthika, S., Masilamani, V.: Goal oriented image quality assessment. *IET Image Process.* 16, 1054–1066 (2022).
<https://doi.org/10.1049/ipt2.12209>