

Evaluation of Image Quality Metrics for Color Prints

Marius Pedersen^{1,2}, Yuanlin Zheng^{1,3}, and Jon Yngve Hardeberg¹

¹ Gjøvik University College, Gjøvik, Norway

² Océ Print Logic Technologies S.A., Creteil, France

³ Xi'an University of Technology, Xi'an Shaanxi, China

Abstract. New technology is continuously proposed in the printing technology, and as a result the need to perform quality assessment is increasing. Subjective assessment of quality is tiresome and expensive, the use of objective methods have therefore become more and more popular. One type of objective assessment that has been subject for extensive research is image quality metrics. However, so far no one has been able to propose a metric fully correlated with the percept. Pedersen et al. (J Elec Imag 19(1):011016, 2010) proposed a set of quality attributes with the intention of being used with image quality metrics. These quality attributes are the starting point for this work, where we evaluate image quality metrics for them, with the goal of proposing suitable metrics for each quality attribute. Experimental results show that suitable metrics are found for the sharpness, lightness, artifacts, and contrast attributes, while none of the evaluated metrics correlate with the percept for the color attribute.

Keywords: Image quality, metrics, print quality, quality attributes, color printing.

1 Introduction

Image Quality (IQ) assessment is an important part in the printing industry. The introduction of new technology and products require assessment of quality to see if the quality is improved over the current technology. When observers judge IQ they base their decision a number of quality attributes, such as colorfulness, contrast, and sharpness. Many researchers have been investigating the importance of different quality attributes and their influence on IQ [14,20,19,24,23]. Knowledge about the importance of quality attributes can be used to achieve an optimal reproduction of an image [8]. However, evaluating all quality attributes in the literature is not practical, therefore most researchers evaluate a subset of quality attributes. A subset of quality attributes helps reduce the complexity of IQ, and the strengths and weaknesses of a system can be modeled using only a few parameters. Recently, Pedersen et al. [24,23] proposed a set of six Color Printing Quality Attributes (CPQAs) for the evaluation of print quality:

- The **color** CPQA contains aspects related to color such as hue, saturation, and color rendition, except lightness.
- The **lightness** CPQA is considered so perceptually important that it is beneficial to separate it from the color CPQA. Lightness ranges from light to dark.

- The **contrast** CPQA can be described as the perceived magnitude of visually meaningful differences, global and local, in lightness and chromaticity within the image.
- The **sharpness** CPQA is related to the clarity of details and definition of edges.
- The **artifacts** CPQA includes noise, contouring, and banding. In color printing, some artifacts can be perceived in the resulting image. These artifacts can degrade the quality of an image if they are detectable.
- The **physical** CPQA contains all physical parameters that affect quality, such as paper properties and gloss.

These were proposed with the intention of being used in both subjective and objective evaluation of quality. Validation of the CPQAs showed that they were suitable to evaluate IQ [24,26]. Not long ago, Pedersen et al. [25,27] evaluated IQ metrics for each CPQA. Their evaluation indicated that metrics based on structural similarity gave good results for the sharpness, contrast, and lightness CPQAs, but for the other CPQAs the results were inconclusive. The conclusion was that further evaluation was needed in order to find suitable metrics to assess the quality of the CPQAs. We continue this work and evaluate IQ metrics for the CPQAs, with the intention of proposing suitable metrics for each CPQA. This work is considered as a part of our long term goal to be able to assess quality without being dependent on human observers.

The remainder of the paper is organized as follows: in the next section we introduce the experimental setup, before we evaluate a set of metrics against the perceptual data from the experiment. Finally we conclude and propose future work.

2 Experimental Setup

We want to investigate the relationship between the percept of the CPQAs and IQ metrics. In order to do this we have carried out an experiment where human observers judge the quality of the CPQAs on a set of printed images.

2.1 Test Images

Ten images (Figure 1) were selected from the ISO standards [12,13]. The number of images follow the recommendation by Field [9], who recommend between five and ten images, and the CIE [5], who recommend at least four images. The images were selected to cover a wide range of characteristics, such as lightness from low to high levels, saturation from low to high levels, contrast from low to high levels, hue primaries, fine details, memory colors as skin tones. These different characteristics will ensure evaluation of many different aspects of IQ.

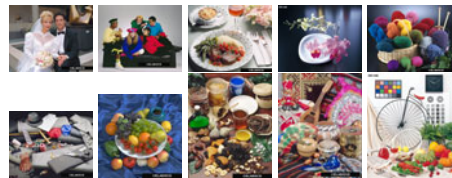


Fig. 1. The ten test images used in the experiment. Each reproduced with four different settings.

2.2 Printing Workflow

Firstly, the color space of all the images was changed to sRGB to define the reference images. Secondly, then the color space was changed to CMYK using the output profile that was generated using a TC3.5 CMYK test target, measured with a GretagMacbeth Eye-One Pro spectrophotometer and generated with ProfileMaker Pro 5.0.8. Finally the CMYK images were printed by a HP DesignJet 10ps printer with the HP software RIP v2.1.1 using four different modes: the best print mode, with the resolution of 1200x1200, and the perceptual intent (abbr. BP), the best mode and relative colorimetric intent (abbr. BR), normal print mode, with the resolution of 600x600 and the perceptual intent (abbr. NP), and the last with normal print mode and relative colorimetric intent (abbr. NR). This resulted in the ten images having four different reproductions, giving a total of 40 images for the observers to judge.

2.3 Observers

Ten observers participated in the experiment, all had normal vision without visual deficits. There were 3 females and 7 males with an average age of 23 years.

2.4 Viewing Conditions

The observers were presented with a reference image on an EIZO ColorEdge CG224 at a color temperature of 6500 K and luminance level of 80 cd/m². The image set was rendered for sRGB display, and therefore a monitor capable of displaying the sRGB gamut was the most adapted reproduction device for the set of images. A hood was fitted to the monitor to prevent glare. The printed images were presented randomly in a controlled viewing room at a color temperature of 5200 K, an illuminance level of 450 ± 75 lux and a color rendering index of 96. The observers viewed the reference image and the printed image simultaneously from a distance of approximately 60 cm. The experiment followed the CIE guidelines [5] as closely as possible.

2.5 Experiment Procedure

The observers were asked to compare one image selected from the ten images at random to its four prints. Sharpness quality, color quality, lightness quality, contrast quality, artifacts quality, and the quality of the main characteristics were evaluated on a five step scale, where 1 indicated best quality and 5 the worst quality. The physical CPQA was not evaluated since no physical parameter was changed.

3 Experimental Results

From the experiment z-scores were calculated using the color engineering toolbox [10], which indicated the perceived differences between the four reproductions. These z-scores were calculated for each CPQA and the main characteristics, both image-wise and for the complete dataset.

It has been suggested in the literature that some regions of the image is more important than others [30,18,43]. In order to investigate the relationship between the CPQAs and different regions of the image, we have calculated the Pearson correlation coefficients [15] between the main characteristics and the CPQAs. This analysis would reveal if the quality of the CPQAs are related to the quality of main characteristics (region-of-interest). From Table 1 we can see that in the different reproductions the main characteristics have varying correlation coefficients with the CPQAs. This indicates that the quality of the CPQAs are not directly linked with main characteristics, but that other characteristics are important for the impression of quality of most CPQAs. However, for some CPQAs and printing modes we see a high correlation between the main characteristics and the CPQAs, this might indicate that IQ metrics performing a weighting of regions could be more suitable than those assigning equal weight to the entire image.

Table 1. Pearson correlation between z-scores of the main characteristics and the z-scores of the CPQAs for each printing mode and for all modes

Mode	CPQAs				
	Color	Lightness	Sharpness	Contrast	Artifacts
BP	0.85	0.47	0.55	0.92	0.28
BR	0.72	0.45	0.48	0.78	0.55
NP	-0.02	0.60	0.30	0.61	0.71
NR	0.31	0.29	0.31	0.88	0.60
All	0.79	0.77	0.71	0.89	0.77

4 Evaluation of Image Quality Metrics

Our long term goal is to be able to automatically evaluate IQ through the CPQAs, more specifically using IQ metrics. In this part we evaluate a set of IQ metrics for each CPQA against the perceptual data from the experiment.

4.1 Preparation of the Printed Images

The printed images cannot be directly used with IQ metrics, since the metrics require a digital input. Because of this the images need to be digitized. To perform this we have adopted the framework by Pedersen and Amirshahi [22]. First the images were scanned at a resolution of 600 dpi using an HP ScanJet G4050. The scanner was characterized with the same test target as used to generate the printer profile. Since the experiment was carried out under mixed illumination, the CIECAM02 chromatic adaptation transform [6] was used to ensure consistency in the calculations for the metrics. The CIE guidelines were followed [6], using the measured reference white point of the monitor and the media were used as input to the adaptation transform.

4.2 Selected Image Quality Metrics

There are a number of IQ metrics proposed in the literature [31]. We cannot evaluate all of these, and because of this we have made a selection based on previous

Table 2. Selected IQ metrics for the evaluation of CPQAs

Metric \ CPQA	Sharpness	Color	Lightness	Contrast	Artifacts
ABF [38]		X	X		X
Busyness [21]	X		X		
blurMetric [7]	X				
Cao [3]	X				X
CW-SSIM[40]	X		X	X	X
Δ LC [2]	X		X	X	X
IW-SSIM [39]	X		X	X	X
LinLab [16]		X	X		X
MS-SSIM [42]	X		X	X	X
M-SVD [34]	X		X		X
PSNR-HVS-M[32]	X		X		X
PSNR-HVS [32]	X		X		X
RFSIM [44]	X		X	X	X
RRIQA [41]	X		X	X	X
S-CIELAB [45]		X	X		X
S-DEE [35]		X	X		X
SHAME [29]		X	X		X
SHAME-II [29]		X	X		X
SSIM [37]	X		X	X	X
VIF [33]	X		X	X	X
VSNR [4]	X		X		X
WLF [36]				X	X
YCXCzLab [17]		X	X		X

evaluations [1,11,22,25,27], the criteria on which the metrics were created, guidelines for metrics for CPQAs [27], and their popularity. Since many of the metrics are designed to account for specific aspects, only the ones suitable for a given CPQA is evaluated. An overview of the 23 metrics selected for the evaluation and the CPQAs they evaluate is found in Table 2.

4.3 Evaluation Method

Three different methods were adopted for the evaluation of the IQ metrics. In order to evaluate all aspects of the metrics we will investigate the performance of the IQ metrics both image by image, and the overall performance over the entire set of images. The Pearson correlation [15] is used for the image-wise evaluation, comparing the calculated quality and observed quality. The mean of the correlation for each image in the dataset and the percentage of images with a correlation above 0.6 is used as a measure of performance. Overall performance is also an important aspect, and for this evaluation we will use the rank order method [28], where the correlation between the z-scores from the observers and the z-scores of the metric is the indication of performance. With only four data points it is important to carry out visual inspections of the z-scores to validate the correlation values.

4.4 Evaluation Results

Due to many IQ metrics and several CPQAs we will only show the results of the best performing metrics for each CPQA.

Sharpness. For sharpness the Structural SIMilarity (SSIM) based metrics perform well (Table 3). The Multi-Scale SSIM (MS-SSIM) has the highest mean correlation with 0.73 and the highest number of images with a correlation above 0.6. It also performs among the best for the rank order correlation. The results show that metrics based on structural similarity are well-suited to measure perceived sharpness quality. However, other approaches as the Δ LC and the Riesz-transform based Feature SIMilarity metric (RFSIM) have very good performance, indicating that these might be suitable as well.

Table 3. Evaluation of IQ metrics for the sharpness CPQA

Metric	Mean correlation	Above 0.6	Rank order	
			Correlation	p-value
CW-SSIM	0.66	70	0.94	0.06
Δ LC	0.43	50	1.00	0.00
IW-SSIM	0.56	70	0.89	0.11
MS-SSIM	0.73	80	0.94	0.06
RFSIM	0.61	70	0.97	0.03
SSIM	0.66	80	0.96	0.04

Table 4. Evaluation of metrics for the color CPQA. Color indicates the color part of the metric.

Metric	Mean correlation	Above 0.6	Rank order	
			Correlation	p-value
ABF	0.07	0	0.23	0.77
LinLab	-0.09	0	0.04	0.96
SCIELAB	-0.27	0	-0.24	0.76
S-DEE _{Color}	-0.38	0	-0.35	0.65
SHAME	0.01	10	0.10	0.90
SHAME _{Color}	0.05	20	0.12	0.88
SHAMEII	0.23	30	0.27	0.73
YCxCzLab	0.24	30	0.33	0.67

Color. For the color CPQA none of the evaluated metrics perform well (Table 4). It should be noted that all of these metrics are based on color differences, and this might be an indication that using only the color difference from the original is not enough to predict perceived color quality. The color CPQA had a fairly high correlation for all modes between the main characteristic and perceived IQ (Table 1), which might indicate that metrics giving more importance to certain regions, such as SHAME and SHAME-II, could perform better than the metrics that equally weight the entire image. The experimental results in Table 4 shows that these metrics do not outperform other metrics.

Lightness. The SSIM based metrics perform very well for the lightness attribute (Table 5), the Complex Wavelet SSIM (CW-SSIM) has a mean correlation 0.86 and all images have a correlation above 0.6. However, other metrics also perform well, such as the RFSIM, Δ LC, Spatial-DEE with only the lightness part (S-DEE_{Lightness}) and Adaptive Bilateral Filter with only the lightness part (ABF_{Lightness}). The results indicate that any of these are appropriate to measure lightness quality. These metrics take different approaches to measure lightness quality, indicating that different strategies are suitable.

Table 5. Evaluation of metrics for the lightness CPQA. Lightness indicates the lightness part of the metric.

Metric	Mean correlation	Above 0.6	Rank order	
			Correlation	p-value
ABF _{Lightness}	0.69	80	0.87	0.13
CW-SSIM	0.86	100	0.93	0.07
Δ LC	0.69	80	0.99	0.01
IW-SSIM	0.85	80	0.95	0.05
MS-SSIM	0.82	90	0.93	0.07
RFSIM	0.86	90	1.00	0.00
S-DEE _{Lightness}	0.80	90	0.89	0.11
SSIM	0.63	70	0.98	0.02

Table 6. Evaluation of metrics for the contrast CPQA

Metric	Mean correlation	Above 0.6	Rank order	
			Correlation	p-value
CW-SSIM	0.72	90	1.00	0.00
IW-SSIM	0.59	70	0.94	0.06
MS-SSIM	0.72	80	1.00	0.00
RFSIM	0.67	80	0.96	0.04
SSIM	0.65	70	0.99	0.01

Contrast. Many metrics perform well for the contrast CPQA (Table 6). The SSIM based metrics all have a correlation above 0.6 in more than 70% of the images, they also have a high mean correlation and excellent rank order correlation. The RFSIM has a similar performance to the SSIM based metrics. All of these metrics would be appropriate for measuring contrast. One should notice that all of the well performing metrics for contrast are based on lightness, and none of them take color information into account. This might make them inappropriate to measure contrast in images where color strongly contributes to the impression of contrast.

Table 7. Evaluation of metrics for the artifacts CPQA

Metric	Mean correlation	Above 0.6	Rank order	
			Correlation	p-value
CW-SSIM	0.83	90	0.97	0.03
Δ LC	0.72	70	0.94	0.06
IW-SSIM	0.83	90	0.99	0.01
MS-SSIM	0.77	90	0.97	0.03
RFSIM	0.82	90	0.99	0.01
SSIM	0.60	70	1.00	0.00

Artifacts. The performance for the artifacts CPQA (Table 7) follow the results of many of the other CPQAs. The SSIM based metrics perform well together with Δ LC and RFSIM. There are only minor differences between these, and any of them seem to be suitable to measure artifacts. However, artifacts can vary significantly and to measure specific artifacts specially designed metrics might be required.

5 Conclusion and Future Work

In this research we focused on quality attributes for automatic assessment of print quality. We evaluated a set of image quality metrics for a set of quality attributes, with the intention of proposing suitable metrics for each attribute. The experimental results show that structural similarity based metrics perform well for the sharpness, contrast, and artifacts attributes, but for the color attribute none of the evaluated metrics correlated with the percept, and for the lightness attribute many different metrics perform well.

Future work should include further investigation of the color attribute in order to find a suitable metric. Another important issue is how to combine the results from the attributes to obtain one number representing overall image quality.

Acknowledgments

The author hereof has been enabled by Océ-Technologies B.V. to perform research activities which underlies this document. This document has been written in a personal capacity. Océ-Technologies B.V. disclaims any liability for the correctness of the data, considerations and conclusions contained in this document.

References

1. Ajagamelle, S.A., Pedersen, M., Simone, G.: Analysis of the difference of gaussians model in image difference metrics. In: 5th European Conference on Colour in Graphics, Imaging, and Vision (CGIV), pp. 489–496. IS&T, Joensuu (2010)
2. Baranczuk, Z., Zolliker, P., Giesen, J.: Image quality measures for evaluating gamut mapping. In: Color Imaging Conference, pp. 21–26. IS&T/SID, Albuquerque (2009)
3. Cao, G., Pedersen, M., Baranczuk, Z.: Saliency models as gamut-mapping artifact detectors. In: 5th European Conference on Colour in Graphics, Imaging, and Vision (CGIV), pp. 437–443. IS&T, Joensuu (2010)
4. Chandler, D., Hemami, S.: VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing* 16(9), 2284–2298 (2007)
5. CIE: Guidelines for the evaluation of gamut mapping algorithms. Tech. Rep., CIE TC8-03 (156:2004) ISBN: 3-901-906-26-6
6. CIE: Chromatic adaptation under mixed illumination condition when comparing softcopy and hardcopy images. Tech. Rep., CIE TC8-04 (162:2004) ISBN: 3-901-906-34-7
7. Crete, F., Dolmiere, T., Ladret, P., Nicolas, M.: The blur effect: perception and estimation with a new no-reference perceptual blur metric. In: Rogowitz, B.E., Pappas, T.N., Daly, S.J. (eds.) *Proceedings of SPIE Human Vision and Electronic Imaging XII*, vol. 6492, p. 64920I (March 2007)
8. Fedorovskaya, E.A., Blommaert, F., de Ridder, H.: Perceptual quality of color images of natural scenes transformed in CIELUV color space. In: Color Imaging Conference, pp. 37–40. IS&T/SID (1993)
9. Field, G.G.: Test image design guidelines for color quality evaluation. In: Color Imaging Conference, pp. 194–196. IS&T, Scottsdale (1999)
10. Green, P., MacDonald, L. (eds.): *Colour Engineering: Achieving Device Independent Colour*. John Wiley & Sons, Chichester (2002)

11. Hardeberg, J., Bando, E., Pedersen, M.: Evaluating colour image difference metrics for gamut-mapped images. *Coloration Technology* 124(4), 243–253 (2008)
12. ISO: ISO 12640-2: Graphic technology - prepress digital data exchange - part 2: XYZ/sRGB encoded standard colour image data (XYZ/SCID) (2004)
13. ISO: ISO 12640-3 graphic technology - prepress digital data exchange - part 3: CIELAB standard colour image data (CIELAB/SCID) (2007)
14. Keelan, B.W.: *Handbook of Image Quality: Characterization and Prediction*. Marcel Dekker, New York (2002)
15. Kendall, M.G., Stuart, A., Ord, J.K.: *Kendall's Advanced Theory of Statistics: Classical inference and relationship*, 5th edn., vol. 2. A Hodder Arnold Publication (1991)
16. Kolpatzik, B., Bouman, C.: Optimized error diffusion for high-quality image display. *Journal of Electronic Imaging* 1(3), 277–292 (1992)
17. Kolpatzik, B., Bouman, C.: Optimal universal color palette design for error diffusion. *Journal of Electronic Imaging* 4(2), 131–143 (1995)
18. Larson, E.C., Chandler, D.M.: Unveiling relationships between regions of interest and image fidelity metrics. In: Pearlman, W.A., Woods, J.W., Lu, L. (eds.) *Visual Communications and Image Processing*. SPIE Proceedings, vol. 6822, pp. 68222A–68222A–16. SPIE, San Jose (2008)
19. Lindberg, S.: *Perceptual determinants of print quality*. Ph.D. thesis, Stockholm University (2004)
20. Norberg, O., Westin, P., Lindberg, S., Klamann, M., Eidenvall, L.: A comparison of print quality between digital, offset and flexographic printing presses performed on different paper qualities. In: *International Conference on Digital Production Printing and Industrial Applications*, pp. 380–385. IS&Ts (May 2001)
21. Orfanidou, M., Triantaphillidou, S., Allen, E.: Predicting image quality using a modular image difference model. In: Farnand, S.P., Gaykema, F. (eds.) *Image Quality and System Performance V*. SPIE Proceedings, vol. 6808, pp. 68080F–68080F–12. SPIE/IS&T, San Jose, USA (2008)
22. Pedersen, M., Amirshahi, S.: A modified framework the evaluation of color prints using image quality metrics. In: *5th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pp. 75–82. IS&T, Joensuu (2010)
23. Pedersen, M., Bonnier, N., Hardeberg, J.Y., Albrechtsen, F.: Attributes of a new image quality model for color prints. In: *Color Imaging Conference*, pp. 204–209. IS&T, Albuquerque (2009)
24. Pedersen, M., Bonnier, N., Hardeberg, J.Y., Albrechtsen, F.: Attributes of image quality for color prints. *Journal of Electronic Imaging* 19(1), 011016–1–13 (2010)
25. Pedersen, M., Bonnier, N., Hardeberg, J.Y., Albrechtsen, F.: Estimating print quality attributes by image quality metrics. In: *Color and Imaging Conference*, pp. 68–73. IS&T/SID, San Antonio (2010)
26. Pedersen, M., Bonnier, N., Hardeberg, J.Y., Albrechtsen, F.: Validation of quality attributes for evaluation of color prints. In: *Color and Imaging Conference*, pp. 74–79. IS&T/SID, San Antonio (2010)
27. Pedersen, M., Bonnier, N., Hardeberg, J.Y., Albrechtsen, F.: Image quality metrics for the evaluation of print quality. In: Gaykema, F., Farnand, S. (eds.) *Image Quality and System Performance*. Proceedings of SPIE. SPIE, San Francisco (2011)
28. Pedersen, M., Hardeberg, J.Y.: Rank order and image difference metrics. In: *4th European Conference on Colour in Graphics, Imaging, and Vision (CGIV)*, pp. 120–125. IS&T, Terrassa (2008)
29. Pedersen, M., Hardeberg, J.Y.: A new spatial hue angle metric for perceptual image difference. In: Trémeau, A., Schettini, R., Tominaga, S. (eds.) *CCIW 2009*. LNCS, vol. 5646, pp. 81–90. Springer, Heidelberg (2009)

30. Pedersen, M., Hardeberg, J.Y., Nussbaum, P.: Using gaze information to improve image difference metrics. In: Rogowitz, B., Pappas, T. (eds.) *Human Vision and Electronic Imaging VIII*, San Jose, CA, USA. SPIE Proceedings, vol. 6806, p. 680611 (January 2008)
31. Pedersen, M., Hardeberg, J.: Survey of full-reference image quality metrics. Høgskolen i Gjøviks rapportserie 5, The Norwegian Color Research Laboratory (Gjøvik University College) (June 2009) ISSN: 1890-520X
32. Ponomarenko, N., Silvestri, F., Egiazarian, K., Carli, M., Astola, J., Lukin, V.: On between-coefficient contrast masking of DCT basis functions. In: *Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics VPQM 2007*, Scottsdale, Arizona, USA, pp. 1–4 (January 2007)
33. Sheikh, H.R., Bovik, A.C.: Image information and visual quality. *IEEE Transactions on Image Processing* 15(2), 430–444 (2006)
34. Shnayderman, A., Gusev, A., Eskicioglu, A.M.: An SVD-based grayscale image quality measure for local and global assessment. *IEEE Transactions On Image Processing* 15(2), 422–429 (2006)
35. Simone, G., Oleari, C., Farup, I.: Performance of the euclidean color-difference formula in log-compressed OSA-UCS space applied to modified-image-difference metrics. In: *11th Congress of the International Colour Association (AIC)*, Sydney, Australia (October 2009)
36. Simone, G., Pedersen, M., Hardeberg, J.Y., Rizzi, A.: Measuring perceptual contrast in a multilevel framework. In: Rogowitz, B.E., Pappas, T.N. (eds.) *Human Vision and Electronic Imaging XIV*, vol. 7240. SPIE, San Jose (2009)
37. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13(4), 600–612 (2004)
38. Wang, Z., Hardeberg, J.Y.: An adaptive bilateral filter for predicting color image difference. In: *Color Imaging Conference*, pp. 27–31. IS&T/SID, Albuquerque, NM, USA (2009)
39. Wang, Z., Li, Q.: Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing* (2010)
40. Wang, Z., Simoncelli, E.: Translation insensitive image similarity in complex wavelet domain. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 2, pp. 573–576 (2005)
41. Wang, Z., Simoncelli, E.P.: Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. In: *Human Vision and Electronic Imaging X. Proceedings of SPIE*, vol. 5666, pp. 149–159. SPIE, San Jose (January 2005)
42. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multi-scale structural similarity for image quality assessment. In: *Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computers*, pp. 1398–1402 (November 2003)
43. Wang, Z., Bovik, A.C., Lu, L.: Wavelet-based foveated image quality measurement for region of interest image coding. In: *International Conference on Image Processing*, pp. 89–92. IEEE, Los Alamitos (2001)
44. Zhang, L., Zhang, L., Mou, X.: RFSIM: A feature based image quality assessment metric using riesz transforms. In: *International Conference on Image Processing*, Hong Kong, pp. 321–324 (September 2010)
45. Zhang, X., Farrell, J., Wandell, B.: Application of a spatial extension to CIELAB. In: *Very high resolution and quality imaging II*, San Jose, CA, USA. SPIE Proceedings, vol. 3025, pp. 154–157 (February 1997)