Journal of Imaging Science and Technology® 61(4): 040402-1–040402-8, 2017.

Image Quality Metrics for the Evaluation and Optimization of Capsule Video Endoscopy Enhancement Techniques

Marius Pedersen▲, Olga Cherepkova, and Ahmed Mohammed
The Norwegian Colour and Visual Computing Laboratory, Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjøvik, Norway
E-mail: marius.pedersen@ntnu.no

Abstract. *Capsule endoscopy, using a wireless camera to capture the digestive track, is becoming a popular alternative to traditional colonoscopy. The images obtained from a capsule have lower quality compared to traditional colonoscopy, and high-quality images are required by medical doctors in order to set an accurate diagnosis. Over the last years several enhancement techniques have been proposed to improve the quality of capsule images. In order to verify that the capsule images have the required diagnostic quality some kind of quality assessment is required. In this work, the authors evaluate state-of-the-art no-reference image quality metrics for capsule video endoscopy. Furthermore, they use the best performing metric to optimize one of the capsule video endoscopy enhancement methods and validate through subjective experiment.*

[DOI: 10.2352/J.ImagingSci.Technol.2017.61.4.040402]

INTRODUCTION

Capsule video endoscopy has proven to be a powerful tool for diagnosis of the digestive tract diseases. It has many advantages over traditional colonoscopy,[1] as being less invasive, the lack of a requirement for sedation, and no need for gas insufflation. Also, since it is less invasive it might also increase participation in colorectal cancer screening. Capsule video endoscopy has been used to diagnose inflammatory bowel disease[2] (i.e., Crohn disease and ulcerative colitis), gastrointestinal bleeding, and polyps. It has been shown to have a high sensitivity for the detection of clinically relevant lesions.[3] The capsule itself is about 11 mm × 32 mm and usually captures images at a rate of 4 frames per second.[4] The images are at a lower resolution compared to traditional colonoscopy (usually full-HD). The images produced by capsule video endoscopy suffers from several problems, such as uneven illumination, low resolution, images taken under low illumination, high compression ratio, and noise. The problem of capsule image quality enhancement has been an active research topic since capsules appeared commercially in 2006. Enhancement techniques for capsule video endoscopy can be categorized based on the image attributes they focus on for accurate diagnosis of pathologies. There are four main categories: (1) making blood vessels visible; (2) removing or de-emphasizing specular reflections and illumination variation; (3) making tissues visible; and (4) keeping the original color tone.

For the first category, the Flexible spectral imaging color enhancement (FICE) emphasizes certain ranges of wavelengths by spectral decompositions.[5] Narrow Band Imaging[6] uses a set of filters to restrict the incident light to two narrow bands of wavelengths. For the second category, Ramaraj et al.,[7] proposed a homomorphic filtering technique to deal with uneven illumination, and claimed better results than contrast limited adaptive histogram equalization (CLAHE).[8] Okuhata et al.[9] applied the retinex theory for capsule video endoscopy enhancement of illumination variation. For the third category, Gopi et al.[10] proposed an image denoising technique using dual tree double density complex wavelet transform. A de-blurring technique based on total variation minimization was proposed by Liu et al.[11] Li et al.[12] proposed a method using adaptive contrast diffusion for enhancing tissue details. Ahmed et al.[13] proposed a method for enhancing the visibility of detail and shadowed tissue surfaces using concentric circles at each pixel for random walks combined with stochastic sampling. For the last category, Vu et al.[14] proposed an image enhancement technique that preserves the original color tones. Imtiaz et al.[15] used a sigmoidal function and space-variant color reproduction for enhancement.

In order to evaluate if the enhancement techniques improve the quality of the image, and leading to a more accurate diagnosis by the doctor, quality assessment methods for capsule video endoscopy are required. So far, studies have relied on psychometric experiments with doctors to evaluate the quality of the enhancement techniques.[13] In this article, we evaluate objective image quality metrics against the results from a psychometric experiment, with the goal of finding a suitable image quality metric to assess the quality of capsule video endoscopy. Further, we use the best performing image quality metric to optimize the parameters of one enhancement technique in order to produce higher quality images. The results of the optimization is verified by a medical doctor.

This article is organized as follows: first we introduce relevant image quality metrics, then we present the experimental setup, followed by results and discussion, and finally optimization of an enhancement method before we conclude.

▲ IS&T Member.

Received Feb. 27, 2017; accepted for publication May 30, 2017; published online June 30, 2017. Associate Editor: Jang Jin Yoo.
1062-3701/2017/61(4)/040402/8/$25.00

BACKGROUND

Image quality metrics can be classified depending on the availability of the reference: full reference, reduced reference, and no reference. In full reference[16] the complete reference is used in the assessment of quality, in reduced reference we have partial information about the reference image, and in no reference[17,18] a reference is not available. In capsule video endoscopy a reference image does not exist, and therefore only no-reference image quality metrics are suitable to evaluate capsule video endoscopic images. We present some of the relevant no-reference metrics for this work.

Blind Image quality index (BIQI)[19] is based on natural scene statistics. BIQI first identifies the presence of a distortion in the image, where the standard set of distortions includes JPEG, JPEG2000, white noise, Gaussian Blur, and Fast fading. The amount or probability of each distortion in the image is measured. Further, the quality of the image is measured along these distortions. Then the overall quality is a probability-weighted summation of the different quality values.

Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE)[20] is a natural scene statistic-based distortion-generic blind/no-reference image quality metrics which operates in the spatial domain. BRISQUE does not compute distortion specific features such as blocking, blur, ringing, but rather uses scene statistics of locally normalized luminance coefficients to quantify possible losses of "naturalness" in the image due to the presence of distortions. It is based on features that are derived from an empirical distribution of locally normalized luminances and products of locally normalized luminances under a natural scene statistic model. BRISQUE is also considered to have low complexity, which makes it ideal for optimization or real-time applications. It has, among others, shown to work well for image denoising.[20]

The **BLind Image Integrity Notator using DCT Statistics-II** (BLIINDS2)[21] is a general-purpose, non-distortion specific, blind/no-reference image quality metric that uses natural scene statistics models of discrete cosine transform (DCT) coefficients to perform distortion-agnostic quality assessment. The metric uses a small number of features, that are computed from a natural scene statistics model of block DCT coefficients. These features are fed to a regression function that computes the final image quality scores.

The **Spatial-Spectral Entropy-based Quality index** (SSEQ)[22] takes advantage of local spatial and spectral entropy features of distorted images to form quality aware features. The features then feed to a regression machine to predict image quality. SSEQ has shown to be capable of assessing the quality of images across multiple distortion categories. In addition, the metric has relatively low complexity.

The **Just Noticeable Blur Metric** (JNBM)[23] is a perceptual-based no-reference objective image sharpness/blurriness metric developed by integrating the concept of just noticeable blur into a probability summation model. The just noticeable blur is the minimum amount of blurriness perceived around an edge with contrast higher than just noticeable difference.

The **No-reference Image Quality Metric for Contrast distortion** (NIQMC)[24] performs quality estimation both locally and globally. Locally, the metric focuses on regions with much information, and computes entropy in these regions. Globally, the image histogram is compared with the uniformly distributed histogram of maximum information via the symmetric Kullback–Leibler divergence.

The **Sharpness Index** (SI)[25] is a variant of the global phase coherence sharpness metric,[26] which instead if using random phase images, the equivalent Gaussian random fields are used. The metric is defined by reference to the regularity (total variation) of Gaussian random fields. The metric has been used for optimization of images,[25] having the advantage of being fast and giving good results.

Anisotropy-based Quality estimation of images (AQI)[27] uses an anisotropic measure to assess the quality of images. AQI measures the variance of the expected entropy of an image upon a set of predefined directions. The metric was made for natural images, and showed similar performance as peak signal-to-noise ratio.

The **no-reference JPEG metric** (JPEG-S)[28] is a computationally inexpensive and memory efficient feature extraction method. Two features, blockiness and activity, are calculated horizontally and vertically, and combined to obtained an overall quality score. The metric is both computationally and memory efficient.

Oriented Gradients Image Quality Assessment (OG-IQA)[29] extracts a 6-dimensional relative gradient feature vector from the image; further it uses an AdaBoosting back-propagation neural network to map image features to image quality. OG-IQA has been shown to be rather independent of the database and have low complexity.

The **No-Reference Perceptually Weighted local Noise** (NRPWN) metric[30] is based on integrating noise variance estimation and contrast sensitivity thresholds into a probability summation model. The metric was shown to perform well for different types of noise.

The **Discrete Cosine Transform Statistics Prediction** (DCTSP) metric[31] is made specifically for blur. It is based on block-based DCT statistics and a linear prediction method. The metric was optimized for the LIVE database,[32] but showed good results in CSIQ[33] and TID2008[34] as well.

The **Cumulative Probability of Blur Detection** (CPBD) metric[35] is a no-reference sharpness metric based on a cumulative probability of blur detection. CPBD splits the image into blocks, and only blocks classified as edge blocks are processed. The metrics also take into account response of the human visual system to blur distortions.

The **Distortion Identification-based Image Verity and INtegrity Evaluation** (DIIVINE) metric[36] uses two-stage framework; first it does identification of the distortion, then it does quality assessment of each identified distortion. Evaluation of the metric[36] showed that it was statistically superior to the signal-to-noise ratio and that it was

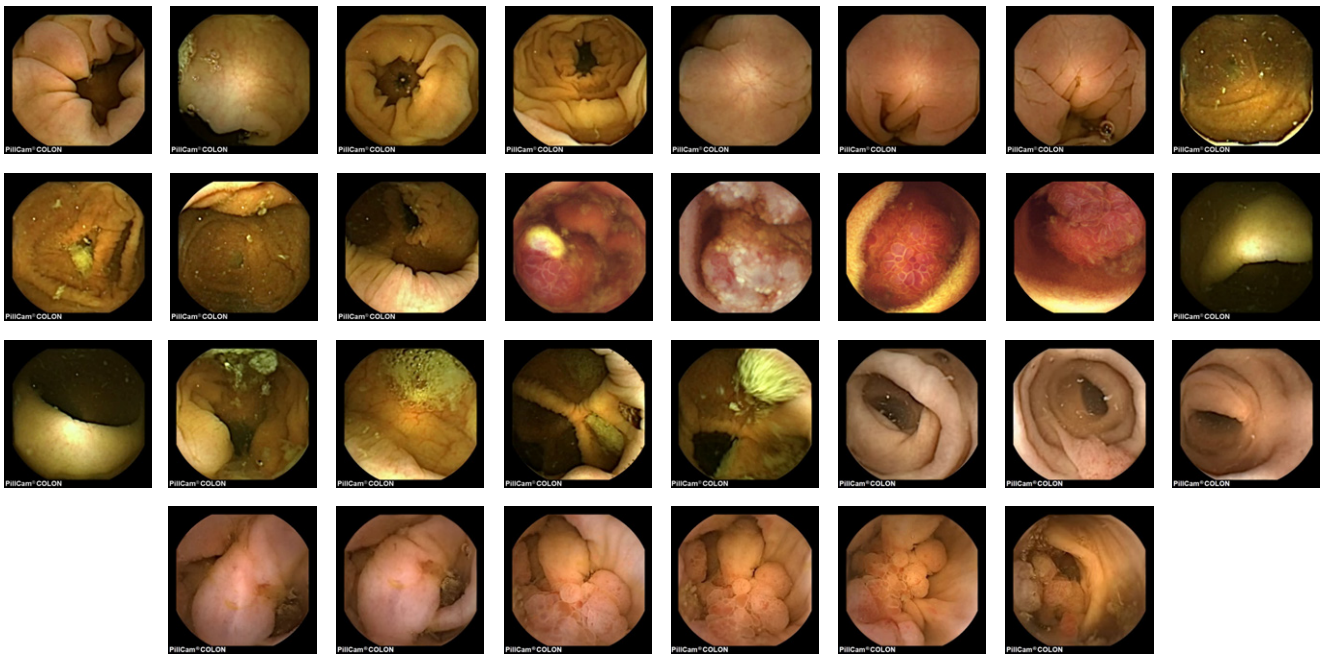


Figure 1. Images in the dataset.

statistically indistinguishable from the full-reference metric SSIM.[37]

The **Integrated Local Natural Image Quality Evaluator** (IL-NIQE)[38] is based on features from natural image statistics, which is used to learn a multivariate Gaussian model of image patches from a collection of pristine natural images. Through the learned multivariate Gaussian model, the quality of patches is calculated; the overall quality is the average of these patches. IL-NIQE is faster than DIIVINE and BLIINDS2, while having a better performance.

The **Weighted Level Framework** (WLF)[39] measures perceptual contrast in digital images by using a pyramidal approach where contrast is computed at different levels. The overall contrast is the weighted combination of the local contrast maps. WLF showed high correlation with perceived contrast, with a higher performance than other contrast measures.[39,40]

LAB variance (LABV)[41] is a contrast metric, which calculates the geometrical mean of the variance in each channel in the CIELAB colorspace. LABV showed similar performance to WLF, but it is computationally faster and less complex.[41] LABV has also shown to correlate well with perceived contrast in projection displays.[42,43]

EXPERIMENTAL SETUP FOR EVALUATION OF METRIC PERFORMANCE

In order to evaluate the performance of the image quality metrics we have used the dataset from Ahmed et al.[13] The 30 sample images (Figure 1) in the dataset were selected by a medical doctor from the KID dataset[44,45] and GivenImaging[46] capsule videos with pathologies and normal images from different parts of the colon. Raw images with original resolution are not available for testing due to proprietary issues. Image resolution reported here are as accessed from KID dataset[44,45] and Given Imaging.[46] 26 images were captured by Given Imaging Pillcam COLON capsules, out of which 17 images have resolution of 576×576 and the remaining images are of resolution 480×480. Moreover, four images are taken with Mirocam capsules with resolution 360×360. These 30 images were processed with three different enhancement techniques, giving 90 enhanced images, resulting in a total of 120 images.

- Bilateral:[47] enhances high-contrast images with high dynamic range while preserving details. For the image to be enhanced the method computes a multiscale decomposition based on the bilateral filter. The method combines detailed information at each scale for reconstructing the enhanced image.
- Weighted least square (WLS):[48] images enhanced with a multiscale edge-preserving smoothing method based on weighted least square image decomposition.
- Stochastic capsule endoscopy image enhancement Method:[13] enhances detail and shadow texture of tissues for capsule video endoscopy images. Based on stochastic sampling and edge-aware smoothing, the proposed method enhances vessels and tissue details for clinical applications. The framework decomposes the image into two detailed layers that contain shadowed tissue surfaces and detail features. Two concentric circles given by radii R_1 and R_2 are used for random walk and stochastic sampling, respectively. In order to smooth and contrast enhance the target pixel simultaneously, similar local neighborhood and local lightness and darkness pixels are explored. Similar local

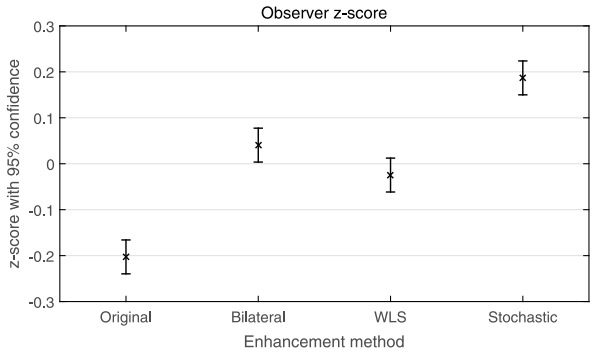


Figure 2. *Z*-scores from the psychometric experiment plotted with a 95% conference interval. A higher *z*-score indicate better diagnostic quality.

neighborhood pixels are explored through random walk within the inner circle which are used to smooth the image locally and get the base layer of the image, while local lightness and darkness pixels are explored in the outer circle resulting in a locally contrast enhanced image. The difference between the original image and base layer gives detail layers and similarly the difference between locally contrast enhanced image and base layer results in shadow texture layer. The final detail and shadow tissue texture enhanced image is obtained by convex linear combination of the two layers of detail and adding them back to the smoothed image.

The psychometric data is from Ahmed et al.[13] In their study five medical doctors who specialized in colonoscopy imaging participated in the experiment. They were asked to "Decide which image has the best diagnostic image quality." The observers ranked the images from best to worst. BENQ BL series UHD display, with screen resolution of 3840 × 2160 were used. The display is color managed for sRGB with luminance level of 80/cd m2. Moreover, to measure screen uniformity, a middle gray patch is used, and three points are sampled from left to right of the display. Benchmark result for screen uniformity shows a 3.8 standard deviation in CIE XYZ values. To measure color uniformity, different patches of red, green, and blue were measured along black and white patches with average CIE2000 value of 1.5686, which is below the just noticeable difference (JND). The images were displayed and positioned side by side in a random order. The rank-order data was further processed into *z*-scores (Figure 2). Each set of 30 images enhanced by the different methods including original images, have been evaluated by the five observers (medical doctors).

To assess the performance of the image quality metrics we use the rank-order method proposed by Pedersen and Hardeberg.[49] The rank for each image quality metric in the different images have been used as basis to generate *z*-scores, similar to what is common from psychometric rank-order experiments.[50] Further, we calculate the Pearson product-moment correlation coefficient and Spearman's rank correlation coefficient between the image quality metrics' *z*-scores and the observer's *z*-scores. In addition, to better analyze the performance for individual images we also investigate the correlation between the sum of scores (ranks) from the observers and the values of the image quality metrics. We have selected the following image quality metrics for the evaluation: JNBM,[23] BLIINDS2,[21] BRISQUE,[20] NIQMC,[24] SI,[25] AQI,[27] JPEG-S,[28] OG-IQA,[29] NRPWN,[30] DCTSP,[31] CPBD,[35], DIIVINE,[36] IL-NIQE,[38] WLF,[39] LABV,[41] BIQI,[19] and SSEQ.[22] These metrics have shown to correlate well with either perceived image quality or with perceived quality of relevant quality attributes, or they obtain properties that could be relevant for optimization. For image quality metrics that requires a grayscale input, we have converted the color images (*RGB*) to grayscale (*G*) using the following Equation:

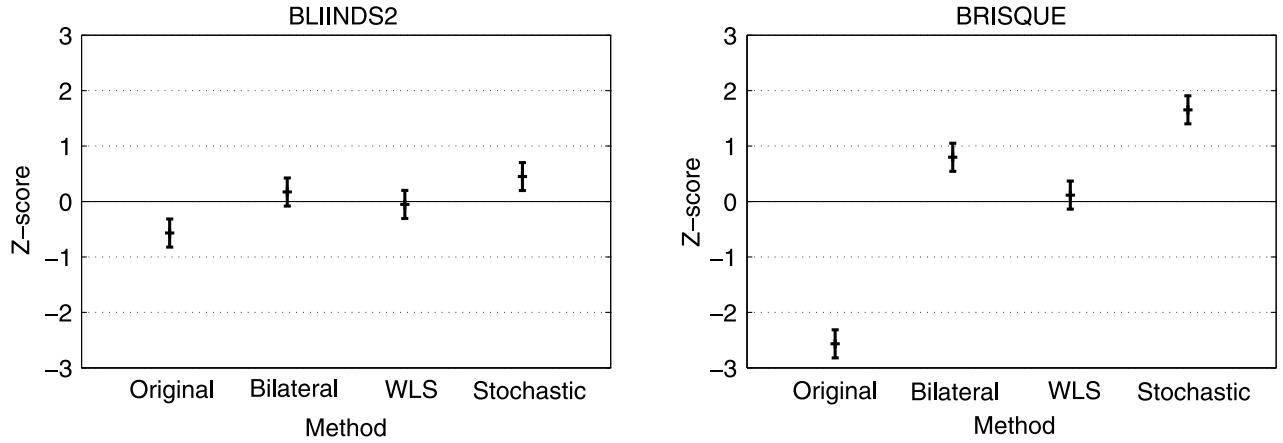
$$G = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B. \quad (1)$$

EXPERIMENT RESULTS AND DISCUSSION FOR METRIC PERFORMANCE

We have ranked all the images according to the quality value obtained after performing image quality assessment. Based on the ranks from 1 to 4 corresponding to the best quality and the worst, respectively, we have computed *z*-scores for each metric, as suggested in Ref. 49. We can see the results for the two metrics with the highest similarity to the observer *z*-scores in Figure 3. Both metrics have a different range compared to the observers, but they both have the same "shape" as the observers, with the stochastic enhancement having the highest *z*-score. Following, the correlation between the metric *z*-scores and the observer *z*-scores have been calculated. The results are shown in Figure 4.

From obtained results in Fig. 4, we can see that the highest Pearson correlation coefficients (greater than 0.8) are obtained by BLIINDS2, BRISQUE, CPBD, and DIIVINE. Investigation of the Spearman correlation identifies that BLIINDS2 and BRISQUE have a perfect rank order with a coefficient of 1, while AQI, CPBD, DIIVINE, OG-IQA, BIQI, and SSEQ have a coefficient of 0.8.

To further determine the metric with the best performance we also analyzed the results for each of the 30 images in the dataset, where the sum of scores (ranks) from the observers were compared to the quality values from the image quality metrics. We will show the results for the two best metrics (Figure 5). We can see that BRISQUE has a better correlation between the sum of scores from the observers and the metric values. BRISQUE has a linear Pearson correlation coefficient of 0.56, while BLIINDS2 has a coefficient of 0.27. For Spearman correlation, BRISQUE has a coefficient of 0.54 and BLIINDS2 has a coefficient of 0.27. It is important to notice that even if BRISQUE has the best performance, there are images where BRISQUE do not correlate completely with the observers. This is most likely due to the fact that BRISQUE is based on natural scene statistics, and that capsule endoscopic images might not be completely "natural." It is also interesting to notice that BRISQUE is only working in grayscale, since it is stated that it is important to keep the original color tone when



(a) BLIINDS2 (b) BRISQUE

Figure 3. *Z*-score plots for BLIINDS2 and BRISQUE according to the method proposed by Pedersen and Hardeberg.[49] Both metrics have a plot similar to that of the observers (Fig. 2).

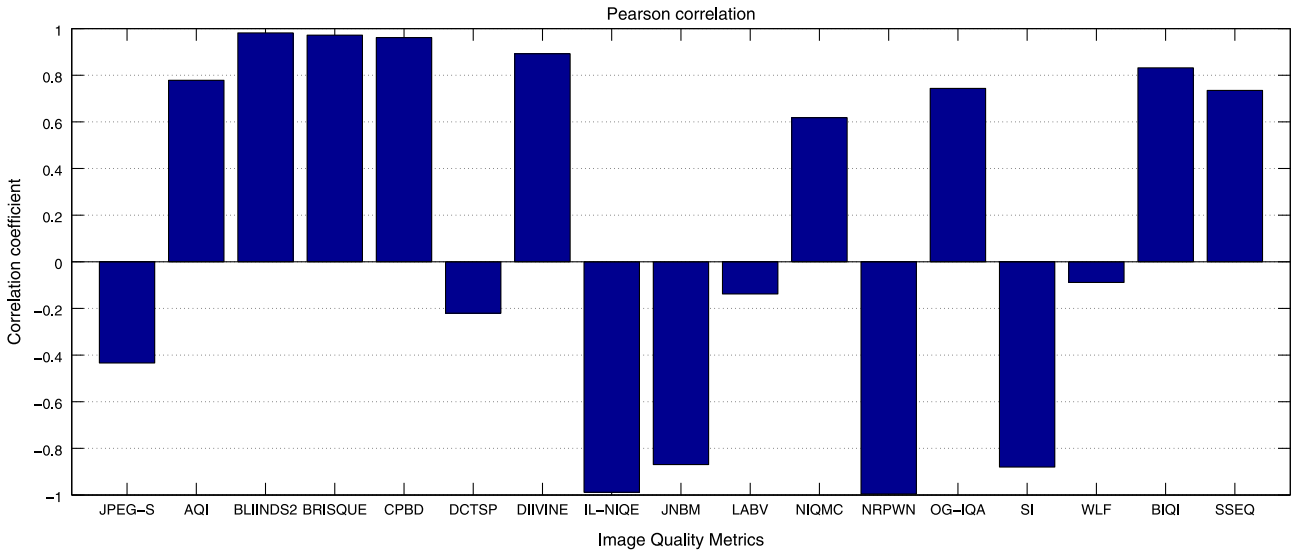


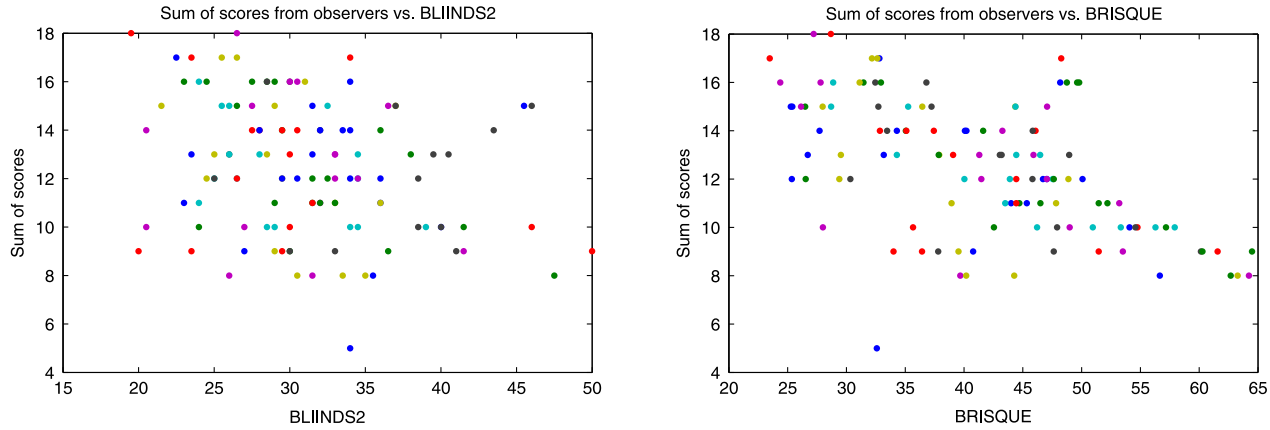
Figure 4. Pearson correlation between observer *z*-scores and metric Z-scores. The best performing metrics are BLIINDS2, BRISQUE, CPBD, DIIVINE and BIQI.

enhancing capsule images.[13] The explanation of this is that the enhancement methods are already incorporating the aspect of maintaining color tone, and because of this they normally do not alter the color and therefore BRISQUE performs well. Overall BRISQUE performs well, and because of its low computational complexity, it is well suited for real-time applications and for optimization. Based on this analysis, we will use BRISQUE in order to optimize the parameters from enhancement of capsule video endoscopy.

OPTIMIZATION OF PARAMETERS USING BRISQUE

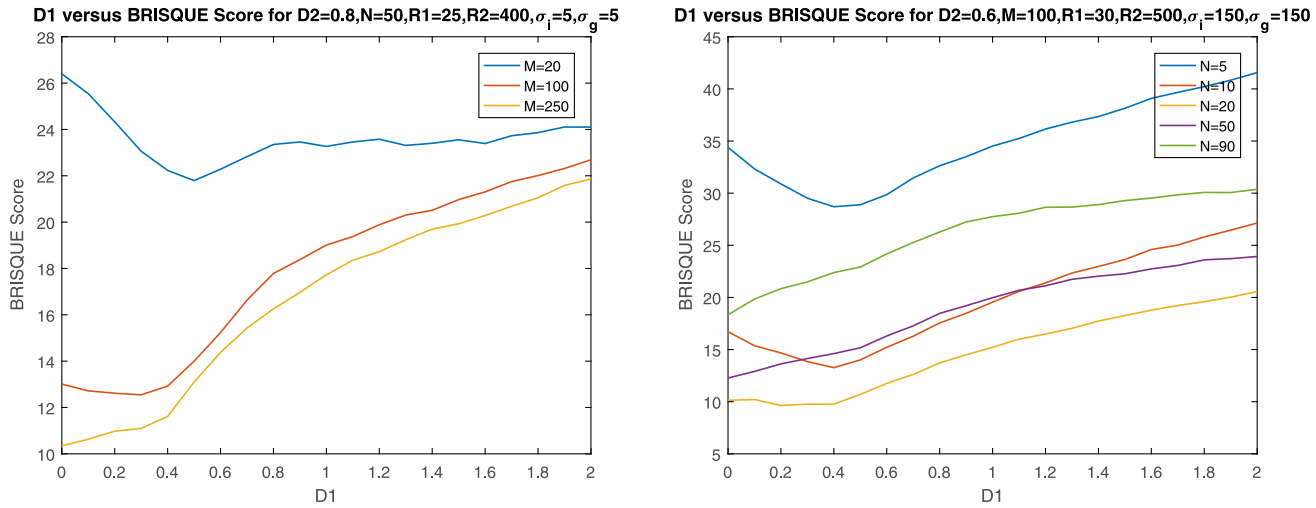
In order for a no-reference quality metric to be a practical tool in evaluating capsule image quality, we adopt a cross-validation technique by optimizing against the metric and validate through subjective experiment. Given the high performance of BRISQUE metric, we have used the metric to find optimal parameters for the best performing enhancement method in the study by Ahmed et al.[13] The objective of this section is to optimize the result of existing method with BRISQUE metric, which is better correlated with the observers.

In their method, Ahmed et al.[13] used two concentric circles given by radii R_1 and R_2 for random walk and stochastic sampling, respectively. Sampling method controls the way the algorithm enhances the contrast and details texture features. Large number of samples N and iteration M gives smoother and better result. The other parameters that are closely related to R_1 are σ_I and σ_g. Taking smaller values of σ_I gives lower weight to neighboring pixels that have significantly different intensity value from the target pixel, while σ_g weights neighboring pixel similarity based on the edges crossed during the random walk within the inner circle R_1. The final enhanced image is obtained by linear combination of detail layer D_1 and lightness layer

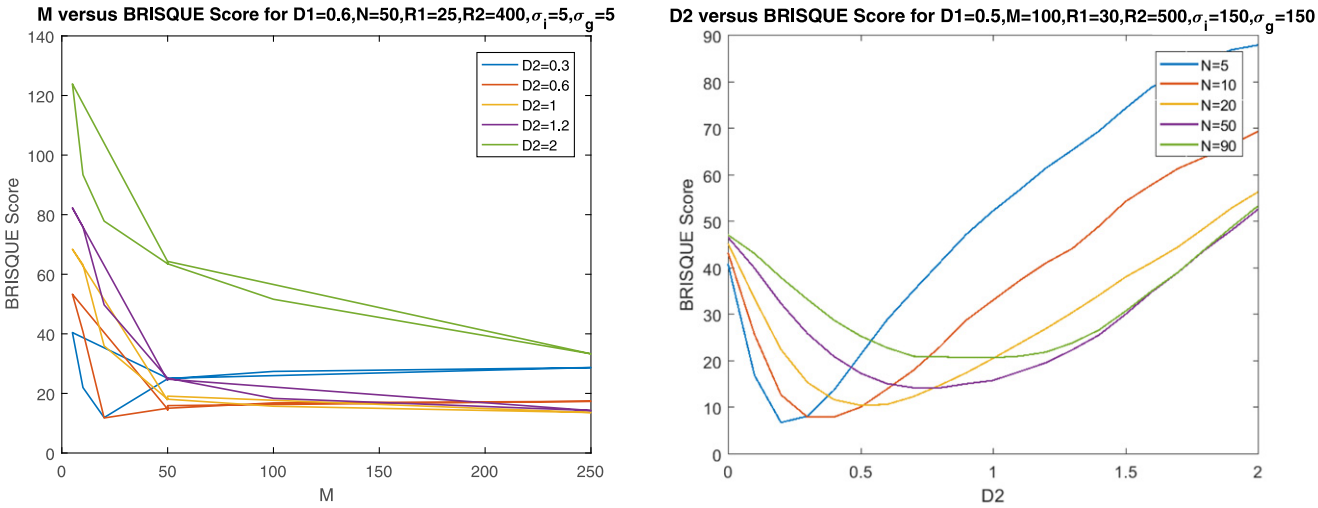


(a) BLIINDS2 (b) BRISQUE

Figure 5. Sum of scores (ranks) from the observers plotted against the quality values from BRISQUE and BLIINDS2. Each dot represents a different image. We can see that BRISQUE has a better correlation with the observers than BLIINDS2. Linear Pearson correlation for BRISQUE is 0.56 and BLIINDS2 is 0.27.

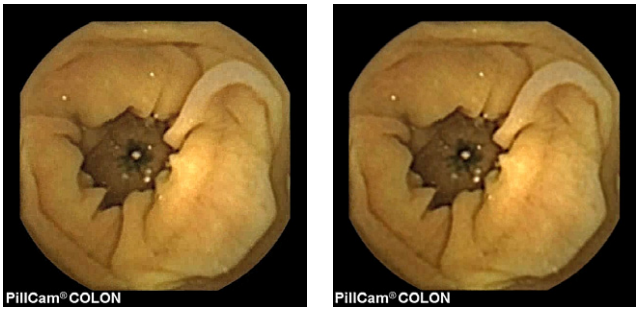


(a) D1 versus BRISQUE for different values of M (b) D1 versus BRISQUE for different values of N



(c) D2 versus BRISQUE for different values of M (d) D2 versus BRISQUE for different values of N

Figure 6. BRISQUE score against parameter space of Ref. 13.



(a) Unoptimized (b) Optimized

Figure 7. Comparison of unoptimized and optimized images.

D_2. For details of the method, we refer the reader to Ref. 13. Due to space limitation, we focus our discussion on parameters that affect the BRISQUE score significantly. In order to optimize using BRISQUE, we run the method for $D_1 = [0, 2]$ and $D_2 = [0, 2]$ with step size of 0.1, $N = [5, 10, 20, 50, 60, 90]$, $M = [5, 10, 20, 50, 100, 250]$, $R_1 = [1, 5, 10, 25, 30, 45]$, $R_2 = [5, 10, 100, 400, 500, 800]$, σ_I and σ_g for [0.5, 1, 2, 5, 150, 10000] values, respectively. The parameter space is chosen as described in Ref. 13. The results are summarized in Figure 6.

As is shown in Fig. 6 optimal BRISQUE score is obtained for number of samples, $M = 250$. Similarly, $N = 10$, number of iteration is required with detail layer and lightness layer mixing multiplier $D_1 = 0.5$ and $D_2 = 0.3$, respectively. Random walk inner circle radius of $R_1 = 10$, with stochastic sampling outer circle of radius $R_2 = 800$ with $\sigma_I = 10$ and $\sigma_g = 10$, gives optimal BRISQUE score.

In order to evaluate the performance of the optimized images, we have conducted a psychometric experiment with a medical doctor using QuickEval.[51] The 30 images shown in Fig. 1 processed with the standard parameters from Ahmed et al.[13] and with the optimized parameters suggested by BRISQUE. A paired comparison experiment was conducted in a dim room, with a viewing distance of approximately 50 cm. The doctor was instructed to select the image with the highest diagnostic quality. In 20 of the 30 images the doctor preferred the optimized image, and in 10 images the unoptimized. Binomial sign test shows statistically significant difference to null hypothesis with 95% confidence interval of [0.472, 0.827]. The result shows that values from the BRISQUE metric is highly correlated with experts rating of capsule images.

Figure 7 shows the unoptimized image and the optimized image using BRISQUE. We can notice that the unoptimized image has more noticeable blocking artifacts. This can also be seen in many of the other test images, especially those with larger areas of the same color.

CONCLUSION

We have evaluated state-of-the-art no-reference image quality metrics on a dataset from capsule video endoscopes. The Blind/Referenceless Image Spatial QUality Evaluator (BRISQUE) has the highest performance with regard to human observers. BRISQUE has further been used in order to optimize one of the recent enhancement methods for capsule video endoscopy. By using BRISQUE the best parameters for the enhancement method was selected. Based on a small psychometric experiment with a medical doctor we show that through optimization using BRISQUE images with higher diagnostic quality can be generated.

ACKNOWLEDGMENT

The authors would like to thank Dr. Øistein Hovde for his assistance and feedback.

This research has been supported by the Research Council of Norway through project no. 247689 IQ-MED: Image Quality enhancement in MEDical diagnosis, monitoring and treatment.

The authors would like to acknowledge the Color in Science and Industry master program (Jean Monnet University, Norwegian University of Science and Technology, University of Granada, and University of Eastern Finland).

REFERENCES

[1] C. Spada, C. Hassan, J. P. Galmiche, H. Neuhaus, J. M. Dumonceau, S. Adler, O. Epstein, G. Gay, M. Pennazio, D. K. Rex, R. Benamouzig, R. De Franchis, M. Delvaux, J. Devire, R. Eliakim, C. Fraser, F. Hagenmuller, J. M. Herrerias, M. Keuchel, F. MacRae, M. Munoz-Navas, T. Ponchon, E. Quintero, M. E. Riccioni, E. Rondonotti, R. Marmo, J. J. Sung, H. Tajiri, E. Toth, K. Triantafyllou, A. Van Gossum, and G. Costamagna, "Colon capsule endoscopy: European society of gastrointestinal endoscopy (esge) guideline," Endoscopy **44**, 527–536 (2012).

[2] A. Kornbluth, P. Legnani, and B. S. Lewis, "Video capsule endoscopy in inflammatory bowel disease. Past, present, and future," Inflammatory Bowel Diseases **10**, 278–285 (2004).

[3] C. Spada, C. Hassan, M. Munoz-Navas, H. Neuhaus, J. Deviere, P. Fockens, E. Coron, G. Gay, E. Toth, M. E. Riccioni, C. Carretero, J. P. Charton, A. Van Gossum, C. A. Wientjes, S. Sacher-Huvelin, M. Delvaux, A. Nemeth, L. Petruzziello, C. P. De Frias, R. Mayershofer, L. Aminejab, E. Dekker, J. P. Galmiche, M. Frederic, G. W. Johansson, P. Cesaro, and G. Costamagna, "Second-generation colon capsule endoscopy compared with colonoscopy," Gastrointestinal Endoscopy **74**, 9 (2011).

[4] I. Fernandez-Urien, C. Carretero, A. Borda, and M. Muñoz-Navas, "Colon capsule endoscopy," World J. Gastroenterol **14**, 5265–5268 (2008).

[5] E. Sakai, H. Endo, S. Kato, T. Matsuura, W. Tomeno, L. Taniguchi, T. Uchiyama, Y. Hata, E. Yamada, H. Ohkubo, T. Higrashi, K. Hosono, H. Takahashi, and A. Nakajima, "Capsule endoscopy with flexible spectral imaging color enhancement reduces the bile pigment effect and improves the detectability of small bowel lesions," BMC Gastroenterology **12**, 83 (2012).

[6] R. Lambert, K. Kuznetsov, and J.-F. Rey, "Narrow-band imaging in digestive endoscopy," Scientific World J. **7**, 449–465 (2007).

[7] M. Ramaraj, S. Raghavan, and W. A. Khan, "Homomorphic filtering techniques for wce image enhancement," *IEEE Int'l. Conf. on Computational Intelligence and Computing Research* (IEEE, Piscataway, NJ, 2013), pp. 1–5.

[8] H. Ibrahim and N. S. Pik Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," IEEE Trans. Consum. Electron. **53**, 1752–1758 (2007).

[9] H. Okuhata, H. Nakamura, S. Hara, H. Tsutsui, and T. Onoye, "Application of the real-time retinex image enhancement for endoscopic images," *35th Annual Int'l. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)* (IEEE, Piscataway, NJ, 2013), pp. 3407–3410.

[10] V. P Gopi, P. Palanisamy, and S. Issac Niwas, "Capsule endoscopic colour image denoising using complex wavelet transform," *Wireless Networks and Computational Intelligence* (Springer, Berlin, Heidelberg, 2012), pp. 220–229.

[11] L. Haiying, W.-S. Lu, and M. Q.-H. Meng, "De-blurring wireless capsule endoscopy images by total variation minimization," *Communications, Computers and Signal Processing (PacRim), 2011 IEEE Pacific Rim Conf.* (IEEE, Piscataway, NJ, 2011), pp. 102–106.

[12] B. Li and M. Q.-H. Meng, "Wireless capsule endoscopy images enhancement via adaptive contrast diffusion," J. Vis. Commun. Image Represent. **23**, 222–228 (2012).

[13] M. Ahmed, I. Farup, M. Pedersen, Ø. Hovde, and S. Yildirim, Stochastic capsule endoscopy image enhancement. Submitted (2016).

[14] V. Hai, T. Echigo, K. Yagi, H. Okazaki, Y. Fujiwara, Y. Yagi, and T. Arakawa, "Image-enhanced capsule endoscopy preserving the original color tones," *Int'l. MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging* (Springer, Berlin, Heidelberg, 2011), pp. 35–43.

[15] M. S. Imtiaz and K. A. Wahid, "Color enhancement in endoscopic images using adaptive sigmoid function and space variant color reproduction," Comput. Math. Methods Med. **2015**, 3905–3908 (2015).

[16] M. Pedersen and J. Y. Hardeberg, "Full-reference image quality metrics: Classification and evaluation," Found. Trends. Comput. Graph. Vis. **7**, 1–80 (2012).

[17] V. Kamble and K. M. Bhurchandi, "No-reference image quality assessment algorithms: A survey," Optik - Int. J. Light Electron Optics **126**, 1090–1097 (2015).

[18] M. Shahid, A. Rossholm, B. Lövström, and H. J. Zepernick, "No-reference image and video quality assessment: a classification and review of recent approaches," EURASIP J. Image Video Process. **40** (2014).

[19] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," IEEE Signal Process. Lett. **17**, 513–516 (2010).

[20] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Trans. Image Process. **21**, 4695–4708 (2012).

[21] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," IEEE Trans. Image Process. **21**, 3339–3352 (2012).

[22] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," Signal Process., Image Commun. **29**, 856–863 (2014).

[23] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb)," IEEE Trans. on Image Process. **18**, 717–728 (2009).

[24] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "No-reference quality metric of contrast-distorted images based on information maximization," IEEE Trans. Cybernetics (2016).

[25] G. Blanchet and L. Moisan, "An explicit sharpness index related to global phase coherence," *IEEE Int'l. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, Piscataway, NJ, 2012), pp. 1065–1068.

[26] G. Blanchet, L. Moisan, and B. Rougé, "Measuring the global phase coherence of an image," *Image Processing, 2008. ICIP 2008. 15th IEEE Int'l. Conf.* (IEEE, Piscataway, NJ, 2008), pp. 1176–1179.

[27] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy," JOSA A **24**, B42–B51 (2007).

[28] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference perceptual quality assessment of jpeg compressed images," *Image Processing. 2002. Proc. 2002 Int'l. Conf.* (IEEE, Piscataway, NJ, 2002), Vol. 1, pp. I–477.

[29] L. Liu, Y. Hua, Q. Zhao, H. Huang, and A. C. Bovik, "Blind image quality assessment by relative gradient statistics and adaboosting neural network," Signal Process., Image Commun. **40**, 1–15 (2016).

[30] T. Zhu and L. Karam, "A no-reference objective image quality metric based on perceptually weighted local noise," EURASIP J. Image and Video Process. **2014**, 1–8 (2014).

[31] Y. Han, X. Xu, and Y. Cai, "Novel no-reference image blur metric based on block-based discrete cosine transform statistics," Optical Eng. **49**, 050501 (2010).

[32] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, Live image quality assessment database release 2 (2005).

[33] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," J. Electron. Imaging **19**, 011006 (2010).

[34] N. Ponomarenko, F. Battisti, K. Egiazarian, J. Astola, and V. Lukin, "Metrics performance comparison for color image database," *Fourth Int'l. Workshop on Video Processing and Quality Metrics for Consumer Electronics, Scottsdale, Arizona, USA* (2009), Vol. 27, p. 6.

[35] N. D. Narvekar and L. J. Karam, "A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection," *Quality of Multimedia Experience, 2009. QoMEx 2009. Int'l. Workshop* (IEEE, Piscataway, NJ, 2009), pp. 87–91.

[36] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," IEEE Trans. Image Process. **20**, 3350–3364 (2011).

[37] 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. **13**, 600–612 (2004).

[38] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," IEEE Trans. Image Process. **24**, 2579–2591 (2015).

[39] G. Simone, M. Pedersen, and J. Y. Hardeberg, "Measuring perceptual contrast in digital images," J. Vis. Commun. Image Represent. **23**, 491–506 (2012).

[40] G. Simone, M. Pedersen, and J. Y. Hardeberg, "Measuring perceptual contrast in uncontrolled environments," *2010 2nd European Workshop on Visual Information Processing (EUVIP)* (IEEE, Piscataway, NJ, 2010), pp. 102–107.

[41] M. Pedersen, A. Rizzi, J. Y. Hardeberg, and G. Simone, "Evaluation of contrast measures in relation to observers perceived contrast," *Proc. IS&T CGIV2008: 4th European Conf. on Colour in Graphics, Imaging, and Vision* (IS&T, Springfield, VA, 2008), pp. 253–258.

[42] Z. Ping, M. Pedersen, J. Y. Hardeberg, and J.-B. Thomas, "Camera-based measurement of relative image contrast in projection displays," *Visual Information Processing (EUVIP), 2013 4th European Workshop* (IEEE, Piscataway, NJ, 2013), pp. 112–117.

[43] P. Zhao, M. Pedersen, J. Y. Hardeberg, and J.-B. Thomas, "Measuring the relative image contrast of projection displays," J. Imaging Sci. Technol. **59**, 30404 (2015).

[44] A. Koulaouzidis and D. K. Iakovidis, Kid: Koulaouzidis-iakovidis database for capsule endoscopy, (2015).

[45] D. K. Iakovidis and A. Koulaouzidis, "Software for enhanced video capsule endoscopy: challenges for essential progress," Nature Reviews Gastroenterology Hepatology **12**, 172–186 (2015).

[46] GivenImaging. Capsule video endoscopy, image atlas. http://www.capsuleendoscopy.org, (2016).

[47] R. Fattal, M. Agrawala, and S. Rusinkiewicz, "Multiscale shape and detail enhancement from multi-light image collections," ACM Trans. Graph. (Proc. SIGGRAPH) **26** (2007).

[48] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," ACM Trans. Graph. **27**, 67 (2008).

[49] M. Pedersen and J. Y. Hardeberg, "Rank order and image difference metrics," *Proc. IS&T CGIV2008: 4th European Conf. on Colour in Graphics, Imaging, and Vision* (IS&T, Springfield, VA, 2008), pp. 120–125.

[50] P. G. Engeldrum, *Psychometric Scaling: A Toolkit for Imaging Systems Development* (Imcotek Press, Winchester, MA, 2000).

[51] K. V. Ngo, J. J. Storvik, C. A. Dokkeberg, I. Farup, and M. Pedersen, "Quickeval: a web application for psychometric scaling experiments," Proc. SPIE **9396**, 93960O (2015).