Human Visual Perception-based Image Quality Analyzer for Assessment of Contrast Enhancement Methods

Soong-Der Chen
College of Information Technology, University Tenaga Nasional, Malaysia

Abstract: *Absolute Mean Brightness Error (AMBE) and entropy are two popular Image Quality Analyzer (IQA) metrics used for assessment of Histogram Equalization (HE)-based contrast enhancement methods. However, recent study shows that they have poor correlation with Human Visual Perception (HVP); Pearson Correlation Coefficient (PCC)<0.4. This paper, proposed a new IQA which takes into account important properties of HVP with respect to luminance, texture and scale. evaluation results show that the proposed IQA has significantly improved performance (PCC>0.9). It outperforms all IQAs in study, including two prominent IQAs designed for assessment of image fidelity in image coding-Multi-Scale Structural Similarity (MSSIM) and information fidelity criterion.*

Keywords: *IQA, visual perception, distortions, contrast enhancement.*

Received May 20, 2013; accepted November 10, 2013; published online April 1, 2015

1. Introduction

Histogram Equalization (HE) is one of the popular methods used to enhance the contrast of image. It has been widely used in many areas such as medical and radar imaging. However, HE could cause undesirable distortions such as:

1. Excessive Brightness Change.
2. Noise.
3. Saturation.

Many variants of HE have been proposed to overcome the above mentioned problem. They can be broadly classified into two categories:

1. *Automatic-Human*: Intervention not required in the process of enhancement ([4, 7, 9, 10, 11, 12, 14, 16, 21, 23, 25, 26]).
2. *Adjustable-User*: Can interactively regulate the degree of enhancement by altering parameter's value ([1, 2, 3, 6, 8, 13, 15, 17, 22]).

An ideal contrast enhancement method must be able to automatically enhance image's contrast without any perceptually annoying distortion. Towards this direction, adjustable methods are left beyond the scope of this paper as they are not automatic.

Although, all automatic methods are designed to overcome the problem of distortion, the extent to which they are resilient to distortions remains questionable. In fact, [5] has reported that the automatic methods [4, 10, 23, 25] are not resilient to noise. This paper, aims to review the existing Image Quality Analyzer (IQA) used to assess HE-based methods. Section 2 reviews and discuss the shortcomings of existing IQAs. Section 3 proposes a new IQA which take into account important properties of Human Visual Perception (HVP). Section 4 describes an experiment setup to evaluate the IQA's correlation with regards to HVP of distortions. The experiment was designed based on recommendations from Video Quality Expert Group (VQEG). Section 5 discusses the results of the experiment and section 6 provides the Conclusions and recommendations of future work.

2. Review of Existing IQAs

Table 1 lists the available automatic HE-based methods together with the IQAs that have been used to evaluate them. Absolute Mean Brightness Error (AMBE) and entropy appear to be the two most frequent used analyzers.

Table 1. List of automatic HE-based methods and their IQA(s).

Methods	IQA
Brightness preserving Bi-HE (BBHE) [10].	•AMBE
Multi-peak HE (Multi-peak) [25].	•AMBE
Equal area Dualistic Sub-Image HE (DSIHE) [23].	•AMBE •Entropy •Background Brightness
Minimum Mean Brightness Error Bi-HE (MMBEBHE) [4].	•AMBE
Brightness Preserving Histogram Equalization with Maximum Entropy (BPHEME) [21].	•AMBE •Entropy
Brightness Preserving Dynamic Histogram Equalization (BPDHE) [7].	•AMBE •Entropy
Multi-Histogram Equalization Methods for Contrast Enhancement and Brightness Preserving (Multi-HE) [11].	•AMBE
[9] Recursively Separated and Weighted Histogram Equalization for Brightness Preservation and Contrast Enhancement (RSWHE)	•AMBE •Entropy •PSNR
Bi-Histogram Equalization with a Plateau Limit for Digital Image Enhancement (BHEPL) [14].	•Average AMBE
Adaptive Contrast Enhancement Methods with Brightness Preserving (DQHEPL and BHEPL-D) [12].	•Average AMBE •Average Entropy •Average PSNR
Fusion Framework of Histogram Equalization and Laplacian Pyramid (FFHELP) [26].	•Standard Deviation of AMBE •Enhancement by Entropy
Image Contrast Enhancement using Bi-Histogram Equalization with Neighborhood Metrics (BHENM) [16].	•Average AMBE

2.1. Absolute Mean Brightness Error

AMBE It is the absolute difference between the mean of input and output image. It is formally defined by Equation 1:

$$AMBE=|E(X)-E(Y)| \quad (1)$$

Where X and Y denotes the input and output image respectively, and $E(.)$ denotes the expected value, i.e., the statistical mean. Equation 1 clearly shows that AMBE is designed to detect one of the distortions-excessive brightness changes. In fact, all the automatic methods so far have been designed to preserve brightness. The idea of preserving brightness was originated by the author of BBHE, who suggested that the fundamental reason HE could produce undesirable distortions is because, it does not take the mean brightness of an image into account. In current practice, lower AMBE implies that the original brightness is better preserved and hence, should yield a better quality output.

Figure 1. Modified image of plane with AMBE=21.81.

Since, AMBE is designed to detect changes in overall brightness, using it to indicate the presence of other distortions such as noise could be misleading. Figures 1 and 2, show the modified images of plane with different AMBE. Despite having much smaller AMBE (1.53), Figure 2 clearly shows the presence of noise (false contour in the background) which is not seen in Figure 1 with a much higher AMBE (21.81). This may explain why the automatic methods evaluated by [5] were found not resilient to noise despite that they were brightness-preserving.

Figure 2. Modified image of Plane with AMBE=1.53.

2.2. Entropy

The entropy here refers to the Shannon entropy. It is a measure of the uncertainty associated with a random variable. It quantifies, in the sense of an expected value, the information contained in an information source (in this case, the image), usually in units such as bits. It is formally defined by Equation 2:

$$H(X)=\sum_{i=1}^{n} p(x_i)I(x_i)=-\sum_{i=1}^{n} p(x_i)\log_b I(x_i) \quad (2)$$

Where X: Image, x_i : Level I, $p(x_i)$: Probability of level I, b: Units, (image pixel is coded in bit, so b=2), And n: Number of levels.

Theoretically, the higher the entropy, the more information is available from the information source. HE is designed to maximize the entropy of an image by remapping the gray levels using the gray levels' probability density function such that they are distributed uniformly. It is assumed that by increasing the entropy, the image could reveal more information. Consequently, an image with higher entropy is regarded to have better quality.

For global gray-level transformation, remapping gray levels using their probability density to obtain uniform distribution can only be achieved if the data is in continuous (non-discrete) form. In discrete form, the mapping using probability density function which is always monotonic can never increase the entropy. Furthermore, HE tends to combine gray levels of relatively low probability density and results in decrease of entropy despite the fact that such action tends to increase the contrast of an image. Figures 3 and 4, show two modified images of monarch but with slightly different entropy. Despite having lower entropy (6.20), Figure 4 clearly shows better contrast than Figure 3 with higher entropy (6.23).

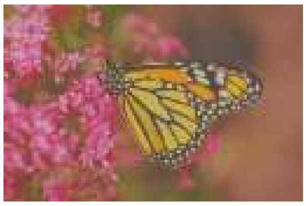

Figure 3. Modified image of Monarch with entropy=6.23.

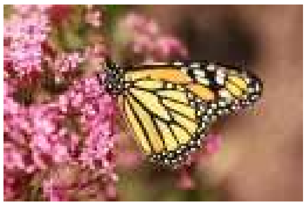

Figure 4. Modified image of Monarch with entropy=6.20.

Based on the above discussion, there is a need to evaluate the correlation of AMBE and entropy with HVP of distortions. The coming section presents the details of the evaluation.

3. The Proposed HVP-based IQA

3.1. Framework

This paper, proposes an IQA which is based on a framework as shown in Figure 5:

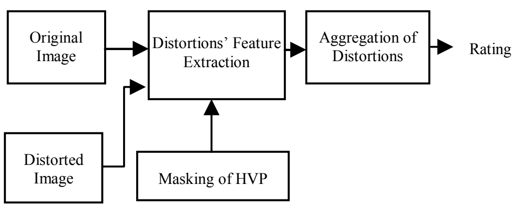

Figure 5. Framework of IQA.

- *Distortion's Feature Extraction*: Although, there are three types of distortions (excessive brightness change, noise artifacts and brightness saturation) which are commonly found in the output images of HE-based techniques, this paper proposes to focus only on the noise artifacts for the following reasons:
  1. It is observed that excessive brightness change may not always cause visual annoyance. This may be due to the fact that excessive brightness change only happens to image of with very low or very high average brightness at which HVP is likely to experience saturation and become less sensitive to brightness change. When the image's average brightness is pushed by HE towards the middle range [1] where HVP has the highest sensitivity, the visibility of the image should likely be enhanced than degraded, though there is significant change in the average brightness.
  2. It is also observed that brightness saturation tends to happen when there are very glaring and annoying noise artifacts. Such phenomena could be explained by the contrast stretching mechanism of HE-based techniques that assigns more contrast-stretching to dominant (frequently occur) gray levels which may amplify noise artifacts, while contrast shrink those of minority which may cause brightness saturation. Since, noise artifacts are made up of dominant gray levels covering much greater portion of the image, they are likely to be much more visible than brightness saturation that are made up of the minority.

  It is observed that the most common and annoying noise artifacts appear in form of edges which are not found in original image. The paper proposes to use edge detection as one of basic steps in extracting feature of annoying distortion. Basically, edge points which are detected in distorted image and not detected in the original image could be classified as noise artifacts. The detection of noise artifacts also incorporates the human visual masking as presented in the following section.

- *Human Visual Masking*: It is essential for an IQA to take into account the important properties of HVP such that it could have good correlation to human perceived annoyance of distortions. Masking designates the reduction in the visibility of one stimulus due to the presence of another. Such masking effect is mainly due to the limitations in the sensitivity of sensor cell in retina in relation to the activity of its surrounding sensor cell. There are three fundamental visual masking effects which are highly related to the visibility of distortions as below:
  1. *Luminance Masking*: It is well known that HVP experiences saturation near the lower and the upper end of the luminance intensity range. Hence, distortions are less visible in area with very low or very high intensity than in area with medium intensity. This paper, proposes to incorporate luminance masking by using different threshold for edge detection according to the average intensity of an area; higher threshold for area with very low or very high average intensity.
  2. *Texture Masking*: It is also know that HVP is more sensitive to changes in area with smooth texture/low activity. Therefore, distortions appear to be less visible in areas with many details/textures compare to those in smooth areas. This paper, proposes to incorporate texture masking by varying weight of each detected edge points according to its surrounding texture; higher weight for edge point in area with low activity.
  3. *Scale Masking*: It is known that HVP is more sensitive to changes of larger scale (size). Consequently, distortions appear to be less visible in small scale. This paper proposes to incorporate scale masking by using multi-scale analysis.

- *Aggregation of Distortions*: In practice, one usually needs a single overall quality measure of the entire image. It can be derived from the aggregation of all the detected noise artefacts in distorted image according to their respective weight.

3.2. Algorithm

Before edge detection, the color image is converted from Red-Green-Blue (RGB) color space to gray scale image using the conversion as shown in Equation 3:

$$I_g(r, c)=0.2989I_r(r, c)+ 0.5870I_g(r, c)+0.1140\, I_b(r, c) \tag{3}$$

For edge detection, this paper proposes to use the well-known Sobel edge operator. The edge magnitude of a point at row r, and column c is as defined in Equation 4:

$$EM(r,c) = \sqrt{S_r^2 + S_c^2} \tag{4}$$

$$S_r=\sum_{i=0}^{n}\sum_{j=0}^{n} I(r-1+i, c-1+j)M_r(i,j) \tag{5}$$

$$S_c=\sum_{i=0}^{n}\sum_{j=0}^{n} I(r-1+i, c-1+j)M_c(i,j) \tag{6}$$

It is basically Euclidean norm of the correlation between sub-image of size $n\times n$ (n=3) centered at the point (r, c) with Sobel row mask, M_r and column mask, M_c as defined below:

$$M_r=\frac{1}{9}\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix} \quad M_c=\frac{1}{9}\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix} \tag{7}$$

A point is classified as edge if its edge magnitude is above a predetermined threshold value, $T(r, c)$. Since, there tends to be contrast difference between original and distorted image, threshold value used for original image, T_o should be lower than the threshold value used for distorted image, T_d. The values have been determined empirically as T_o=0.0001 and T_d=0.0002. In order to account for the effect of luminance masking, $T(r, c)$ is set according to sub-image's average luminance level, $m_l(r, c)$. For 8 bits/pixel resolution, the values have been determined empirically as L_{low}=40 and L_{high}=245.

$$T(r,c)=\begin{cases} T', \text{ for } m_l(r,c) \geq L_{low} \text{ and } m_l(r,c) \leq L_{high} \\ 2T', \text{for } m_l(r,c) < L_{low} \text{ or } m_l(r,c) > L_{high} \end{cases} \tag{8}$$

$$T'=\begin{cases} T_o, \text{for } I(r,c)= \text{original image} \\ T_d, \text{for } I(r,c)= \text{distorted image} \end{cases} \tag{9}$$

$$m_l(r,c)=\frac{1}{n^2}\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} I(r-1+i,\ c-1+j) \tag{10}$$

A point is classified as noise if there is edge detected in distorted image ($EM_d(r,c) \geq T(r,c)$) but not detected in original image ($EM_o(r,c) < T(r,c)$), where $EM_d(r, c)$ and $EM_o(r, c)$ denotes the edge magnitudes of distorted and original image respectively. In order to account for the effect of texture masking, local entropy $H_l(r, c)$ is computed to measure the activity of a sub-image. Low entropy indicates low level of activity. Only noise in sub-image with activity level below a predetermined threshold, H_T is considered visible. Empirically, H_T=2.5.

$$I_{noise}(r,c)=\begin{cases} 1 \quad \text{for} \quad EM_o(r,c) < T(r,c) \text{ and } \\ \qquad EM_d(r,c) \geq T(r,c) \text{ and } \\ \qquad H_l(r,c) < H_T \\ 0 \quad \text{otherwise} \end{cases} \tag{11}$$

$$H_l(r,c)= -\sum_{g=0}^{L-1} p_l(g)\log_2 p_l(g) \tag{12}$$

$p_l(g)$ denotes the probability of gray level g which is the total number of pixels with gray level g, $N_l(g)$ divided by the total number of pixels within the sub-image of size $n \times n$ (n=9).

$$p_l(g)= \frac{N_l(g)}{n^2} \tag{13}$$

$$N_l(g)=\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} I_g(r-1+i,\ c-1+j) \tag{14}$$

$$I_g(r,c)=\begin{cases} 1, \text{for } I(r,c)=g \\ 0, \text{for } I(r,c)\neq g \end{cases} \tag{15}$$

The rating, R: Is the ratio of total number of pixels with noise artifacts detected to the image's total number of pixels.

$$R=\frac{\sum_{i=11}^{height-1}\sum_{j=0}^{width-1} I_{noise}(i,j)}{height \times width} \tag{16}$$

The above rating is repeated using images of smaller scale. This paper proposes to repeat for 3 different scales, each is half of the previous scale as shown in Figure 6. The final rating is maximum rating of the rating from the 3 scales as in Equation 17.

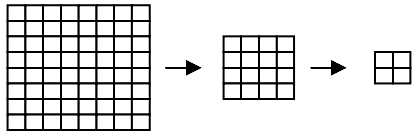

Figure 6. Reduce image size to half for each scale.

$$R_{final}=\max\left\{R_{S1}, R_{S2}, R_{S3}\right\} \tag{17}$$

4. Performance Evaluation

4.1. Image Database

A set of 9 source images of diverse image content was selected from "Lossless True Color Image Suite" provided by Rich Franzen and "LIVE Image Quality Assessment Database" provided by Laboratory of Image and Video Engineering at University of Texas, Austin. Figure 7 shows these source images. These images were preprocessed to simulate poor contrast image which shows distortions after HE as follows:

- Original images were JPEG compressed at quality Q=50.
- The JPEG compressed images were contrast-reduced such that the new dynamic range is between from 0.2-0.8 of the old dynamic range.

Each contrast-reduced image was then processed using SGHESE [6] with different parameter's values to generate one stimulus for each of the distortion levels below [20]:

1. Very Annoying.
2. Annoying.
3. Slightly Annoying.
4. Perceptible but not Annoying.
5. Imperceptible.

a) Caps (512×768). b) Flower (512×768). c) Sailboats (720×480).

d) Plane (512×768). e) Parrots (512×768). f) Buildings (505×634).

g) Lighthouse (512×768). h) Monarch (512×768). i) Carnival (488×610).

Figure 7. The 9 source images and their spatial resolution (height×width).

A total of 43 stimuli were generated (stimulus with distortion level "imperceptible" is not available for caps and lighthouse). The parameters' values were chosen by human expert by interacting with graphical user interface as shown in Figure 8. Table 2 lists the chosen parameter's values.

Figure 8. Graphical user interface of SGHESE.

Table 2. List of parameters used to generate the 43 stimuli.

	Imperceptible		Perceptible but not Annoying		Slightly Annoying		Annoying		Very Annoying	
	Min	Max	Min	Max	Min	Max	Min	Max	Min	Max
Caps	-	-	0.01	0.2	0.01	0.7764	0.01	0.7764	0.01	1
Carnival	0.01	0.1	0.01	0.17	0.01	0.3	0.01	0.5	0.05	1
Plane	0.01	0.03	0.01	0.05	0.01	0.16	0.01	0.3098	0.01	1
Lighthouse	-	-	0.01	0.12	0.01	0.23	0.01	0.6	0.2	1
Flower	0.01	0.1	0.01	0.2025	0.01	0.3544	0.01	0.6667	0.6	1
Parrots	0.01	0.1	0.01	0.32	0.01	0.5	0.01	1	0.88	1
Sailboats	0.01	0.12	0.01	0.18	0.01	0.28	0.01	0.5	0.01	1
Monarch	0.01	0.17	0.01	0.22	0.01	0.32	0.01	0.6	0.01	1
Buildings	0.01	0.1	0.01	0.3	0.01	0.55	0.01	0.83	0.57	1

4.2. Test Methodology

The experiment was designed based on double-stimulus method. The original and distorted images were displayed on two separate screens placed next to each other, both connected to a PC that supports extended display. The screens used were both 17" liquid-crystal display screens, with a native resolution of 1920×1080 pixels and a screen refresh rate of 60 Hz. The experiment was conducted in a standard office environment and the viewing distance was approximately 60 cm. Subjects rated the level of distortion for each stimulus by comparing it with the corresponding contrast-reduced image. The scoring scale ranged from 0 to 100, where "0-20" means "Very annoying", "21-40" means "Annoying", "81-100" means "Imperceptible" and so on.

The subjects participated in the experiment were bachelor of IT students from University Tenaga National. The 45 students (35 male and 10 female) were inexperienced with image quality assessment and distortions. They were briefed about the objective and procedures of the experiment. No training session were given as that could influence subject's opinion. After briefing, the stimuli were presented to subject in a random order and the ratings were recorded in a score sheet.

4.3. Raw Data Processing

1. Outlier Detection and Subject Rejection: Simple outlier detection and subject rejection procedures were carried out on the raw scores before the actual data analysis. Raw score for an image was considered to be an outlier if it was outside an interval of 2.33 standard deviations about the mean score for that image [19]. Also, all scores of a subject were rejected if more than 6 of his scores were outliers. Overall, only 1 subject was rejected and only less than 4% of the total scores were rejected.
2. MOS Scores: In order to, compute the Mean Opinion Score (MOS), the raw scores were first converted into Z scores after outlier removal and subject rejection. The Z scores for i^{th} subject and j^{th} image is as defined in Equation 18:

$$Z_{ij} = \frac{(S_{ij} - \overline{S_i})}{\sigma_i} \tag{18}$$

Where S_{ij}: The raw score for i^{th} subject and j^{th} image, $\overline{S_i}$: The average of all the scores rated by subject i, σ_i: The standard deviation all the scores rated by subject i, and Z: scores were then averaged across subjects to yield the MOS for the j^{th} image as defined in Equation 19:

$$MOS_j = \frac{1}{S} \sum_{i=1}^{S} Z_{ij} \tag{19}$$

Where S: The total number of subjects after subject's rejection.

5. Results and Discussions

5.1. Performance Metrics

According to the recommendations from VQEG [20] the performance of an IQA can be quantitatively evaluated with respect to its ability to predict subjective quality rating in the following three aspects:

1. Prediction Accuracy: The ability to predict the subjective quality score with low error. The metrics used were: Pearson Correlation Coefficient (PCC) and Root Mean Squared Error (RMSE).
2. Prediction Môn tonicity: The degree to which the IQA's prediction agrees with the relative magnitudes of the subjective quality rating. The metric used was Spearman Rank Order Correlation Coefficient (SROCC).
3. Predication Consistency: The degree to which the IQA maintains prediction accuracy over different types of images and not to fail excessively for a subset of images. The metric used were Outlier Ratio (OR) (ratio of outlier to total scores). Outlier score is a score outside an interval of two times the standard deviation about the MOS.

The evaluation was done using MOS after non linear regression using a five-parameter logistic function (a logistic function with an added linear term, constrained to be monotonic) [19] as defined by (18):

$$R(x) = b_1\left(\frac{1}{2} - \frac{1}{1+e^{[b_2(x-b_3)]}}\right) + b_4 x + b_5 \tag{20}$$

This nonlinearity was applied to the MOS or its logarithm, which ever gave a better fit for all data.

In this experiment, Peak Signal to Noise Ratio (PSNR), Multi-Scale Structural Similarity (MSSIM) [24] and Information Fidelity Criterion (IFC) [18] metrics were also evaluated besides AMBE and Entropy. MSSIM and IFC are two prominent IQAs designed to measure image fidelity in image compression. Table 3 shows the results obtained.

Table 3. The results of Pearson CC, RMSE, SROCC and OR for AMBE, Entropy, MSSIM and the proposed IQA.

	Pearson CC	RMSE	SROCC	OR
AMBE $\|E(X) - H(Y)\|$	0.1346	0.7809	0.0802	0.2791
Entropy $\|H(X) - H(Y)\| \div H(X)$	0.3291	0.7510	0.7682	0.2093
IFC	0.4025	0.7705	0.3961	0.1860
PSNR	0.7009	0.6074	0.7274	0.0930
MSSIM	0.7174	0.5840	0.7628	0.1395
Proposed IQA (without scale masking)	0.8687	0.3815	0.8990	0
Proposed IQA (with scale masking)	0.9036	0.358	0.9088	0

5.2. Discussions

Table 4 shows the interpretations of the value of Pearson CC and SROCC which are widely accepted and used by many scientific journals.

Table 4. Interpretation of correlation values.

Value of Correlation	Interpretation
0.00-0.40	Poor
0.41-0.75	Fair
0.76-0.85	Good
0.86-1.00	Excellent

Based on the interpretation in Table 4, the results in Table 3 show that:

- All the four metrics consistently indicate that AMBE has poor correlation with MOS. This is consistent with our discussion in section 2A that brightness change does not always cause annoying effect.
- All metrics except SROCC indicate that Entropy has poor correlation with MOS. This indicates that Entropy agrees well with relative magnitude of MOS but fails to accurately predict the MOS. Further study and modification on entropy is required to improve its predication accuracy.
- The performance of IFC-based measure is unexpectedly poor given that it is one of the best performing measures reported by [18]. Such result hints that the performance of IQA could be application-dependant; IQA designed for image coding may not necessary work well for image contrast enhancement.
- PSNR: All the four metrics consistently show that the conventional PSNR and is fairly correlated to MOS and its performance is comparable to those of MSSIM.
- MSSIM only outperforms PSNR marginally in most metrics except OR.
- All the four metrics consistently show that the proposed IQA clearly outperforms all other IQAs in study. All the four metrics consistently show that the proposed IQA correlates excellently with MOS with zero outlier. It is worth highlighting that incorporating scale-masking clearly improves the prediction accuracy of the proposed IQA.

6. Conclusions and Recommendations

In this paper, the existing IQAs (AMBE and entropy) used to evaluate the output of HE-based contrast enhancement methods have been reviewed and their shortcomings have been highlighted. A subjective quality assessment was conducted in which image quality data obtained from 1935 human observer opinion scores were used to evaluate the IQAs. The experiment results showed that AMBE and Entropy have poor correlation with HVP. This paper has also proposed a new IQA based on a generic framework which takes into account of HVP such as luminance, texture and scale. Experiment results show that the proposed IQA clearly outperform the rest of the IQAs in study, including the two prominent IQAs-MSSIM and IFC. Future work recommended is reduce the complexity of the proposed IQA (while maintaining the predication accuracy) to achieve near real-time processing.

The data used in this experiment is made publicly available to the research community for further scientific study in the field of image quality assessment.

Acknowledgements

I would like to thank Nur Hanani Azami and all the UNITEN B.I.T second and final year students in year 2011 who have contributed to the compilation of opinion score.

References

[1] Abdullah-Al-Wadud M., Kabir M., Dewan M., and Chae O., "A Dynamic Histogram Equalization for Image Contrast Enhancement," *IEEE Transaction Consumer Electron*, vol. 53, no. 2, pp. 593-600, 2007.

[2] Arici T., Dikbas S., and Altunbasak Y., "A Histogram Modification Framework and Its Application for Image Contrast Enhancement," *IEEE Transaction Image Processing*, vol. 18, no. 9, pp. 1921-1935, 2009.

[3] Chen S. and Ramli A., "Contrast Enhancement using Recursive Mean-Separate Histogram Equalization for Scalable Brightness Preservation," *IEEE Transaction Consumer Electron*, vol. 49, no. 4, pp. 1301-1309, 2003.

[4] Chen S. and Ramli A., "Minimum Mean Brightness Error Bi-Histogram Equalization in Contrast Enhancement," *IEEE Transaction Consumer Electron*, vol. 49, no. 4, pp. 1310-1319, 2003.

[5] Chen S. and Sidhu M., "Re-evaluation of Automatic Global Histogram Equalization-based Contrast Enhancement Methods," *SOURCE Electronic Journal of Computer Science and Information Technology*, vol. 1 no. 1, pp. 13-17, 2009.

[6] Chen S. and Suleiman A., "Scalalable Global Histogram Equalization with Selective Enhancement for Photo Processing," *in Proceedings of the 4th International Conference on Information Technology and Multimedia*, Malaysia, pp. 744-752, 2008.

[7] Ibrahim H. and Kong N., "Brightness Preserving Dynamic Histogram Equalization for Image Contrast Enhancement," *IEEE Transaction Consumer Electron*, vol. 53, no. 4, pp. 1752-1758, 2007.

[8] Ibrahim H. and Kong N., "Image Sharpening using Sub-Regions Histogram Equalization," *IEEE Transaction Consumer Electron*, vol. 55, no. 2, pp. 891-895, 2009.

[9] Kim M. and Chung M., "Recursively Separated and Weighted Histogram Equalization for Brightness Preservation and Contrast Enhancement," *IEEE Transaction Consumer Electron*, vol. 54, no. 3, pp. 1389-1397, 2008.

[10] Kim Y., "Contrast Enhancement using Brightness Preserving Bi-Histogram Equalization," *IEEE Transaction Consumer Electron*, vol. 43, no. 1, pp. 1-8, 1997.

[11] Menotti D., Najman L., Facon J., and Araujo A., "Multi-Histogram Equalization Methods for Contrast Enhancement and Brightness Preserving," *IEEE Transaction Consumer Electron*, vol. 53, no. 3, pp. 1186-1194, 2007.

[12] Ooi C. and Isa N., "Adaptive Contrast Enhancement Methods with Brightness Preserving," *IEEE Transaction Consumer Electron*, vol. 56, no. 4, pp. 2543-2551, 2010.

[13] Ooi C. and Isa N., "Quadrants Dynamic Histogram Equalization for Contrast Enhancement," *IEEE Transaction Consumer Electron*, vol. 56, no. 4, pp. 2552-2559, 2010.

[14] Ooi C., Kong N., and Ibrahim H., "Bi-histogram with a Plateau limit for Digital Image Enhancement," *IEEE Transaction Consumer Electron*, vol. 55, no. 4, pp. 2072-2080, 2009.

[15] Pichai S., Krishnasamy B., and Karuppanagounder S., "Bi-Level Weighted Histogram Equalization for Scalable Brightness Preservation and Contrast Enhancement for Images," *the International Arab Journal of Information Technology*, vol. 10, no. 6, pp. 603-609, 2013.

[16] Sengee N., Sengee A., and Choi H., "Image Contrast Enhancement using Bi-Histogram Equalization with Neighbourhood Metrics," *IEEE Transaction Consumer Electron*, vol. 56, no. 4, pp. 2727-2734, 2010.

[17] Sheet D., Garud H., Suveer A., Mahadevappa M., and Chatterjee J., "Brightness Preserving Dynamic Fuzzy Histogram Equalization," *IEEE Transaction Consumer Electron*, vol. 56, no. 4, pp. 2475-2480, 2010.

[18] Sheikh H., Bovik A., and Veciana G., "An Information Fidelity Criterion for Image Quality Assessment using Natural Scene Statistics," *IEEE Transaction Image Processing*, vol. 14, no. 12, pp. 2117-2128, 2005.

[19] Sheikh H., Sabir M., and Bovik A., "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," *IEEE Transaction Image Processing*, vol. 15, no. 11, pp. 3440-3451, 2006.

[20] VQEG., "VQEG Final Report of FR-TV Phase II Validation Test," available at: http://www.its.bldrdoc.gov/vqeg/reports.aspx, last visited 2003.

[21] Wang C. and Ye Z., "Brightness Preserving Histogram Equalization with Maximum Entropy: A Variational Perspective," *IEEE Transaction Consumer Electron*, vol. 51, no. 4, pp. 1326-1334, 2005.

[22] Wang Q. and Ward K., "Fast Image/Video Contrast Enhancement Based on Weighted Thresholded Histogram Equalization," *IEEE Transaction Consumer Electron*, vol. 53, no. 2, pp. 757-764, 2007.

[23] Wang Y., Chen Q., and Zhang B., "Image Enhancement based on Equal Area Dualistic Sub-image Histogram Equalization method," *IEEE Transaction Consumer Electron*, vol. 45, no. 1, pp. 68-75, 1999.

[24] Wang Z., Bovik A., Sheikh H., and Simoncelli E., "Image Quality Assessment: From Error Visibility to Structural Similariy," *IEEE Transaction Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.

[25] Wongsritong K., Kittayaruasiriwat K., Cheevasuvit F., Dejhan K., and Somboonkaew A., "Contrast Enhancement using Multi-peak Histogram Equalization with Brightness Preserving," *in Proceedings of IEEE Asia-Pacific Conference on Circuits and Systems*, Chiangmai, pp. 455-458, 1998.

[26] Yun S., Kim J., and Kim S., "Image Enhancement using a Fusion Framework of Histogram Equalization and Laplacian Pyramid," *IEEE Transaction Consumer Electron*, vol. 56, no. 4, pp. 2763-2771, 2010.

Soong-Der Chen is a senior lecturer in the Department of Graphics and Multimedia, College of Information Technology, Universiti Tenaga Nasional, Malaysia. He received his BE in Electronics/Computer in 1997, MSc in Electronic Imaging in 2000 and PhD in 2008 from the Department of Computer and Communication Engineering in University Putra Malaysia, Serdang, Malaysia. Chen has more than 15 years of research experience in the area of image processing. He has secured funding for three research projects related to image quality from Malaysia Ministry of Higher Education since 2010. Chen supervises three PhD and in image processing, artificial intelligence and green IT. He has graduated four master students since 2008 His research interest includes image quality assessment, image enhancement, computer vision and image compression. He also, leads and teaches modules at both undergraduate and postgraduate levels in IT.