Full-Reference Contrast-Distortion Image Quality Assessment Algorithm

Hasan Thabit Rashid Kurmasha
Computer Department
College of Education for Women, Kufa University
Najaf, Iraq
hasant.kurmasha@uokufa.edu.iq

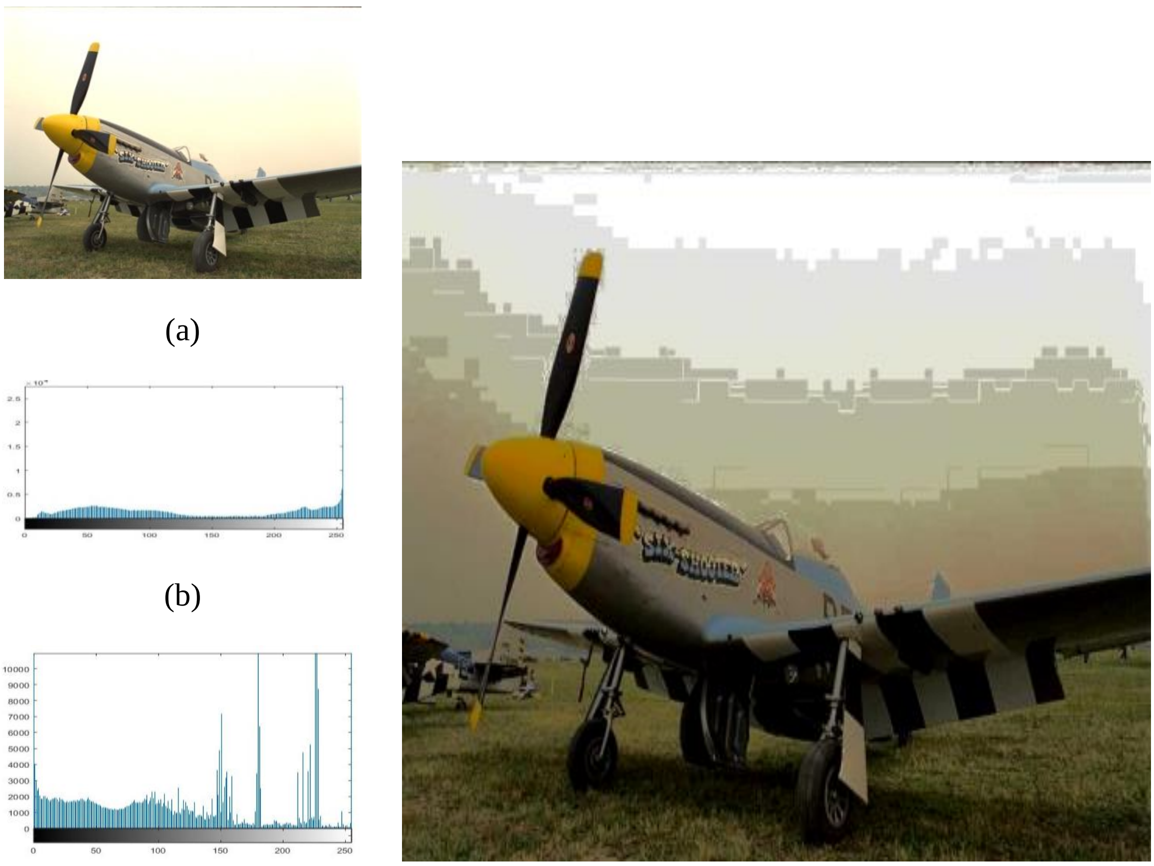
Abstract. Image enhancement is a popular technique used for process transferring distortions; contrast-distorted images CDI are treated using different images quality assessment IQA algorithms as full-reference FR-IQA. In this paper, a new FR-IQA is proposed to determine the noise artifacts caused by compression, blurring and other types of distortion as well as the brightness saturation (i.e. no details in the region) which happen for the same reasons above. The proposed FR-IQA is designated based on the variance moment which has the feature of fast computation, very sensitive to represent the dispersion of image pixel intensity (change in details), and very sensitive to the overall brightness of images in correlation to the human visual sensitive HVS. The statistical evaluation outcomes expression that most of the FR-IQAs have humble correlation along with human mean opinion scores MOS although they are using the original image in comparing with the distortion one. The proposed algorithm takes into account the significant properties of human visual perception HVP. The algorithm objective and subjective results have very well scores. It has significantly well correlation for person correlation coefficient PCC > 0.87, spearman rank order SROCC > 0.91 and outlier ratio (OR) equal to 0%.

Keywords: Histogram Equalization HE, Distortions, Contrast Enhancement, Image Quality Measures, Human Visual Perception

Introduction

The quick manufacturing rollout of internationally omnipresent community media platforms and compressed picture broadcast technologies in recent years has to deal with a slew of IQAs degradations originating during the content acquisition, transport, and storage processes (McCarthy, 2012; Kim et al., 2019). The creation of models that can predict perceived image quality as accurately as feasible would tremendously aid efforts to improve these systems. As a result, many FR-IQAs techniques have been developed during the last few decades based on perceptual models that attempt to emulate the HVS reactions to distortion. Though, creating a sufficiently complete and integrative HVS model is a tough task that is still unsolved (Kim et al., 2020). The degree of visual degradations could be happen to an image is referred to as image quality as in Figure 1.

Noise, blocking artifacts, blurring, fading, and other factors can cause degradation. These flaws appear during image processing operations. Detecting image degradation during acquisition might let you take countermeasures to diminish poverty during saving the image as a file. As a result, a quantitative image evaluation system that is automated is needed (Kamble & Bhurchandi, 2015). FR-IQA, which is best suited for comparing multimedia service quality performance, can perform objective picture quality assessment (Yang, Sun, & Wang, 2018), see full-reference FR-IQAs in Sheikh and Bovik (2006), Wang and Bovik (2009), Zhang et al. (2011), Liu, Lin, and Narwaria (2011). Image grayscale fluctuations are typically employed to describe spatial structure information in FR-IQA approaches. Furthermore, image distortions may create changes of structural information in the same areas among the reference image and the distorted image, for example, blur alterations may alter an area with a very strong grayscale fluctuation into a flat zone. The proposed FR-IQA approach measures image structure information changes induced by distortion in both structural intensity and structural distributions. Absolute mean brightness error AMBE and Entropy are most FR metrics used in assessing histogram equalization-based approaches, however they have long been challenged for their weak association with human mean opinion scores MOS. Mean square error MSE and peak signal-to-noise ratio PSNR reveal the similar deficit in perceived image quality. This paper are organized as follow, the FR-IQM algorithm details in section III, section IV setting the subjective experiment, analysing the outcomes, and comparing them to the existing matrices, section V has the final conclusion and suggested future work, and the history of IQAs-based approaches, evaluations of AMBE, entropy, HPV overviews, and the distortions are in section II.



(a) (b) (c) (d)

Figure 1. Plane original image (512×768) in (a) (Sheikh, Sabir, & Bovik, 2006) and its histogram in (b) while the very annoying version in (d) and its histogram in (c).

Image Distortion Review

The automatically assessment of the quality of images is the objective of an IQA, consequently a countless effort has been made to improve IQA algorithms that associate fine with HVP for distortions. HE technique is a very popular for CDI and it is commonly used in many fields as radar and medical image handling and therefore, a lot of adjustments of HE-based techniques over the years have been proposed. Normally, they are automatic (no need user interference into process) and adjustable (user needs to adjust the parameter) to control enhancement' amount. However, investigate the problem of distortions in such techniques based the correlation between the IQA's and the MOS for predict the changes encourage this work to focus on the problem of noisy and saturation distortions in such automatic techniques. Figure 2 shows the influence of distortion in the histogram of Caps image.

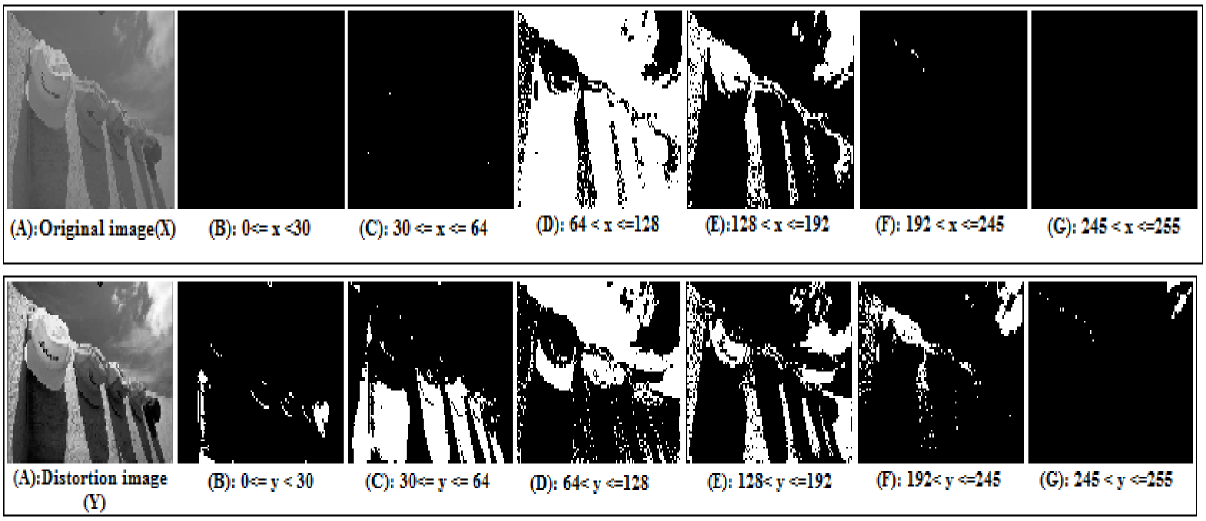


Figure 2. Histogram distribution of Caps image (X) and its distorted version (Y).
Note: The detected pixels (white) in the range; while black is for no-pixel, caps image (X) caps with less AMBE (14.02), Caps image (Y) has high AMBE (37.31). Notice that most of image pixels concerning in the middle range in X while the pixels are distributed along the grayscale range (0…255) in Y.

Luminance HVP masking states the influence of distortion' visibility is advanced for middle background intensity while it is reduced in little or in height intensity. Texture HVP masking states that distortion' visibility is advanced in homogenous regions than in coarse or details regions (texture areas). In Figure 3 (a), the noise visibility is high in the right side which has medium intensity while it is very low in the left side which is dark. In Figure 3 (b), the noise visibility is high in the left side which is smoothly than in the right side which is coarse region. However, IQAs is stronger well when it models the non-linear responses of the visual system extremely to detect distortions and spatial appearances.

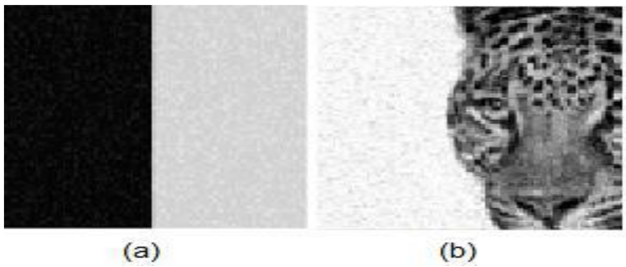


Figure 3. The effect of luminance masking HVP in (a), while the effect of texture masking HVP in (b). (Chen, 2012)

The Proposed FR IQM

This paper proposes to analysis the distortion image contents at pixel level as for distortions' features extraction based on HVS. It focuses on the noisy and saturation pixels which set up in the production images of CDI techniques such as HE-based techniques. In such techniques, the distortion phenomena tend to distribute the brightness average along the range of the grayscale levels from 0 to 255, and the degree of the distribution depend on the degree of the distortion for that image and however, if there is highly distortion image, then the distribution of the average brightness pixels will be increased while it will be more normally in original image excepted the image that has more brightness scene which tend to move on to the higher score of the range. HVP highest sensitivity is where the image's brightness average is constrained into the middle range and then, the image' visibility should be improved, and yet there is influenced alteration in the illumination's average. However, when the appearance of the noise are becoming very obvious and clearing and the dominant grayscales levels are contrast stretching and leading to extension of the noise pixels while the minority of noise pixels are shrinks by contrast and may cases the brightness saturation. The developed generic framework of the new FR-IQM in Figure 1 and algorithm details are as following.

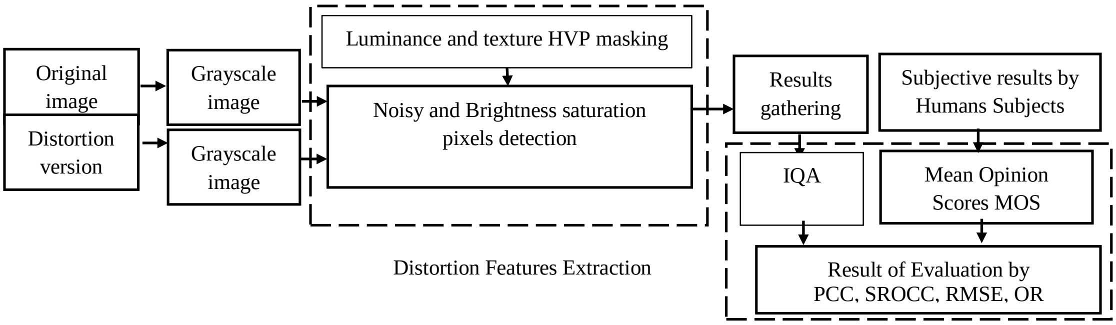


Figure 4. The generic framework of the proposed FR-IQA

It is basically, if the variance moment is higher in image; that's mean it is more enhanced and has more visible based on luminance HVS masking as well as the difference or any change in variance between both original and distorted images will lead to make an influence in the texture of image' details and result in low visible based on texture HVS masking. However, for noisy pixels detection the pixels that certify the two compound conditions in Equation 1 will consider as noisy-pixels and otherwise will not be.

$$I_{noise}(r,c) = \begin{cases} 1, If\ \left(V_O(r,c) > V_D(r,c)\right) \\ and\ (T1 < Difference_V < T2) \\ 0, Else \end{cases} \quad (1)$$

The method proposes to measure the variance moment at each pixel (sub-image 3×3) and compare between the original pixel $V_o(r,c)$ and the corresponding distorted one $V_D(r,c)$ as the first compound condition. The second condition is the difference $_V$ among the variance of both original and distorted pixels should be higher than *T1* and lower than *T2* to be considered as noisy pixel since the variance based on HVS is related and measured by these two thresholds. The two thresholds are chosen empirically to maximize the IQA's correlations with MOS, to normalize distorted image, and to use it as standard for all dataset of tested images.

$$V(r,c) = \sqrt[2]{\sum_{g=0}^{l-1} (g-\bar{g})^2 P_l(g)} \quad (2)$$

$$\bar{g} = Mean = \sum_{g=0}^{l-1} g\, P_l(g) \quad (3)$$

$P_l(g)$ is the probability of grayscale level g which is the whole amount of pixels with grayscale level $N_l(g)$ divided by the whole amount of pixels in the of size $n \times n$ region.

$$P_l(g) = \frac{N_l(g)}{n^2} \quad (4)$$

$$N_l(g) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} I_g(r-1+i, c-1+j) \quad (5)$$

$$I_g\,(r,c) \;=\; \begin{cases} 1, for\ I(r,c) = \ g \\ 0, for\ I(r,c) \neq \ g \end{cases} \tag{6}$$

High variance means that the image has a good contrast. Therefore, brightness saturation change detection for each image pixel with less contrast in distorted image as compared to contrast in original image then, this pixel will be considered as saturation pixel as Equation 7 below.

$$I_s(r,c) = \begin{cases} 1, \left(\left(V_O(r,c)\ ./\ V_D(r,c)\right) > T\_lum\right) \\ \quad And\ (V_O(r,c) > T\_textu) \\ \quad 0, Else \end{cases} \tag{7}$$

The first condition is in relating to the effect of the luminance-HVP masking, the distortion pixel contrast is lower the original pixel contrast under the luminance saturation threshold *T_Lum* which is chosen empirically to ensure to ensure that the saturation is observable, while the brightness saturation account in relating for the effect of the texture-HVP masking is considered to the original image under the texture saturation threshold *T_textu*, as a result that there is adequate activity or information to be perceived in the original image. The grouping among $I_n(r,c)$ and $I_s(r,c)$ is done by using Equation 8 below.

$$I_{ns}(r,c) = \sum_{r=0}^{n}\sum_{c=0}^{m} I_n(r,c) \mid I_s(r,c) \tag{8}$$

As a result, the rating *R* is the ratio of the aggregate sum of noisy and saturation pixels noticed to the image' of pixels ($h\times w$).

$$R \;=\; \frac{\sum_{i=0}^{h-1}\sum_{j=0}^{w-1} I_{ns}(i,j)}{h\times w} \tag{9}$$

Evaluation Results

A nine source images dataset showing adequate various in image content was chosen (Franzen, 1999; Sheikh et al., 2005). 1032 human viewer MOS (24 persons) are collected from a subjective quality assessment (Kurmasha & Alharan, 2017). Figure 8 below shown scaling rang of five levels of distortions which applied to the each source image into the experiment.

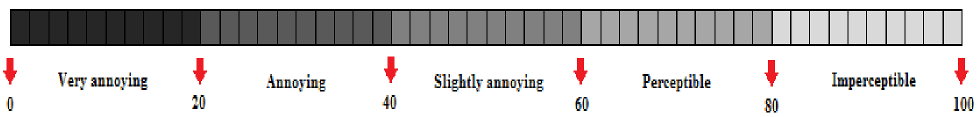


Figure 8. Scaling rang of five levels of distortions.

The evaluations of performance metrics are used as recommended in Recommendation of VQEG (2003); the "PCC" and "root mean squared error RMSE" are applied to expect the low errors of the subjective quality scores. "SROCC" metric used to quantity the grade to which the model' expectation agrees with the relative scales of the individual quality assessment. "Outlier ratio OR" metric is the ratio of outlier to entire marks which measure the grade to which the model preserves expectation precision over diverse forms of images and not to flop extremely for a subgroup of images. Also, to compare the performance with the proposed IQM; current full references metrics are used: AMBE, Entropy, "PSNR", "multiscale structural similarity MSSIM" (Wang et al., 2004) and FR-edge IQA (Chen, 2012). All of these metrics are designed to measure image quality. Table 1 shows the outcomes achieved by the objective measures and Table 2 presents the proposed IQM rating and the MOS of subjects for the applied image dataset after normalizing the scores between 0...1 while Figure 9 shows the linear correlation between it. AMBE has poor correlation with MOS as the fact that the intensity change does not continuously basis annoying result. Entropy also indicates poor correlation with MOS. For "PSNR and MSSIM", the performance metrics demonstration a properly correlations to MOS. FR edge-IQM shows an very good correlation to MOS for detecting the noisy pixels based on edge detection and by taking into account the important properties of HVP. The proposed FR-IQM shows that very good correlation to MOS comparing to all others in detecting the noisy and brightness saturation pixels.

Table 1. Objective metrics used to predict the correlation and accuracy.

IQMs	PCC	RMES	SROCC	OR	GRAPH REPRESENTATION OF THE RESULTS
AMBA	0.1349	0.0408	0.9316	0.201	
Entropy	0.2968	0.3424	0.2057	0.213	
PSNR	0.6701	0.713	0.5529	0.042	
MSSIM	0.7738	0.7403	0.1424	0.1951	
FR Edge IQM	0.8021	0.839	0.4964	0.023	
Proposed FR IQM	0.8754	0.8951	0.2191	0	

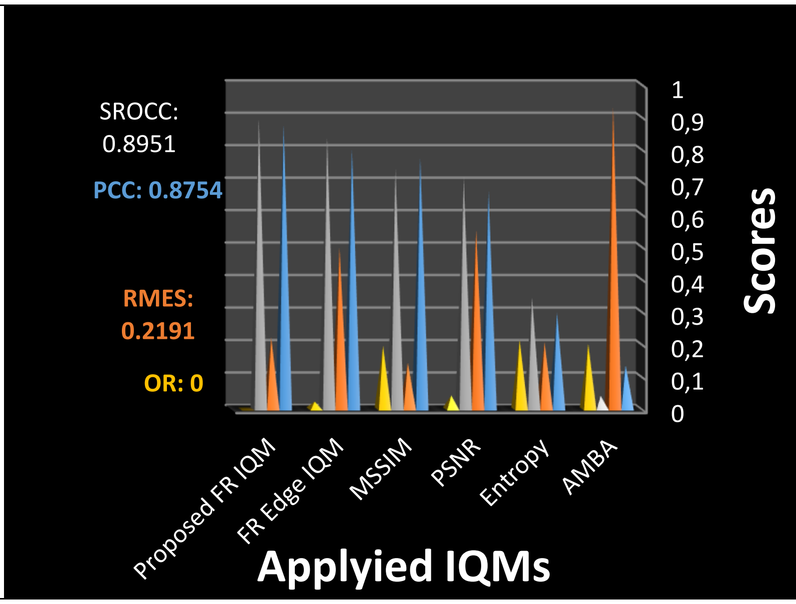


Table 2. The proposed IQM rating and the subjects' MOS after normalization for the image dataset.

	IQM rating	MOS		IQM rating	MOS		IQM rating	MOS
1	0.000323	0.645976	**16**	0.004168	0.729814	**31**	0.00297	0.231554
2	0.036924	-0.55915	**17**	0.033325	-1.69051	**32**	0.003082	0.457774
3	0.016571	0.404486	**18**	0.001016	0.58115	**33**	0.002815	0.971765
4	0.001885	0.827896	**19**	0.019964	-0.25042	**34**	0.02136	-0.60624
5	0.011241	0.05724	**20**	0.023974	-1.08139	**35**	0.037793	-1.63486
6	0.015493	0.29554	**21**	0.001972	0.669857	**36**	0.026878	-1.41897
7	0.000295	0.864254	**22**	0.000881	1.160653	**37**	0.01946	-0.78353
8	0.012227	-0.23787	**23**	0.020368	0.030567	**38**	0.003443	0.098345
9	0.004486	0.080548	**24**	0.001417	1.009186	**39**	0.010879	-0.24554
10	0.000239	0.929048	**25**	0.002045	1.292883	**40**	0.001455	0.933573
11	0.020681	-1.18261	**26**	0.041677	-1.75727	**41**	0.003215	0.733185
12	0.018113	-0.75938	**27**	0.003981	0.080777	**42**	0.003263	-0.04165
13	0.000596	0.596587	**28**	0.00369	0.476495	**43**	0.010729	-0.53926
14	0.011543	-0.47005	**29**	0.001534	0.251798			
15	0.014414	-0.79047	**30**	0.023492	-0.32089			

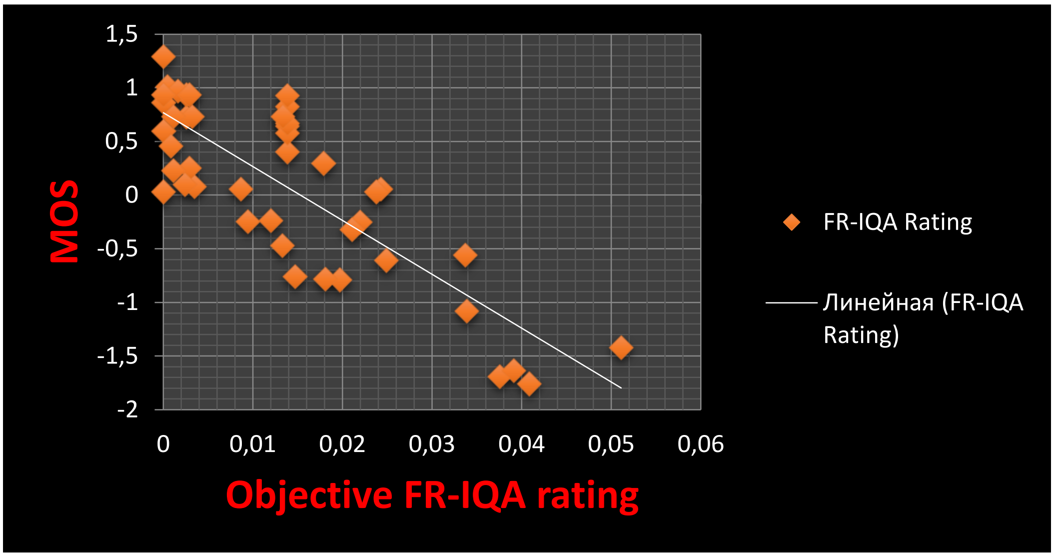


Figure 9. The linear correlation' scatter-plot chart for the proposed IQM rating and the MOS as of Table 2.

Conclusion

To develop a perfect FR-IQM in fully automated and fast computation for in-time applications, this paper propose a new IQM based on a basic structure which takes into account the significant properties of luminance and texture HVP masks and the variance moment measure which is more fast in computation than edge-magnitude and entropy. Assessment outcomes show that the proposed FR-IQM significantly outdoes the rest of the IQMs and metrics, including the prominent MSSIM and PSNR and the two popular AMBE and Entropy. Future work recommendation is to apply the deep learning technique for the proposed FR-IQM.

References

Chen, S. D. (2012). A new image quality measure for assessment of histogram equalization-based contrast enhancement techniques. *Digital Signal Processing*, *22*(4), 640-647. http://dx.doi.org/10.1016/j.dsp.2012.04.002

Franzen, R. (1999). *Kodak lossless true color image suite*. http://r0k.us/graphics/kodak/.

Kamble, V., & Bhurchandi, K. M. (2015). No-reference image quality assessment algorithms: A survey. *Optik*, *126*(11-12), 1090-1097.

Kim, W. et al. (2019). Modern trends on quality of experience assessment and future work. *APSIPA Trans. Signal Inf. Process.*, *8*, 1–19.

Kim, W., Nguyen, A. D., Lee, S., & Bovik, A. C. (2020). Dynamic receptive field generation for full-reference image quality assessment. *IEEE Transactions on Image Processing*, *29*, 4219-4231.

Kurmasha, H. T. R., & Alharan, A. F. (2017). A Review and New Subjective Evaluation Experiment of Objective Metrics used to Evaluate Histogram Equalization-based Contrast Enhancement Techniques. *International Journal of Computer Techniques, 4*(4), 131-136.

Liu, A., Lin, W., & Narwaria, M. (2011). Image quality assessment based on gradient similarity. *IEEE Transactions on Image Processing*, *21*(4), 1500-1512.

McCarthy, S. (2012). Quantitative evaluation of human visual perception for multiple screens and multiple CODECs. In *Proc. SMPTE Annu. Tech.Conf. Exhib.*, pp. 1–8.

Recommendation of VQEG (2003). VQEG final report of FR-TV phase II validation test.

Sheikh, H. R., & Bovik, A. C. (2006). Image information and visual quality. *IEEE Transactions on Image Processing*, *15*(2), 430-444.

Sheikh, H. R., Sabir, M. F., & Bovik, A. C. (2006). A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing*, *15*(11), 3440-3451.

Sheikh, H.R., Wang, Z., Cormack, L., & Bovik, A.C. (2005). *LIVE image quality assessment database release 2.* http://live.ece.utexas.edu/research/quality.

Wang, Z., & Bovik, A. C. (2009). Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, *26*(1), 98-117.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600-612.

Yang, X., Sun, Q., & Wang, T. (2018). Image quality assessment via spatial structural analysis. *Computers & Electrical Engineering*, *70*, 349-365.

Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE transactions on Image Processing*, *20*(8), 2378-2386.

Hasan Thabit Rashid is currently a lecturer with University of Kufa, Iraq. He received his Bachelor degree in Computer Science from Babylon University, Iraq in 2005, and his Master degree in Information Technology from UNITEN University, Malaysia in 2013, and his Ph.D from Babylon University in computer science at 2021. His research interests include Intelligent Video Surveillance Systems, Image and Video Analysing, Computer Vision.