No-Reference Perceived Image Quality Algorithm for Demosaiced Images

Lamb Anupama Balbhimrao

Electronics &Telecommunication Dept.
College of Engineering Pune
Pune, Maharashtra, India

Madhuri Khambete

Electronics &Telecommunication Dept.
Cummins College of Engineering
Pune, Maharashtra, India

Abstract—**Visual image quality assessment (IQA) plays a key role in every multimedia application, as end user to it is a human-being. Real time applications demand no reference (NR) IQA, due to unavailability of the reference image. Today, most of the perceived/visual NR-IQA algorithms developed are for distortions like blur, ringing, and blocking artifacts. Very few are available for color distortions. Visible color distortions, such as false color, and zipper are produced in the demosaiced image due to incorrect interpolation of missing color values. In this paper, state of the art zipper and false color artifact quantification algorithms, general purpose NR-IQA algorithms are evaluated for visual quality assessment of demosaiced images. Separate NR- IQA algorithms are proposed for zipper and false color artifact quantification scores are then combined to obtain final quality score for demosaiced image. Zipper algorithm quantifies zipper artifact by searching for zipper pixels in an image, while false color algorithm finds correlation between local high frequency region's color planes to quantify false color.**

Keywords—Demosaicing; Correlation; False color; Image quality; Regression; Zipper

I.   INTRODUCTION

Image quality assessment plays key role in image processing pipeline for testing and validation. End user to every multimedia application is human being, so visual quality assessment is of great importance. Visual quality scores can be obtained either by conducting subjective experiment or by developing objective IQA algorithms. Conducting subjective experiment is very difficult and is not the real time solution. So many objective IQA algorithms, correlating well with subjective ratings, are developed. These objective IQA algorithms are categorized into full Reference (FR), reduced reference (RR), and no reference (NR) algorithms.

FR and RR algorithms require original image or partial information of it respectively for IQA. On the other hand NR algorithms assess quality without any reference image and only with distorted image in hand. For real time applications NR algorithms are needed as making reference image available is not always possible. Most of the NR-IQA algorithms developed are for distortions like, blur, and distortions due to compression algorithms. Few general purpose algorithms are also proposed but they are evaluated only for distortions like blur, white noise, and distortions due to JPEG, JPEG-2000.

Digital cameras capture real world scenes using a single sensor in order to reduce the size and cost of the camera. This single sensor is overlaid with a color filter array (CFA) that restricts the sensor elements to storing only red (R), green (G), or blue (B) color information. A mosaic/bayer image is subsequently generated at the output of this sensor. Although various types of CFA patterns are available, the Bayer CFA pattern [1] is the one most commonly used. Figure 1a shows the Bayer CFA pattern, where green color information is sampled more because the human visual system is more sensitive to green color information.

Demosaicing process reconstructs full color image from this bayer image by interpolating missing color values. If interpolation for missing color values is incorrect color distortions like false color, zipper, water color, or grid pattern can appear in the demosaiced image. Of these artifacts, zipper and false color are the most common and appear in high frequency areas of the image. Misinterpolation in high frequency areas, owing to limited correlation between neighboring pixels, causes zipper and false color artifacts to appear in those areas. Zipper artifacts appear as on-off-on or off-on-off patterns, whereas false color artifacts appear as aberrant colors. Figures 1b and 1c show examples of zipper and false color artifacts respectively [2].
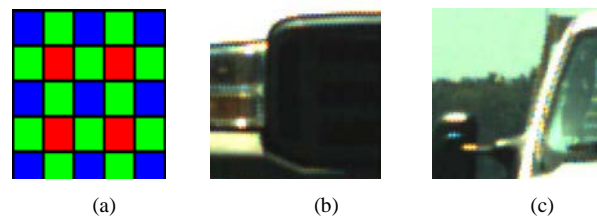

Fig. 1.   (a) Bayer CFA pattern.   (b) Zipper artifact.   (c) False color artifact

Today, quality assessment of demosaiced image and testing of demosaicing algorithm is performed in the RGB color space or International Commission on Illumination's L* a* b* (CIELAB) color space with FR objective quality metrics such as color peak signal to noise ratio (CPSNR), color mean square error (CMSE), CIELAB color difference deltaE (CIELAB ΔE), or spatial extension of the CIELAB color difference deltaE (SCIELAB ΔE) [3].

As these metrics are FR, so they have limitation for real time application. Also, it is stated in the literature that FR color difference metrics in CIELAB space indicate perceived quality, but we obtained very less correlation (around 37%) between FR color difference metrics in CIELAB and subjective opinions.

Considering all the limitations of FR-IQA algorithms for demosaiced image quality assessment, NR-IQA algorithms are proposed in the literature but they are very few. Furthermore, no correlation is exhibited between their objective scores and subjective opinions.

In this paper, state of the art NR demosaicing IQA algorithms and general purpose NR-IQA algorithms are evaluated for visual quality assessment of demosaiced images. Separate novel visual NR-IQA algorithms are proposed for zipper and false color artifacts. These scores are then combined to obtain final visual quality score for demosaiced image. The remainder of this paper is organized as follows: Section 2 offers an overview of previous work. Section 3 presents proposed NR-IQA algorithm. Section 4 outlines the performance evaluation. Section 5 concludes the paper.

II. PREVIOUS WORK

NR-IQA algorithms are categorized in to distortion specific and general purpose ones. Distortion specific NR-IQA algorithms predict perceived quality of an image distorted with specific distortion [4-7], while general purpose algorithms predict perceived quality of an image for any type of distortion. General purpose algorithms like [8-9] are trained on human rated database, while general purpose algorithm like [10] do not require training also for visual quality assessment. Very few algorithms are reported for NR quality assessment of demosaiced image. Also these algorithms are not evaluated for their correlation with subjective ratings.

A. State of the Art NR-General Purpose Algorithms

General purpose algorithms quantify visual quality for any type of distortion. In this section very popular general purpose algorithms [8-10] are presented. General purpose algorithms [8-9] are trained on distortions from LIVE database [11] and general purpose algorithm [10] is not trained on any human rated database. Blind image quality index (BIQI) [8] general purpose algorithm is a two step framework for quality assessment. The two steps are image distortion classification based on a measure of how the wavelet statistics is modified, followed by quality assessment, using an algorithm specific to the decided distortion. Blind referenceless image spatial quality evaluator (BRISQUE) [9] general purpose algorithm uses scene statistics of locally normalized luminance coefficients to quantify possible losses of naturalness in the image due to the presence of distortions thereby leading to a holistic measure of quality. Natural Image Quality Evaluator (NIQE) general purpose algorithm do not require training on human ratings and is based on collection of quality aware statistical features from a corpus of natural undistorted images. The quality of the test image is expressed as the distance between multivariate Gaussian fits of the Natural Scene Statistic (NSS) features extracted from test image and the multivariate Gaussian model of the quality aware features extracted from the corpus of natural images.

B. State of the Art Algorithms for Zipper Artifact Quantification

Very popular state of the art zipper artifact quantification algorithms [12-13] are overviewed in this section.

Lu and Tan [12] proposed a FR algorithm for quantifying zipper artifacts that is currently used by most of the demosaicing algorithms for performance evaluation. Their algorithm is outlined below:

1) Find a pixel with minimum CIELAB color difference deltaE (ΔE) with respect to the center pixel in a 3 × 3 neighborhood in the original image (image having no demosaicing artifacts is referred to as original image here). Equation (1) gives the CIELAB color difference formula, where l1, a1, and b1 are the CIELAB color values for pixel1 and l2, a2, and b2 are the CIELAB color values for pixel 2.

$$\Delta E^{*}{}_{ab} = \sqrt{(l2 - l1)^2 + (a2 - a1)^2 + (b2 - b1)^2} \qquad (1)$$

2) Compute CIELAB $\Delta E^{*}{}_{ab}$ between the same pair of pixels in the demosaiced image.

3) Find the difference between $\Delta E^{*}{}_{ab}$ of the original and the demosaiced image; if this difference is greater than 2.3, then that pixel is affected by zipper.

4) Compute percentage of such pixels which is the zipper score for that demosaiced image.

Losson, Macaire, and Yang [13] proposed a directional alternation measurement algorithm for zipper artifact measurement. Their proposed algorithm is outlined below:

5) Compute the green color variance in the original image along the vertical as well as horizontal direction at the center pixel in the 3 × 3 neighborhood.

6) Consider the lowest green color variance direction to find the alternation amplitude at the center pixel. If the lowest variance direction is horizontal, then the alternation amplitude is given by (2). However, if the lowest variance direction is vertical, then the alternation amplitude is given by (3). The alternation amplitude value is always positive if a "high-low-high" or "low-high-low" pattern exists at the considered pixel or the center pixel in the local neighborhood.

$$\alpha^{i} = |g(i-1,j) - g(i,j)| + |g(i,j) - g(i+1,j)| + |g(i-1,j) - g(i+1,j)| \qquad (2)$$

$$\alpha^{j} = |g(i,j-1) - g(i,j)| + |g(i,j) - g(i,j+1)| + |g(i,j-1) - g(i,j+1)| \qquad (3)$$

7) Compare this alternation amplitude to the alternation amplitude at the same pixel in the demosaiced image. If the value obtained is greater, then the alternation amplitude of the green levels has been amplified by demosaicing.

C. State of the Art Algorithms for False Color Artifact Quantification

State of the art false color artifact quantification algorithms [14-15] are presented in this section.

A NR algorithm is proposed in [14] for quantifying false color. The algorithm proposed is as follows:

1) Use the Prewitt operator to find edge points in the Gaussian filtered image.

2) At each edge point, compute the difference between the G and R values.

3) Compute the median of the (G-R) values in a 5 × 5 neighborhood centered on the edge point.

4) Finally, compute the MSE between the median and difference between G and R values at the edge point over all edge points.

5) Repeat the above procedure for the G and B planes.

6) Compute the average of these MSE values for the (G, R) and (G, B) planes. This average value is the false color measure.

In [15] a FR algorithm for false color quantification is proposed, where the absolute difference between the pixel values of the demosaiced and corresponding original image is computed for the R, G, and B color planes, then the maximum color difference value among the three color planes is found. Subsequently, if that difference is greater than a specified threshold value, then that pixel is considered as being affected by the false color artifact. The percentage of such pixels is quantified as the false color score. Threshold value of 2.3 is used for implementation of this algorithm. All these state of the art zipper and false color algorithms are implemented in Matlab 7.

III. PROPOSED NR-IQA ALGORITHMS FOR DEMOSAICING ARTIFACTS

Proposed perceived NR-IQA algorithm for visual quality assessment of demosaiced images is presented in this section.

A. Proposed NR Algorithm for Zipper Artifact Quantification

Zipper artifacts predominantly appear near high frequency regions such as edges, and zipper pixels appear perpendicular to edge pixels. This fact is utilized for quantification of zipper artifacts. The proposed algorithm derived is outlined below:

1) Convert demosaiced image to gray image, compute gradient magnitude and gradient direction.

2) Use the Sobel operator to find edge pixels in the horizontal as well as vertical direction.

3) Consider a pixel in 3 × 3 neighborhood around each edge pixel as a zipper pixel, if it satisfies the following conditions:

a) It is not an edge pixel.

b) Its gradient magnitude is greater than gradient magnitude of the edge pixel.

c) It is perpendicular to edge pixel.

4) Repeat this process for all edge pixels to find zipper pixels.

5) Compute the ratio of the total number of zipper pixels to the total number of edge pixels, which is considered as zipper score for demosaiced image.

B. Proposed NR Algorithm for False Color Artifact Quantification

The high frequency details in an image are more affected by false color than smooth regions. Consequently, for false color quantification, the correlation of the high frequency details of the green (G) plane to the high frequency details of the blue (B) and red (R) planes is found. G plane is used as reference because green color information is sampled more than red and blue colors in bayer image. The resulting algorithm proposed is outlined below:

1) Divide Demosaiced image in to blocks of 64×64. Block size of 64 is selected as it corrosponds to foveal region [6].

2) For each block

a) Obtain HH, HL, LH, and LL sub-bands for G, and R planes.

b) Obtain correlation between the HH, HL, and LH sub-bands of the G plane and the respective HH, HL, and LH sub-bands of the R plane.

c) Compute the average of these three correlation values, which is correlation value between block's high frequency details of G and R planes.

3) Repeat step 2 for all blocks.

4) Take average of block's high frequency details of G and R planes (obtained in step 2C), which is correlation value for high frequency regions of the G and R planes.

5) Repeat steps 2-4 for G and B planes.

6) Average out the correlation values obtained for high frequency regions of G&R (value obtained in step 4) and G&B planes to obtain proposed false color quantification score.

C. Proposed perceived NR-IQA Algorithm for Demosaiced Images

Zipper and false color scores obtained with proposed algorithms are combined, with non-linear equation as in (4), to obtain final quality score for demosaiced image. In (4) z & fc are zipper and false color scores respectively. Constant values in (4) were obtained with multiple non-linear regression.

$$\text{Quality Score} = c + c1*z + c2*fc + c3*z^2 + c4*fc^2 + c5*f*z \quad (4)$$

Table 1 gives details of the regression statistics for the multiple nonlinear regression.

TABLE I. MULTIPLE NON LINEAR REGRESSION STATISTICS

Statistics	Values	Constant Values for (4)	
Multiple R	0.955562	C	37.98871
R Square	0.9131	C1	315.2318
Adjusted R Square	0.90032	C2	-200.859
F	71.45053	C3	-1009.65
Significance F	4.74E-17	C4	300.213
		C5	-213.677

IV. PERFORMANCE EVALUATION AND COMPARISON

For experimentation, forty demosaiced images, which were utilized for subjective experiment1 in [16] are used. The subjective scores for the images are made available from Brainard, one of the authors of [16]. In [16], ten high resolution images were acquired using a Kodak DCS200 camera and 50 mm lens under various daylight conditions. From each captured image, four demosaiced images were then obtained by applying Bilinear, Freeman1 [17], Bayesian1, and Bayesian2

[18], [19] demosaicing algorithms to the bayer image. Forty demosaiced images were thus obtained from the ten original high resolution images.

For performance evaluation four (IQA) performance parameters are computed: 1) Spearman rank order correlation coefficient (SROCC); 2) Kendall's rank order correlation coefficient (KROCC); 3) Pearson linear correlation coefficient (PLCC); and 4) Root mean squared error (RMSE). Nonlinear logistic mapping function as given in [20] is used to compute PLCC and RMSE. Tables 2 and 3 give performance evaluation and comparison of proposed zipper and false color algorithms respectively. Table 4 gives performance evaluation of general purpose NR-IQA algorithms on demosaiced images. Performance evaluation of proposed NR-IQA algorithm is given in table 5. Figure 2 gives scatter plot for proposed algorithm.

TABLE II.   PERFORMANCE COMPARISON OF ZIPPER ALGORITHMS

Zipper Algorithms	Performance Parameters			
	PLCC	*SROCC*	*KROCC*	*RMSE*
Zipper algorithm in [12]	0.6439	0.3712	0.2662	23.50
Zipper algorithm in [13]	0.5443	0.3344	0.23	25.76
Proposed zipper algorithm	0.6730	0.6901	0.4874	22.72

TABLE III.   PERFORMANCE COMPARISON OF FALSE COLOR ALGORITHMS

False Color Algorithms	Performance Parameters			
	PLCC	*SROCC*	*KROCC*	*RMSE*
False color algorithm in [14]	0.4224	0.4567	0.3204	27.84
False color algorithm in [15]	0.5637	0.4422	0.3101	25.40
Proposed false color algorithm	0.9555	0.8078	0.6056	9.05

TABLE IV.   PERFORMANCE EVALUATION OF GENERAL PURPOSE ALGORITHMS FOR VISUAL QUALITY ASSESSMENT OF DEMOSAICED IMAGES

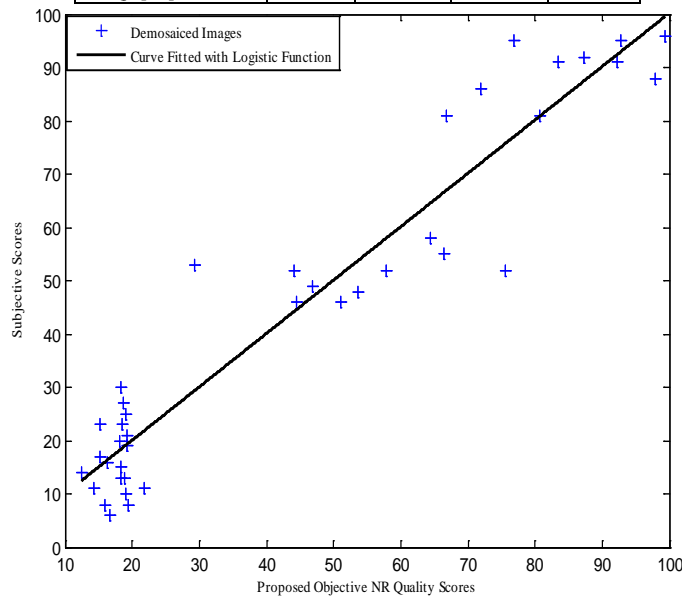| General Purpose Algorithms	Performance Parameters			
	PLCC	*SROCC*	*KROCC*	*RMSE*
BIQI [8]	0.5055	0.4261	0.2905	26.50
BRISQE [9]	0.3956	0.4548	0.3163	28.19
NIQE[10]	0.4838	0.5355	0.3783	26.86

Fig. 2.   Scatter plot of proposed NR-IQA algorithm for demosaiced images

TABLE V.   PERFORMANCE EVALUATION OF PROPOSED NR-IQA ALGORITHM

Proposed NR-IQA Algorithm	Performance Parameters			
	PLCC	*SROCC*	*KROCC*	*RMSE*
	0.9552	0.8601	0.6770	9.0866

Statistical significance testing is also required in order to evaluate the performance of the IQA algorithm, if the number of tested images are less. So, we performed F-test and t-test between predicted and the actual subjective scores at 5% significance level. The F-test was used to check whether the difference between the variance of the predicted and actual subjective scores is statistically significant or not, whereas the t-test was used to check whether the difference between the means of the two is statistically significant or not. Table 6 gives details of the F-test and t-test for proposed NR-IQA algorithm for demosaiced images. Results in table 6 shows that proposed algorithm's scores are statistically significant as F & t statistic values are less than critical values.

TABLE VI.   STATISTICAL SIGNIFICANCE TESTING

Statistical Parameters	Proposed NR-IQA Algorithm for Demosaiced Images
F-statistic	1.090163
F-critical one tail	1.704465
t-statistic	0.00011
t-critical two tail	1.990847036

V.   CONCLUSION

Separate NR-IQA algorithms are proposed for quantification of zipper and false color artifacts then these scores are combined to obtain final visual quality score for demosaiced image. Experimental results points towards usability of proposed algorithm for visual quality assessment of demosaiced images. State of the art demosaicing artifact quantification algorithms as well as general purpose NR-IQA algorithms fail to quantify visual quality of demosaiced image. Statistical significance testing also confirms validity of proposed algorithm's scores. So, proposed algorithm can be utilized for testing and validation of demosaicing algorithm and for visual quality assessment of demosaiced image.

REFERENCES

[1] Bayer, B.E.: Color imaging array. U.S. Patent 3 971 065, July 20, (1976).

[2] Maschal Jr., R.A.Young, S.S. Reynolds, J. Krapels, K.Fanning J., and Corbin T., "Review of Bayer pattern color filter array (CFA) demosaicing with new quality assessment algorithms," ARL-TR-5061, January (2010).

[3] Franck JOURNES (2014), "A study of image quality assessment and color image reconstruction algorithms for mono-sensor camera," (Master's thesis, Université Jean Monnet Saint-Etienne).

[4] H.R. Sheikh, A.C. Bovik, and L.Cormack, "No Reference Quality Assessment Using Natural Scene Statistics: JPEG 2000," IEEE Trans. Image Process. , **4**(11), (2005).

[5] P.Marziliano, F. Dufaux, S.Winkler, and T.Ebrahimi, "Perceptual blur and ringing metrics: Applications to Jpeg 2000," Signal process.Image Commun,**19** (02),163-172 (2004).[10.1016/j.image.2003.08.003]

[6] N. D. Narvekar, and L. J. Karam, "A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD)," IEEE Trans. Image Process, **20** (9), (2011). [10.1109/TIP.2011.2131660]

[7] R. Hassen, Z. Wang, M. Salama, "Image Sharpness Assessment Based on Local Phase Coherence," IEEE Trans. Image Process.**22** (7), 2789-2810 (2013). [10.1109/TIP.2013.2251643]

[8] A. K. Moorthy, and A. C. Bovik, "A Two step Framework for Constructing Blind Image Quality Indices," *IEEE Signal Processing Letters,* **17** (5), (2010). [10.1109/LSP.2010.2043888]

[9] A. Mittal, A. K. Moorthy, A. C. Bovik, "No reference Image Quality Assessment in the Spatial Domain," IEEE Transactions on Image Processing, **21**(12), (2012). [10.1109/TIP.2012.2214050]

[10] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a Completely Blind Image Quality Analyzer," IEEE Signal Processing Letters, **20** (3), (2013). [10.1109/LSP.2012.2227726]

[11] H. R. Sheikh, Z. Wang, L. Cormack ,and A. C. Bovik, "LIVE Image Quality Assessment Database Release 2", http://live.ece.utexas.edu/research/quality.

[12] Lu W., and Tan Y.P, "Color filter array demosaicking: New method and performance measures," IEEE Trans Image Process, 12(10), 1194–1210, (2003).

[13] Losson O., Macaire L., and Yang Y., "Comparison of color demosaicing methods," Advances in Imaging and Electron Physics, Elsevier, 2010, 162, pp.173-265.

[14] Maschal Robert A., Susan Y., Reynolds Joseph P., Krapels K., Fanning J., and Corbin, T., "New Image Quality Assessment Algorithms for CFA Demosaicing," IEEE sensors journal, vol. 13, January 2013.

[15] Yang Y., Losson O., and Duvieubourg L., "Quality evaluation of color demosaicing according to image resolution," In : Proceedings of the 3rd International Conference on Signal-Image Technology & Internet-based Systems (SITIS'07). Shanghai Jiaotong University, China, pp. 689–695.

[16] Longere P., Xuemei Z., Delahunt P., and Brainard D., "Perceptual assessment of demosaicing algorithm performance," In: Proceedings of the IEEE, vol. 90, no. 1, pp. 123–132 (2002).

[17] Freeman W., and Polaroid C., Method and apparatus for constructing missing color samples. US Patent 4 774 565, 1988.

[18] Brainard D., and Sherman D., " Reconstructing image from trichromatic samples from basic research to practical applications," In: Proc. IS&T/SID 1995 Color Imaging Conf., 1995.

[19] Brainard D., "Bayesian method for reconstructing color images from trichromatic samples," In: Proc. IS&T 47th Annual Meeting, Rochester, NY, 1994, pp. 375–380.

[20] Sheikh H., Sabir M., and Bovik A., "A statistical evaluation of recent full reference image quality assessment algorithms," IEEE Trans. on Image Processing, 15(11), 3440-3451, (2006) .