

LETTER

No-Reference Quality Metric of Blocking Artifacts Based on Color Discontinuity Analysis

Leida LI[†], Member, Hancheng ZHU[†], Jiansheng QIAN^{†a)}, and Jeng-Shyang PAN^{††}, Nonmembers

SUMMARY This letter presents a no-reference blocking artifact measure based on analysis of color discontinuities in YUV color space. Color shift and color disappearance are first analyzed in JPEG images. For color-shifting and color-disappearing areas, the blocking artifact scores are obtained by computing the gradient differences across the block boundaries in U component and Y component, respectively. An overall quality score is then produced as the average of the local ones. Extensive simulations and comparisons demonstrate the efficiency of the proposed method.

key words: image quality assessment, blocking artifact, color discontinuity

1. Introduction

Image quality is important for many image processing applications. The key of image quality assessment (IQA) is to obtain a quality measure that is consistent with human perception. According to the availability of a reference image, the current IQA methods can be classified into full-reference (FR), reduced-reference (RR) and no-reference (NR) metrics [1]. While most of the existing methods are FR and RR ones, NR metrics are more useful because a reference image is not always available in practice.

Blocking artifacts are mainly caused by JPEG compression, where the blocks are quantized individually without considering the high correlations of adjacent blocks. Recently, several methods have been proposed to evaluate the blocking artifacts in images. Perra *et al.* [2] extracted image edges using the Sobel operator. The luminance variations of both block boundary pixels and inner block pixels were calculated to produce the blockiness score. Pan *et al.* [3] computed the edge directions from the edge image. The population of the edges with 0° and 90° orientations was calculated to produce the pixel discontinuity measure. The signal activity measure was computed as the population of the pixels with 180° orientations. The quality metric was obtained combining the discontinuity measure and the activity measure. Liu *et al.* [4] computed the blockiness metric in gradient domain with a simplified visual masking model. In [5], the difference image was processed along each row and column, producing one-dimensional signals. Then dis-

crete Fourier transform was adopted to analyze the periodic peaks, which were signs of blocking artifacts. Recently, Lee *et al.* [6] addressed a method in the spatial domain. Candidate block boundaries were determined using the pixel gradients. The boundaries with actual blocking artifacts were detected by investigating the pixel gradients on both sides of the boundary. The quality score was defined as log of average strength of blockiness over the entire image.

The current methods evaluate the blocking artifacts in the luminance domain. In JPEG, coarse quantization is applied to the color component so that color discontinuity is produced, which is also a representative feature of blocking artifacts. In this work, blocking artifacts are estimated by analyzing the color discontinuities in YUV color space. For the areas with shifted color, the blocking artifacts are evaluated in U channel. For the areas missing color, it is done in Y channel. The performance of the proposed method is verified by extensive experiments and comparisons.

2. Color Discontinuities in JPEG Images

In JPEG compression, the luminance component and the color component are quantized individually. Fine quantization is applied to the luminance component, while coarse quantization is applied to the color component. Due to the coarse quantization, JPEG produces color discontinuities, where the image takes on blocking artifacts. Typically, two kinds of areas may appear in JPEG images, namely color-shifting area (CSA) and color-disappearing area (CDA).

Figure 1 shows an example of color discontinuity in a

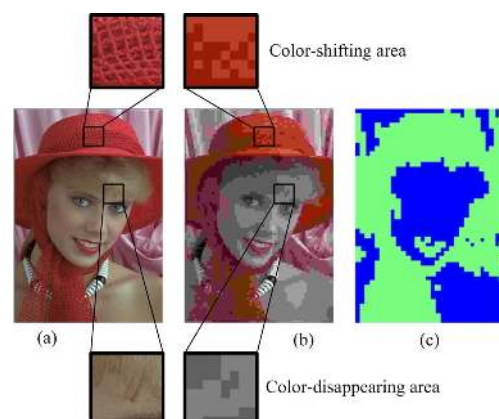


Fig. 1 An example of color discontinuity. (a) Original image, (b) Low bit rate JPEG image, (c) Color distortion map.

Manuscript received August 12, 2013.

Manuscript revised October 30, 2013.

[†]The authors are with the School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou 221116, China.

^{††}The author is with the Department of Computer Science and Technology, Harbin Institute of Technology Shenzhen Graduate School, Shenzhen 518055, China.

a) E-mail: iqaqian@gmail.com (Corresponding author)

DOI: 10.1587/transinf.E97.D.993

low bit rate (0.1577 bpp) JPEG image. It is observed that in the hat area, the color of a coding block shifts to another value, which is called color shift. By contrast, the color disappears in the forehead area. Both color shift and color disappearance are caused by the quantization operation in JPEG coding, which is lossy and irreversible. While color shift occurs in all JPEG images, color disappearance only occurs when heavy compression is applied to the images with large homogeneous areas. In this regard, color disappearance is a special case of color shift. It is also found that for low bit rate JPEG images, region merging occurs in both color-shifting and color-disappearing areas, which rarely occurs in natural images and is a sign of heavy blocking artifacts. Figure 1 (c) shows the color distortion map, where green denotes the CSA and blue denotes the CDA.

Color discontinuity is a representative feature of blocking artifact. In this letter, the characteristics of both color shift and color disappearance are utilized to estimate the blocking artifacts in images.

3. Blocking Artifact Metric

The diagram of the proposed method is shown in Fig. 2. The image is first converted into YUV color space, and the gradient is computed. Then CSA and CDA are identified, and the blockiness scores are computed in U channel and Y channel, respectively. Finally, an overall quality score is obtained.

The proposed method is applicable to the general luminance and chroma color models, for example YUV and YCbCr. Although YCbCr is the model that JPEG compression uses, we find that YUV produces slightly better results, which will be shown in the experiment section. Therefore, YUV color space is employed in this work. For an image in RGB model, it is first converted into YUV color space:

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B, \\ U &= -0.147R - 0.289G + 0.436B, \\ V &= 0.615R - 0.515G - 0.100B, \end{aligned} \tag{1}$$

where Y denotes the luminance component, U and V denote the chroma components. For CSA, the metric is computed in U channel. For CDA, the metric is computed in Y channel. While both U and V channels can be used for CSAs, we find that U channel has better capability to describe the image content. In order to demonstrate this point, an example is shown in Fig. 3 (images are from [7]).

It is easily observed from Fig. 3 that the image content is better represented in U channel. Therefore, the blocking artifact scores of the CDAs are computed in U channel.

Identification of CSA and CDA is done in HSI color space. For CDA, the H and S components should be zeros. Otherwise, it belongs to the CSA. For both CSA and CDA, the blocking artifact metrics are computed in gradient domain, so the gradient image is first calculated. For an $M \times N$ image $f(x, y)$, the gradient image is defined by

$$\begin{cases} G_h(i, j) = |f(i, j + 1) - f(i, j)| \\ G_v(i, j) = |f(i + 1, j) - f(i, j)|, \end{cases} \tag{2}$$

where i and j denote the row and column of the image, G_h and G_v denote the gradients in horizontal and vertical directions. In this work, the gradients of U channel are denoted by GU_h and GU_v , and the gradients of Y channel are denoted by GY_h and GY_v .

Since the blocking artifacts appear at the boundaries of coding blocks, the target blocks for blockiness estimation should cover these boundaries. Figure 4 shows the relation between two horizontally neighboring coding blocks and the target block in the gradient domain. For vertical direction, the target blocks can be determined similarly. The blocking

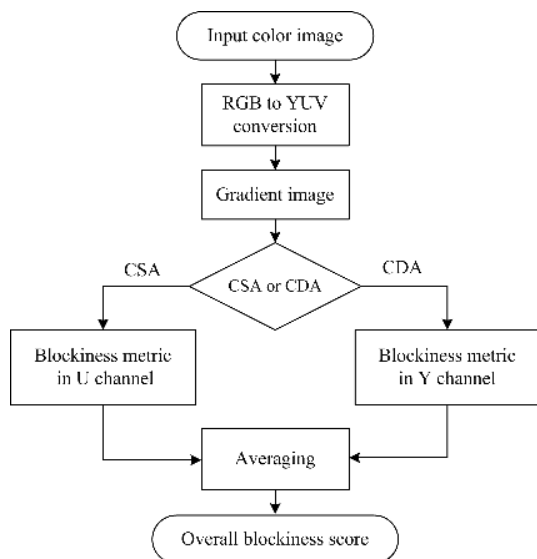


Fig. 2 Diagram of the proposed method.

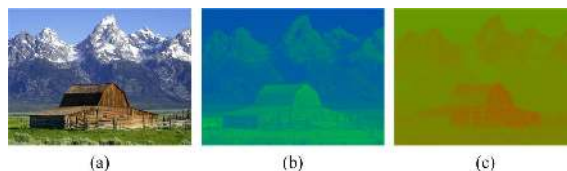


Fig. 3 Illustration of U channel and V channel. (a) Original image, (b) U channel, (c) V channel.

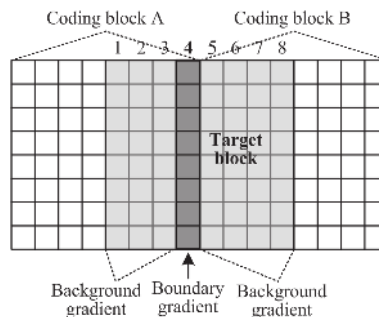


Fig. 4 Illustration of a target block in horizontal direction.

artifact scores are computed based on the target blocks.

In order to evaluate the blocking artifacts, a target block is further divided into boundary area and background area, as illustrated in Fig. 4. For a color-shifting block, the average of background gradients and relative gradient of the block boundary are first calculated:

$$BGU_h(i) = \frac{1}{7} \sum_{j=1,2,3,5,6,7,8} GU_h(i, j), \quad (3)$$

$$RGU_h(i) = GU_h(i, 4) - BGU_h(i). \quad (4)$$

where $BGU_h(i)$ and $RGU_h(i)$ denote the average of background gradients and relative gradient of the block boundary in horizontal direction, with $i = 1, 2, \dots, 8$. Then the blocking artifact score for each row of the target block is defined as

$$CS_h(i) = \begin{cases} 0.5, & \text{if } RGU_h(i) = 0, BGU_h(i) = 0 \\ RGU_h(i), & \text{if } RGU_h(i) \neq 0, BGU_h(i) = 0 \\ \frac{RGU_h(i)}{BGU_h(i)}, & \text{if } RGU_h(i) \neq 0, BGU_h(i) \neq 0, \end{cases} \quad (5)$$

where $i = 1, 2, \dots, 8$. The case $RGU_h(i) = BGU_h(i) = 0$ indicates that the gradients are all zeros for the target block. This rarely appears in natural images, so it belongs to the merged region, which is caused by heavy JPEG compression. Therefore, a fixed score 0.5 is set. The case $RGU_h(i) \neq 0, BGU_h(i) = 0$ indicates that the background pixels are constant and nonzero gradient only appears at the block boundary. In this case, the relative gradient of the boundary is adopted as the blockiness score. The last case, $RGU_h(i) \neq 0, BGU_h(i) \neq 0$ denotes the most common natural images, and the relative gradient ratio between the relative gradient and the averaged background gradient is computed as the blockiness score.

Another problem that should be considered is that when the blocking artifact exceeds a certain extent, human eyes tend to give a constant score. Based on this characteristic, a threshold $T = 1$ is applied to the blockiness scores, and the maximum quality score is set to 1:

$$CS_h(i) = \begin{cases} CS_h(i), & \text{if } CS_h(i) \leq T \\ 1, & \text{if } CS_h(i) > T. \end{cases} \quad (6)$$

Next, the eight scores are averaged to obtain the blockiness score of the target block:

$$CS_h = \frac{1}{8} \sum_{i=1}^8 CS_h(i). \quad (7)$$

The above operations are used to compute the blocking artifact score in the horizontal direction. The score in the vertical direction, CS_v , can be obtained similarly. By computing the average of all the scores in horizontal and vertical directions, the overall blockiness score for color-shifting area can be obtained, which is denoted by CSH and CSV . Then the average value of them is used to evaluate the blocking artifacts in the color-shifting area:

$$CS = (CSH + CSV)/2. \quad (8)$$

For the color-disappearing blocks, the blocking artifact scores are computed in Y channel similarly. The only difference applies to Eq. (5). Specifically, when $RGY_h(i) = BGY_h(i) = 0$, the blockiness score is set to 1 instead of 0.5, because for the region missing color, humans tend to believe that the blocking artifact is heavier. The blocking artifact score of the color-disappearing area is denoted by CD ,

Finally, an overall blocking artifact score is obtained:

$$\text{Score} = (CS + CD)/2. \quad (9)$$

The maximum value of the blocking artifact score is 1. High score indicates that the blocking artifact is severe.

4. Experimental Results

In this section, four image quality databases are employed to evaluate the performance of the proposed method, including LIVE [8], [9], MICT [10], IVC [11], [12] and CSIQ [13], [14]. The LIVE database contains 29 original images and 779 distorted images. There are five types of distortions in LIVE, including JPEG compression (169 images), JPEG2000 compression (175 images), additive Gaussian white noise (145 images), Gaussian blurring (145 images), and JPEG2000 with bit errors (145 images). The subjective scores are measured using difference mean opinion score (DMOS). The MICT database contains 14 original images and 196 distorted images. There are two types of distortions in this database: JPEG compression (98 images) and JPEG2000 compression (98 images). The subjective scores are measured using mean opinion score (MOS). The IVC database contains 10 reference images and 185 distorted images. There are five types of distortions in this database: JPEG compression (50 images), JPEG compression of only the luminance component (25 images), JPEG2000 compression (50 images), locally adaptive-resolution coding (40 images), and Gaussian blurring (20 images). The subjective scores are measured using MOS. The CSIQ database contains 30 reference images and 866 distorted images. There are six distortion types in this database: JPEG compression (150 images), JPEG2000 compression (150 images), additive Gaussian white noise (150 images), additive Gaussian pink noise (150 images), Gaussian blurring (150 images), and global contrast decrements (116 images). The subjective scores are measured using DMOS.

Since blocking artifact is the main source of quality degradation in JPEG compression, the JPEG images in each database are employed to conduct the experiments. Pearson correlation coefficient (CC) and the root mean-squared error (RMSE) are employed to evaluate the predication accuracy. Spearman rank-order correlation coefficient (SROCC) is employed to evaluate the predication monotonicity. These criterions are computed after a nonlinear fitting between the subjective scores and the predicted scores.

The first experiment is to test the performance of our method on YCbCr color space, which is used in JPEG compression. Table 1 lists the simulation results of the proposed method and the results on Cb channel of YCbCr model.

Table 1 Comparison of the results on YUV and YCbCr.

Database	Model	CC	RMSE	SROCC
LIVE	YUV	0.9628	6.5514	0.9420
	YCbCr	0.9625	6.5783	0.9413
MICT	YUV	0.9560	0.3871	0.9411
	YCbCr	0.9552	0.3903	0.9415
IVC	YUV	0.9468	0.3743	0.9351
	YCbCr	0.9497	0.3642	0.9399
CSIQ	YUV	0.9622	0.0833	0.9415
	YCbCr	0.9616	0.0840	0.9407

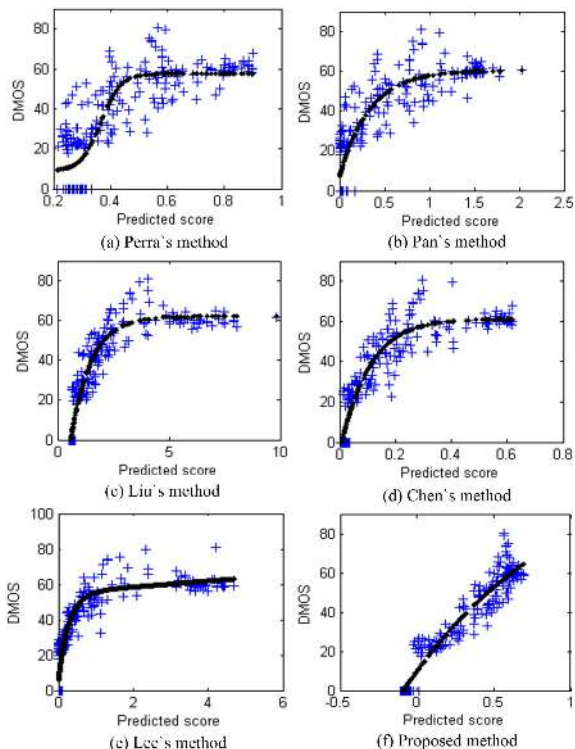


Fig. 5 Fitted curves for different methods on LIVE database.

It is observed from Table 1 that the results on YUV and YCbCr are quite similar. In fact, YCbCr is a scaled and offset version of YUV model. As our method calculates the blocking artifact score in gradient domain, scaling and offset have little effect on the overall performance. A further observation on the results indicates that YUV model produces slightly better results. Therefore, YUV color space is adopted in the proposed method.

The next experiment is to evaluate the overall performance of the proposed method. Five existing no-reference blocking artifact metrics are included for comparison, including Perrra's [2], Pan's [3], Liu's [4], Chen's [5] and Lee's [6]. Two popular full-reference image quality metrics, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [8], are also included to verify their abilities to evaluate the blocking artifacts.

The relations between the predicted scores and the subjective ratings are calculated to have an intuitive view of the prediction accuracy. Figure 5 shows the nonlinearly fitted curves of the proposed method and the compared five no-

Table 2 Simulation results on public databases.

Database	Metric	Type	CC	RMSE	SROCC
LIVE	PSNR	FR	0.8366	8.7591	0.8097
	SSIM [8]	FR	0.9682	6.0790	0.9587
	Perra [2]	NR	0.8570	12.5003	0.8268
	Pan [3]	NR	0.8887	11.1252	0.8728
	Liu [4]	NR	0.9455	7.8989	0.9278
	Chen [5]	NR	0.9356	8.5235	0.9193
	Lee [6]	NR	0.9426	8.0976	0.9296
	Proposed_G	NR	0.9636	6.4900	0.9429
	Proposed	NR	0.9628	6.5514	0.9420
MICT	PSNR	FR	0.5919	1.0636	0.5098
	SSIM [8]	FR	0.7457	0.8792	0.7392
	Perra [2]	NR	0.7993	0.7930	0.7517
	Pan [3]	NR	0.8350	0.7260	0.8253
	Liu [4]	NR	0.8130	0.7684	0.8137
	Chen [5]	NR	0.8381	0.7198	0.8228
	Lee [6]	NR	0.7625	0.8538	0.8097
	Proposed_G	NR	0.9517	0.4051	0.9392
	Proposed	NR	0.9560	0.3871	0.9411
IVC	PSNR	FR	0.6503	0.8838	0.6741
	SSIM [8]	FR	0.8323	0.6448	0.8067
	Perra [2]	NR	0.8046	0.6909	0.8035
	Pan [3]	NR	0.8017	0.6953	0.7974
	Liu [4]	NR	0.8291	0.6504	0.8296
	Chen [5]	NR	0.8835	0.5450	0.8840
	Lee [6]	NR	0.8982	0.5114	0.8968
	Proposed_G	NR	0.9351	0.4124	0.9225
	Proposed	NR	0.9468	0.3743	0.9351
CSIQ	PSNR	FR	0.8905	0.1392	0.8882
	SSIM [8]	FR	0.9403	0.1041	0.9223
	Perra [2]	NR	0.8904	0.1393	0.8518
	Pan [3]	NR	0.8888	0.1402	0.8624
	Liu [4]	NR	0.9483	0.0971	0.9151
	Chen [5]	NR	0.9414	0.1032	0.9227
	Lee [6]	NR	0.9767	0.0657	0.9479
	Proposed_G	NR	0.9506	0.0950	0.9290
	Proposed	NR	0.9622	0.0833	0.9415

reference methods on LIVE database. It is easily observed that the predicted scores of our method are more concentrated around the fitted curve, indicating that the predicted scores are highly consistent with the subjective ratings.

Based on the nonlinear fittings, CC, RMSE and SROCC are computed for each metric. Table 2 lists the experimental results on four databases. In order to demonstrate the advantage of evaluating the blocking artifacts in chroma component, the proposed method is also implemented only in gray channel, and the results are denoted by Proposed_G in Table 2. For each performance criterion, the three metrics producing the best results are highlighted in boldface.

From Table 2, we have the following findings:

(1) Comparison with PSNR and SSIM. PSNR and SSIM are full-reference image quality metrics. Due to the availability of reference images, they are expected to perform better than the proposed no-reference. However, the experimental results show that the proposed method outperforms PSNR and SSIM in MICT, IVC and

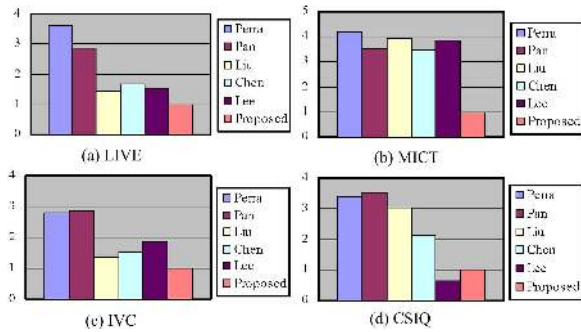


Fig. 6 F statistics on four databases.

CSIQ databases. In LIVE database, SSIM outperforms our method slightly. Therefore, the proposed method outperforms PSNR and SSIM, although they are full-reference image quality metrics.

(2) Comparison with Proposed.G. In MICT, IVC and CSIQ databases, the proposed method outperforms Proposed.G. In LIVE database, Proposed.G outperforms the color space implementation very slightly, with an amount of 0.0008. These results show that it is more effective to evaluate the blocking artifacts in chroma component.

(3) Comparison with peer no-reference metrics. For LIVE, MICT and IVC databases, the proposed method outperforms the other methods in terms of both prediction accuracy and prediction monotonicity. For CSIQ, Lee's method performs the best, and the proposed method ranks second. Therefore, the proposed method achieves the best overall performance.

Finally, we employ F test [15] to evaluate the statistical significance of each no-reference metric as compared to the proposed method. The F statistic between a compared metric and the proposed one is defined as:

$$F = \sigma_x^2 / \sigma_o^2, \quad (10)$$

where σ_x^2 and σ_o^2 denote the variances of predicted errors for the compared metric and the proposed metric, respectively.

Figure 6 shows the experimental results of F statistics on four databases. It is easily observed that the proposed method performs statistically the best, which is also consistent with the results in Table 2.

5. Conclusion

In JPEG compression, coarse quantization is applied to the color component. As a result, color discontinuities exist in the compressed images, which then causes the blocking artifacts. In this work, color shift and color disappearance, two representative forms of color discontinuity, are first analyzed. Then the characteristics of color discontinuities

are analyzed to generate the blocking artifact scores. Different from the existing methods that evaluate the blocking artifacts in illuminance domain, the proposed method achieves this goal in color space. Extensive experiments and comparisons demonstrate that the proposed method achieves very promising results.

Acknowledgments

The authors would like to thank the anonymous reviewer for the constructive comments to improve the quality of this paper. This work is supported in part by the Fundamental Research Funds for the Central Universities (2012QNA59) and National Natural Science Foundation of China (61379143, 51204175, 51204176, U1261105).

References

- [1] W.S. Lin and C.J. Kuo, "Perceptual visual quality metrics: A survey," *J. Vis. Commun. Image R.*, vol.22, no.4, pp.297–312, 2011.
- [2] C. Perra, F. Massidda, and D.D. Giusto, "Image blockiness evaluation based on Sobel operator," *Proc. IEEE Int. Conf. Image Processing*, pp.389–392, Sept. 2005.
- [3] F. Pan, X. Lin, S. Rahardja, E.P. Ong, and W.S. Lin, "Using edge direction information for measuring blocking artifacts of images," *Multidim. Syst. Sign. P.*, vol.18, no.4, pp.297–308, Dec. 2007.
- [4] H.T. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," *EURASIP J. Adv. Sig. Pr.*, vol.2009, Article ID: 263540, pp.1–14, 2009.
- [5] C.H. Chen and J.A. Bloom, "A blind reference-free blockiness measure," *Proc. Pacific-Rim Conf. Advances in Multimedia Information Processing*, pp.112–123, 2010.
- [6] S.W. Lee and S.J. Park, "A new image quality assessment method to detect and measure strength of blocking artifacts," *Signal Process.-Image*, vol.27, no.1, pp.31–38, Jan. 2012.
- [7] YUV color space, <http://en.wikipedia.org/wiki/YUV>
- [8] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol.13, no.4, pp.600–612, April 2004.
- [9] LIVE image quality assessment database release 2, <http://live.ece.utexas.edu/research/quality>
- [10] Y. Horita, K. Shibata, Y. Kawayoke, and Z.M.P. Sazzad, MICT: Image quality evaluation database 2000 [Online]. Available: <http://mict.eng.u-toyama.ac.jp/mict/index2.html>
- [11] A. Ninassi, P.L. Callet, and F. Atrusseau, "Pseudo no reference image quality metric using perceptual data hiding," *Proc. SPIE Human Vision and Electronic Imaging*, pp.146–157, San Jose, CA, USA, 2006.
- [12] Subjective quality assessment IRCCyN/IVC database, <http://www2.irccyn.ecnantes.fr/ivcdb>
- [13] E.C. Larson and D.M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electronic Imaging*, vol.19, no.1, 011006, pp.1–21, March 2010.
- [14] Categorical image quality database, <http://vision.okstate.edu/csiq>
- [15] C.Y. Wee, R. Paramesran, R. Mukundan, and X.D. Jiang, "Image quality assessment by discrete orthogonal moments," *Pattern Recognit.*, vol.43, no.12, pp.4055–4068, Dec. 2010.