BLIND IMAGE QUALITY METRIC FOR BLACKBOARD LECTURE IMAGES

Ali Shariq Imran, Faouzi Alaya Cheikh

Gjøvik University College
P.O.Box-191, N-2802, Gjøvik, Norway
phone: + (47) 96822072, email: ali.imran@hig.no

ABSTRACT

This paper proposes a reference free perceptual quality metric for blackboard lecture images. The text in the image is mostly affected by high compression ratio and de-noising filters which cause blocking and blurring artifacts. As a result the perceived text quality of the blackboard image degrades. The degraded text is not only difficult to read by humans but it also makes the optical character recognition task even more difficult. Therefore, we put our effort firstly to estimate the presence of these artifacts and then we used it in our proposed quality metric. The blocking and blurring features are extracted from the image content on block boundaries without the presence of reference image. Thus it makes our metric reference free. The metric also uses the visual saliency model to mimic the human visual system (HVS) by focusing only on the distortions in perceptually important regions, i.e. those regions which contains the text. Moreover psychophysical experiments are conducted that show very good correlation between the mean opinion score and quality scores obtained from our reference free perceptual quality metric (RF-PQM). The correlation results are also compared with standard reference and reference free metric.

1. INTRODUCTION

There are two common ways to measure the quality of the images under observation. First by conducting subjective tests and the second by objective metric. Subjective test which is also known as psychophysical experiment involves human viewers. The viewers assess the quality of the image based on ITU-R BT recommendations. While in objective metric, the algorithm tries to assess the quality of the image. Most of the objective assessment methods try to assess the quality of the degraded image by using some sort of reference. The reference can either be original undistorted signal source or some extracted features from the original signal, to be used later. The objective metric which use original signal as a reference are known as full reference metric while those that rely on some extracted features are called reduced reference metrics.

There are many situations in which the original signal is not present. In case of images, the signal can be corrupted either because of noise over communication channel or by the introduction of some coding artifacts as a result of compression. To assess the quality of such type of images where a reference image is not present, reference free quality metrics are used. These metrics rely on finding some known artifacts to estimate the quality of the image. Wang and Alan proposed the initial work for finding the blocking artifacts in [1]. Recently, different techniques have been adopted to find these artifacts based on edge sharpness level [2], contrast similarity [3], geometric moments [4] and based on other spatial features [5]. These metrics however do not account for salient regions.

The rest of this paper is organized as follows: Section 2 gives an overview for the need of a new quality metric. Section 3 shows the work flow which is followed by proposed model in section 4. Section 5 shows the experimental setup and in section 6 comparisons are made with other metrics. The last section gives the conclusion.

2. MOTIVATION

Now-a-days many learning system use images and videos as part of their educational program. The educational content usually consists of handwritten text, power point slides, figures and graphs. Standard coding mechanisms introduce certain artifacts that highly affects the visual quality of the content. This ultimately leads to poor readability of text. Blocking and blurring are the two common artifacts caused by compression standards and de-noising filters. Psychophysical experiment show that the presence of these two types of artifacts highly degrade the readability of the text. Thus the demand for quality evaluation mechanism of these degraded content become inevitable.

Moreover, the text is the core part of the content on these educational media and is most salient to attract visual attention. It is also widely accepted that under normal condition human eye tends to follow visually salient regions [6]. Thus visual attention plays an important role in determining the perceived image quality. On the other hand reference free objective quality metrics can play an important role in evaluating the quality. Perkis combined the attention model for saliency with perceptual quality in [7]. However the metrics that are available today do not incorporate text as salient regions. Which makes them less desirable to use for such widely available media. Therefore the need for such an objective metric is desirable. Hence, we propose a visual attention based RF-PQM focusing on text as salient region.

3. WORK FLOW

The work flow of our proposed model is shown in Figure 1. The distorted images are processed by our reference free quality metric. The metric estimates the amount of blurring and blocking artifacts present in the image and give a quality score. The same set of images are subjected to psychophysical experiment where expert and non-expert viewers rate the quality of the image on a scale of 1-5. The mean opinion score (MOS) is obtained from the subjective experiment. The quality score obtained from the reference free metric is then correlated with mean opinion score. Some of the distorted images were used as training dataset for quality prediction model (QPM). Three correlation parameters

$a = -0.24$, $b = -0.16$ and $c = 0.06$ were obtained as a result of non-linear regression routine during training process. These parameters were then used in the QPM.

4. PROPOSED MODEL

The computation of the RF-OQM consists of 5 major processing steps. In the first step, the saliency map is computed followed by visibility map. Saliency map consists of blocks containing text regions while the visibility map contains regions where the artifacts may be present. From the visibility map, the blocking and blurring maps are generated. The blocking map contains block features, from which the blocking score is obtained. While the blurring map contains blur features from which the blurring score is obtained. Lastly the blocking score and the blurring score is combined in the QPM to obtain overall quality score. The processing steps are as follows:

4.1 Saliency Map

Our proposed quality metric tries to mimic the human perceptual mechanism by introducing the saliency map. The traditional bottom up saliency models [8] use color, intensity and orientation as features to compute the salient regions. These however fail when applied to lecture images with text because of minor variation in color and intensity. To be able to detect text as salient regions, we have used the difference of gaussian (DoG) for identifying text regions from those of background and noise in form of chalk dust. Following values of $sigma_1 = 8.0$ and $sigma_2 = 0.5$ give best results for all the dataset images. The window size was chosen to 27 to accurately identify text as salient region. Figure 2 shows the 8×8 representation of saliency and visibility map.

In future, this method will be replaced by a recent saliency model involving contour integration [9] for creating saliency maps. This model shows that human attention mechanisms also include the contour linking mechanism. Such mechanisms should enhance contours against background noise, and are thus very relevant to the comprehension of text documents.

4.2 Visibility Map

The next step is to identify artifact features from the saliency map. To do this, we have computed the visibility map (VM) by analyzing the local contrast as in [10]. The local contrast is given as:

$$C(i,j) = \frac{|S_M(i,j) - S_{M.avg}(i,j)|}{S_{M.avg}(i,j)} \tag{1}$$

Where $S_M(i,j)$ is the luminance of the pixel (i,j) and $S_{M.avg}(i,j)$ is the average luminance of the eight neighbors of (i,j). The pixel is visible if the contrast is higher then just noticeable contrast (JNC)[10]. The visibility map is then computed as:

$$VM(i,j) = \begin{cases} 1, & if\ C(i,j) \geq JNC(i,j) \\ C(i,j), & otherwise \end{cases} \tag{2}$$

4.3 Blocking Map

Blocking artifact is introduced by compression standards. Most of the standard uses 8×8 pixel blocks for quantization that creates artifacts in compressed image at the edges of these blocks. To extract blocking map from the visibility map we use same 8×8 pixel blocks. The extrapolated discontinuity between neighboring block is calculated for this purpose: first across vertical blocks and then horizontal.

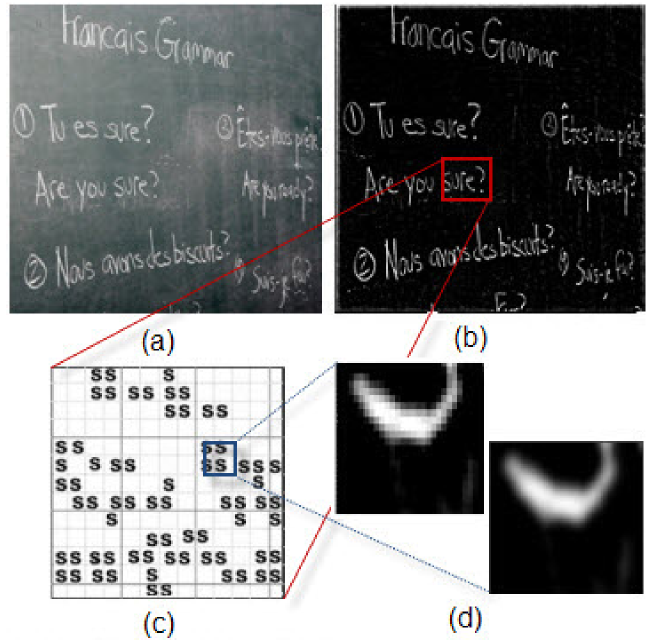

Figure 2: Representation of text regions. (a) original image (b) saliency image (c) saliency map representation (d) visibility map; blocking and blurring artifacts.

Let B_{ij} is an 8×8 block of pixels from the salient regions of the image.

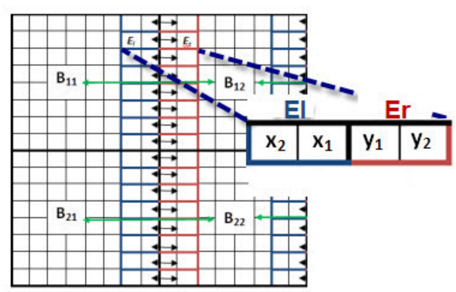

Figure 3: Representation of pixel values for blocking artifacts.

The value of the blocking artifact Blc_v across two horizontally adjacent blocks B_{11} and B_{12}, as illustrated in Figure 3, represents a measure of the discontinuity at the vertical boundary between the two blocks. This value is computed in the following way, first the vertical discontinuity is evaluated for each line across the two blocks. This vertical discontinuity is computed as the absolute difference of the two extrapolated values, E_l and E_r, across the boundaries of two adjacent blocks. E_l and E_r are calculated using first order extrapolator given as:

$$E_l = \frac{3}{2} * x_1 - \frac{1}{2} * x_2 \tag{3}$$

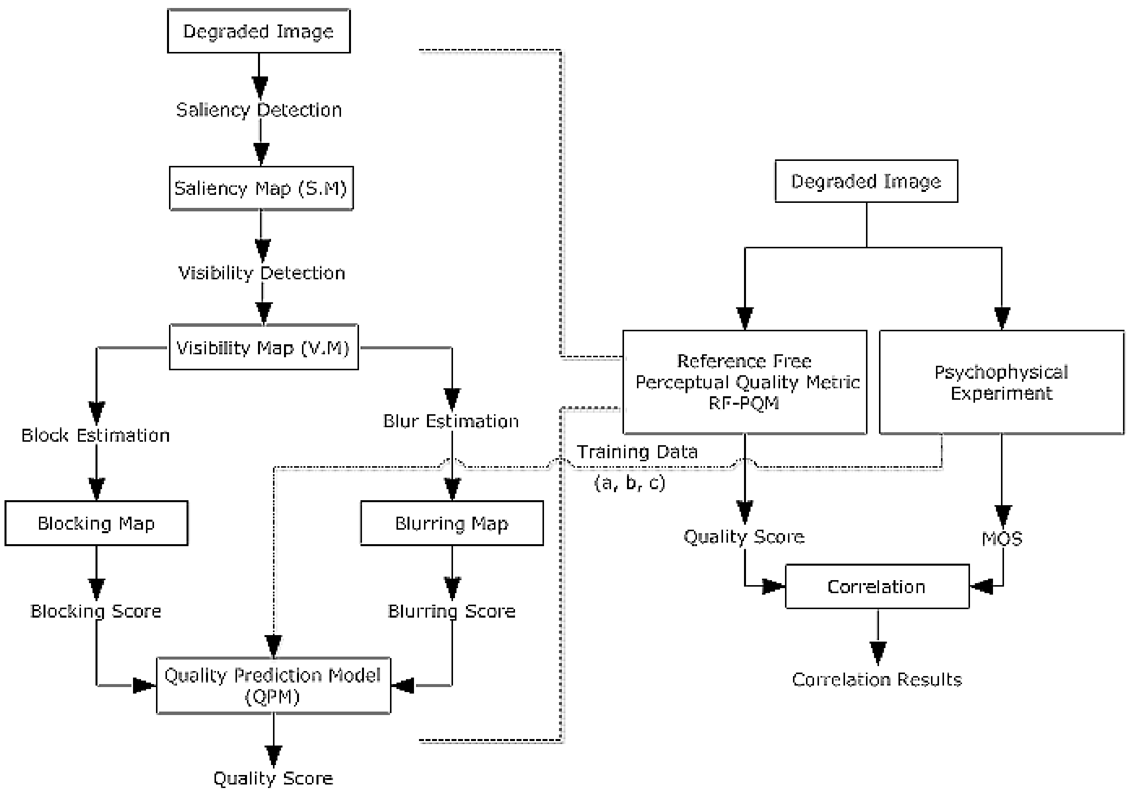

Figure 1: An overview of work flow.

$$E_r = \frac{3}{2} * y_1 - \frac{1}{2} * y_2 \quad (4)$$

Where x_1, x_2 and y_1, y_2 are the pixel values at the boundary of the blocks as illustrated in Figure 3. E_l and E_r are the left and right extrapolated values.

The vertical artifact value is the mean of the eight discontinuities within a single block.

$$Blc_v = \frac{1}{8} \sum_{j=0}^{7} \left| (E_r)_j - (E_l)_j \right| \quad (5)$$

Where $(E_r)\,j$ is the i^{th} row extrapolated values. The values for the horizontal artifacts can be calculated in similar fashion.

4.3.1 *Blocking Score*

A blocking score can be estimated by summing up the vertical and horizontal blocking values.

$$B_S = Blc_v + Blc_h \quad (6)$$

4.4 Blurring Map

Blur normally occurs as a result of de-blocking filter and quantization process which removes the high frequency information from an image. Blur can be hard to estimate in regions with too plain or complicated textures [11]. Thus by introducing VM, the possibility of getting accurate results is maximized. The blurring map is also extracted from the VM by analyzing the local variance across horizontal and vertical blocks. First the variance is calculated across the horizontal blocks and then the vertical. The local variance at the first row of blocks B_{11} and B_{12} is given by:

$$\sigma_{11} = \frac{\sqrt{\sum_{i=1}^{n} |x_i - y_i|}}{n-1} \quad (7)$$

Where $n = 2$ in the example in Figure 3. Next we compute the average of these local variances along row j in the image, as follows:

$$\overline{\Delta\sigma_j} = mean\{\sigma_{ji} - \sigma_{j(i+1)} | i \in \{1, 2, ..., K\}\}, \quad (8)$$

Where K is the number of 8×8 blocks in the horizontal direction of the image. The sum of total blur across vertical direction is given by

$$Blr_v = \sum_{j=1}^{N} \overline{\Delta\sigma_j} \quad (9)$$

Where N is the total number of rows in the image.

The value for the horizontal blur is calculated in similar fashion.

4.4.1 *Blurring Score*

The blurring score can be estimated by summing up the vertical and horizontal blurring values.

$$B_S = Blr_v + Blr_h \quad (10)$$

4.5 Quality Prediction Score

The last step is the QPM which combines the blocking and blurring features extracted from the dataset. First the average blocking and blurring values are obtained by combining the vertical and horizontal artifacts.

$$Blc = \left(\frac{Blc_v + Blc_h}{2}\right), Blr = \left(\frac{Blr_v + Blr_h}{2}\right) \qquad (11)$$

Then the following prediction model is used to combine the artifacts

$$QPM = 10 \times \left(\alpha + \delta \times Blc^a \times Blr^b\right) \times T^c \qquad (12)$$

Where $\alpha = 0.5$ and $\delta = 2.356$ are adjusted based on the opinion score on training dataset. a, b and c are correlation parameters. A perceptual threshold value $T = 2$ is also obtained for acceptable amount of blocking and blurring artifacts in an image via subjective test questionnaire.

5. EXPERIMENTAL SETUP

We conducted the psychophysical experiment where 17 non-experts viewers participated. Most of the viewers were students at NISlab in Høgskolen i Gjøvik. The monitor white point, light intensity and color quality was adjusted as per standard ITU-R BT.500-11 recommendations. Viewers were asked to rate 130 different images on a scale of 1 to 5. Where 1 corresponds to very annoying image quality and 5 corresponds to imperceptible image quality.

The image dataset for the experiment was created from the blackboard lecture videos. In Figure 4, left column, some sample images are shown. The acquired images were organized into three different categories as shown in Table 1. Category 1 and category 2 dataset consist of 5 sets of different images. Each set contains further 10 degraded variations of those images. Category 1 images were introduced with blocking artifact using imwrite routine in Matlab. In category 2 images, blur was introduced using gaussian filter. The third category consist of 3 sets of different images with each set having 10 degraded variations. These images were subjected to both blocking and blurring artifacts.

Table 1: Classification of artifacts into different categories.

	Artifacts Type	Sets	Images	Total Image
Cat. 1	blocking	5	10	50
Cat. 2	blurring	5	10	50
Cat. 3	block/blur	3	10	30

6. EXPERIMENTAL RESULTS

From the psychophysical experiment we obtained the MOS. The MOS of 45 images from the dataset were used to obtain the correlation parameters a, b and c during the training of QPM. Once the QPM is trained, we obtained the quality score from our metrics for all 130 images. The quality scores from peak signal to noise ratio (PSNR), structural similarity index metrics (SSIM), universal image quality index (IQI), Delta Eab with LAB, and non-reference perceptual quality analysis metric (NR-PQA) were also obtained. The scores were then correlated with MOS to obtain the correlation results for comparison.

Table 2: Comparison of correlations results.

	Pearson correlation	Kendall correlation	Spearman correlation
PSNR	0.791	0.714	0.752
RF-PQM	0.841	0.779	0.842
NR-PQA	0.679	0.532	0.710
SSIM	0.579	0.563	0.585
IQI	0.521	0.533	0.510
Delta Eab	0.433	0.408	0.413

Table 3: Individual pearson correlation per dataset.

	Blocking				
Image Dataset No.	*1*	*2*	*3*	*4*	*5*
PSNR	0.77	0.63	0.89	0.54	0.76
RF-PQM	0.87	0.85	0.93	0.63	0.80
NR-PQA	0.70	0.67	0.80	0.65	0.83
SSIM	0.56	0.60	0.53	0.49	0.61
IQI	0.48	0.59	0.55	0.33	0.65
Delta Eab	0.55	0.52	0.48	0.51	0.68
	Blurring				
PSNR	0.67	0.61	0.78	0.59	0.71
RF-PQM	0.65	0.59	0.80	0.87	0.77
NR-PQA	0.34	0.25	0.48	0.31	0.38
SSIM	0.54	0.72	0.58	0.61	0.69
IQI	0.59	0.65	0.84	0.51	0.48
Delta Eab	0.64	0.55	0.38	0.51	0.29
	Blocking/Blurring				
PSNR	0.81	0.62	0.85	-	-
RF-PQM	0.75	0.79	0.89	-	-
NR-PQA	0.64	0.65	0.58	-	-
SSIM	0.44	0.70	0.68	-	-
IQI	0.54	0.45	0.78	-	-
Delta Eab	0.34	0.69	0.58	-	-

A very good correlation result of 0.841 is obtained from our proposed metric RF-PQM for all 130 images as can be seen in Table 2. From the results obtained it is evident that the proposed metric seems to correlate well with HVS unlike others. There are some cases where other metrics show improved results depicted in Table 3. For example, NR-PQA in case of blocking artifacts show good results on two occasions. It is actually hard to accurately estimate the amount of blurring without the presence of original signal. Which is evident from PSNR, SSIM and IQI that show high correlation in some cases because of significant color and structural degradation in particular image dataset. PSNR on the other hand seems to perform better compare to the others for image datasets with combined artifacts. Nevertheless the RF-PQM without the presence of reference image still shows improvement in images having both blurring and blocking artifacts compared to other metrics. The scatter plots shown in Figure 4 also verify the overall correlation results. Where RF-PQM show very high correlation results compared to the others.

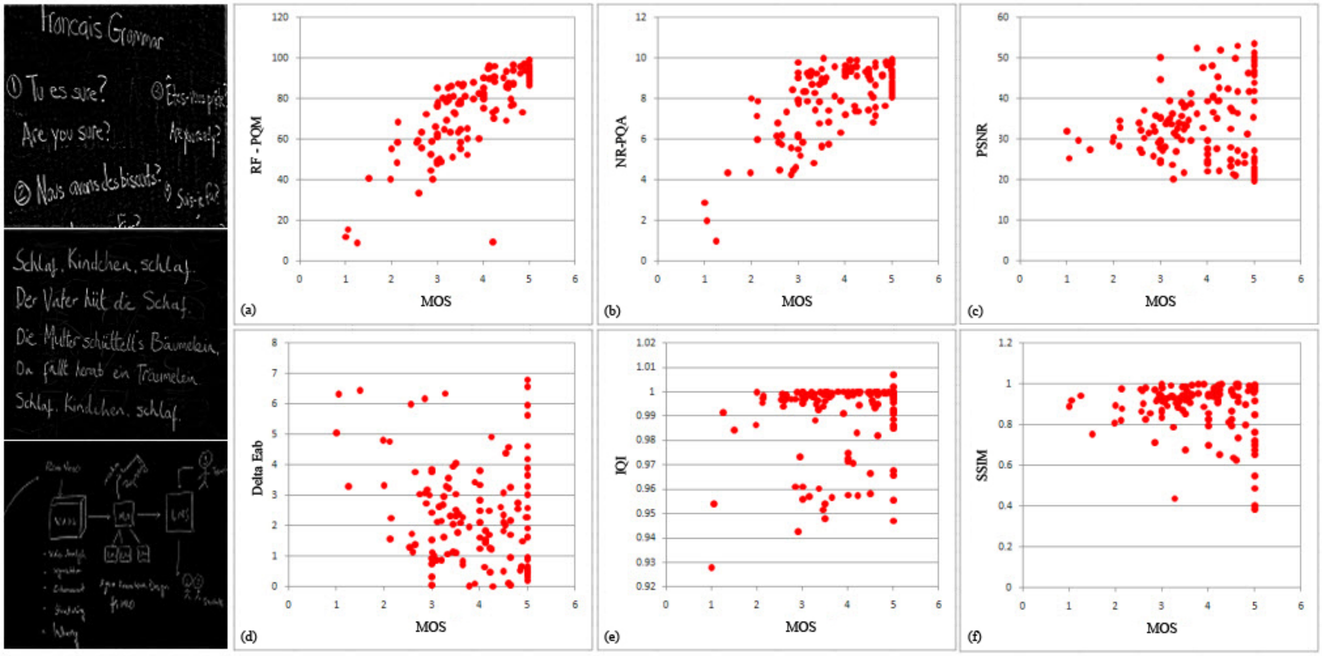

Figure 4: Sample of images from dataset *(left column)*. Scatter plots between: (a) RF-PQM vs MOS (b) NR-PQA vs MOS (c) PSNR vs MOS (d) Delta Eab vs MOS (e) IQI vs MOS (f) SSIM vs MOS

7. CONCLUSION

A reference free objective quality prediction metric is proposed that has been designed for text based images. In text based images people mostly focus on the text rather than the uniform background. Thus RF-PQM relies on saliency detection for text which makes it correlate better with the HVS. Subjective tests are also conducted where users are asked to rate the quality of image. The obtained results were then compared with standard reference and non reference metrics. However the overall image quality for lecture image is highly depended upon the text size, orientation and handwriting style which makes it difficult to adopt a metric for generalized lecture images.

REFERENCES

[1] Zhou Wang, H.R. Sheikh, and A.C. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, 2002, vol. 1, pp. I–477 – I–480 vol.1.

[2] Xin Wang, Baofeng Tian, Chao Liang, and Dongcheng Shi, "Blind image quality assessment for measuring image blur," in *Image and Signal Processing, 2008. CISP '08. Congress on*, 27-30 2008, vol. 1, pp. 467 –470.

[3] Wei Fu, Xiaodong Gu, and Yuanyuan Wang, "Image quality assessment using edge and contrast similarity," in *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on*, 1-8 2008, pp. 852 –855.

[4] Chong-Yaw Wee, R. Paramesran, and R. Mukundan, "Quality assessment of gaussian blurred images using symmetric geometric moments," in *Image Analysis and Processing, 2007. ICIAP 2007. 14th International Conference on*, 10-14 2007, pp. 807 –812.

[5] Z.M.P. Sazzad, Y. Kawayoke, and Y. Horita, "Spatial features based no reference image quality assessment for jpeg2000," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, sept. 2007, vol. 3, pp. III –517 –III –520.

[6] Timothée Jost, Nabil Ouerhani, Roman von Wartburg, René Müri, and Heinz Hügli, "Assessing the contribution of color in visual attention," *Comput. Vis. Image Underst.*, vol. 100, no. 1-2, pp. 107–123, 2005.

[7] Junyong You, Andrew Perkis, Miska M. Hannuksela, and Moncef Gabbouj, "Perceptual quality assessment based on visual attention analysis," in *MM '09: Proceedings of the seventeen ACM international conference on Multimedia*, New York, NY, USA, 2009, pp. 561–564, ACM.

[8] L. Itti and C. Koch, "Computational modelling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 3, pp. 194–203, Mar 2001.

[9] Wen Lu, Kai Zeng, Dacheng Tao, Yuan Yuan, and Xinbo Gao, "No-reference image quality assessment in contourlet domain," *Neurocomput.*, vol. 73, no. 4-6, pp. 784–794, 2010.

[10] K. Belkacem-Boussaid and A. Beghdadi, "A new image smoothing method based on a simple model of spatial processing in the early stages of human vision," *Image Processing, IEEE Transactions on*, vol. 9, no. 2, pp. 220 –226, feb 2000.

[11] Liu Debing, Chen Zhibo, Ma Huadong, Xu Feng, and Gu Xiaodong, "No reference block based blur detection," in *Quality of Multimedia Experience, 2009. QoMEx 2009. International Workshop on*, 29-31 2009.