

PAPER • OPEN ACCESS

Image quality assessment (IQA) using high-frequency and image variance (HFIV) for colour image

To cite this article: Li Chien Tan *et al* 2019 *J. Phys.: Conf. Ser.* **1372** 012034

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Image quality assessment (IQA) using high-frequency and image variance (HFIV) for colour image

Li Chien Tan^{1*}, Haniza Yazid² and Yen Fook Chong³

School of Mechatronic Engineering, Universiti Malaysia Perlis, Pauh Putra Campus, 02600, Arau, Perlis, Malaysia.

*LCtan94@gmail.com

Abstract. Image quality is often lost during image acquisition, transmission, and compression. Therefore, image quality assessment (IQA) is crucial in image processing. Currently, image quality can be measured from the frequency domain features, but it only applicable to blurred grayscale images. Nevertheless, noise distortion is also a common problem in digital images, and colour also affects the perception of image quality. Therefore, this paper proposes an enhanced blur and noise specific colour image quality assessment that measures high-frequency components and image variance. The number of high-frequency components is related to the edge and noise. In order to distinguish the distortion of the image, the image variance estimation is included. Experiments on public databases have shown that this method outperforms PSNR and SSIM in terms of noise and blur distortion and has low processing time of 0.0941 s/img.

1. Introduction

In most electronic imaging applications, high-quality images are requested and often required. However, the visual quality of the image may be corrupted during the process of image acquisition, transmission and so on. The process of evaluating the quality of a distorted image with respect to the original image is referred as IQA. IQA is important because if a poor medical image is not identified immediately, misdiagnosis and delayed treatment may occur, which may result in death. On the other hand, when the same situation occurs in the industrial sector, misdetection will result business and reputation losses to the company. Therefore, there is a great need for IQA to measure the image quality.

In general, IQA can be divided into two categories, subjective and objective method. The subjective method evaluates image quality by human observers. Human is the final receivers in most of image processing application [1] and we can easily evaluate images without referring any reference, hence, the subjective method is the most reliable [2]. The popular subjective quality evaluators, such as Mean Opinion Score (MOS) and Difference Mean Opinion Score (DMOS) have been used in various image processing applications. However, these methods always rely on the visual capability of the viewer and the luminance of the environment. Thus, it is inconvenient, time consuming and costly. Also, it is usually too slow to be used in real-world applications.

In consequence, objective IQA methods are mainly used in most applications. They can usually be subdivided into three types: Full-Reference (FR), such as Mean-Squared Error (MSE) and Structural Similarity Index (SSIM), which require a complete reference image and measure the similarity between the test and reference image; Reduced-Reference (RR), only part of the reference image is available; No-Reference (NR), no reference picture is required. Both FR and RR require a reference



image. However, in most cases, there is no reference image. Therefore, NR can meet people's requirements at this time.

NR methods are typically designed for specific distortion. In this paper, our work focuses on blur and noise estimation in images, which is the most common type of distortion in digital images. Therefore, let us briefly summarize the relevant existing literature in this regard. The sharpness of the image can be measured by computing the wavelet decomposition of the images [3], measuring the edge spread in the image [4] and also the high-frequency (HF) components in the frequency domain [5]. Both methods assume uniform blur.

Besides, some methods focus on both noise and blur distortion, for example, the metrics proposed by Zhu and Milanfar [6], H metric, can detect blur and noise based upon the singular value decomposition (SVD) of the local image gradient matrix based on the assumption that the noise variance is known or estimable. In 2010, they modified the method by introducing Q metric [7,8]. This method is easy to compute, has a reasonable response to both blur and random noise, but the long processing time is the weak point. In addition, Cohen and Yitzhaky [9] estimated the noise in both spatial and frequency domain while blur is only estimated in the frequency domain. This method worked well with additive white noise, Gaussian blur and de-focus blur. But, the performance of the method may be degraded when the image contains strong periodic patterns and when the blurred image contains high noise power. However, all of the methods mentioned above were only suitable to evaluate grayscale images. Not only that, the accuracy, complexity and processing time of the existing metrics for blur and noise images always directly proportional. Therefore, they are not suitable for integration in real-time systems.

Inspired by [5], we propose an improved frequency domain based colour NR-IQA algorithm for blur and noise images. First, the ratio of the HF components in each colour component (YCbCr) is measured separately and then combined by the weight of the rod and the cone receptor in the human retina. Second, measure the image variance of each set of images. If the minimum image variance of each set of images is less than a predetermined threshold, then that particular set of images is considered blurry, otherwise is noisy. The experiments showed that our proposed method was accurate, effective, well-correlated with the human visual system and has short processing time. The rest of the paper is organized as follows: In Section 2, the methodology of the proposed method is discussed, and in Section 3, we present experimental results and correlations with subjective IQA. Finally, in Section 4, the entire paper will be finalized and possible future work is discussed.

2. Proposed colour NR-IQA

Our proposed method is the hybrid of *HF measurement and image variance estimation (HFIV)*, mainly composed of three stages. First, the algorithms described in [5] was implemented in the YCbCr colour space, followed by image variance estimation and lastly the final image quality calculation. The block diagram of the proposed algorithm is shown in Figure 1.

2.1. High-Frequency Measurement in YCbCr colour space

Figure 2 represents the overall flow of the first stage. First, split the colour image into Y, Cb and Cr colour components, and then compute their respective Fourier Transform representation individually. Next, calculate the threshold and measure the ratio of HF components in each colour components. Lastly, combine all together for final image quality score calculation. In this paper, the YCbCr colour space is chosen because it always used as a part of the colour image pipeline in video and digital photography systems, therefore is more suitable for IQA [10]. Y represents the luminance, while Cb and Cr are the blue-difference and red-difference chroma components.

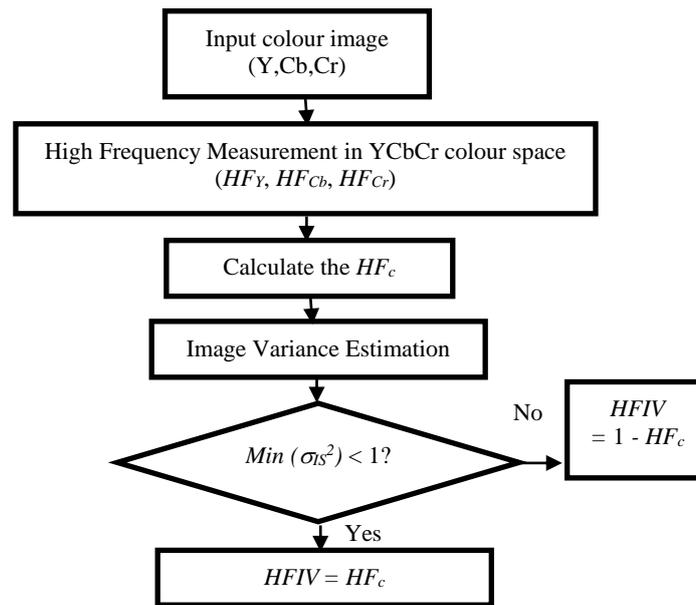


Figure 1. Block diagram of the proposed *HFIV* algorithm.

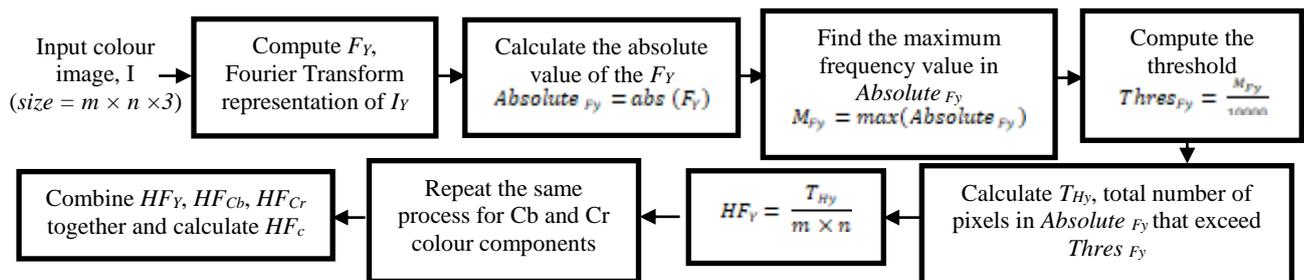


Figure 2. Block diagram of the Stage I (High-Frequency Measurement in YCbCr colour space).

Human retina has two types of photoreceptors, rods and cones. The rod is responsible for scotopic vision, very sensitive to light, and does not mediate colour perception, while on the other hand, cones provide colour perception and the spatial perception of the eye. The fovea part of the human eye has approximately 7 million cones and 120 million rods [11,12]. Therefore, the weight of the luminance component (L_w) and the weight of the colour components (C_w) are calculated as follows, Eq. (1) and Eq. (2). Then, the number of HF component of the colour image (HF_c) can be easily obtained using Eq. (3) [13].

$$L_w = \frac{120,000,000}{(120,000,000 + 7,000,000)} = 0.9449 \quad (1)$$

$$C_w = \frac{7,000,000}{(120,000,000 + 7,000,000)} = 0.0551 \quad (2)$$

$$HF_c = 0.9449 (HF_Y) + 0.0551 \left(\frac{HF_{Cb} + HF_{Cr}}{2} \right) \quad (3)$$

As we know that, as the blur in the image increases, the number of HF components in the image is suppressed. A clearer, the higher quality image will have a greater number of HF components than a blurred image. However, on the contrary, noise is the high frequency in nature. The more the image is contaminated by noise, the greater the number of HF components. Then the problem arises. The image

with the highest number of HF components can be the sharpest image or the noisiest image. In order to solve this problem, the image variance estimation phase is added.

2.2. Image Variance Estimation

In this stage, this paper implements the fast noise variance estimation algorithm proposed by John Immerkaer [14]. The reason for choosing this method is that it is very simple. Grayscale images, rather than colour images, are used at this stage to reduce the processing time. Apply a noise estimation operator N that is almost insensitive to the image structure to the input grayscale image, I_G .

$$N = \begin{array}{|c|c|c|} \hline 1 & -2 & 1 \\ \hline -2 & 4 & -2 \\ \hline 1 & -2 & 1 \\ \hline \end{array} \quad (4)$$

Next, compute the variance of each output image, $\sigma_{g,n}^2$ and find the minimum variance of each group of image, σ_{kmin}^2 using Eq. (5) and (6), respectively.

$$\sigma_{g,n}^2 = \left(\frac{\pi}{2} \frac{1}{6(W-2)(H-2)} \sum_I |I_G(x,y) * N| \right)^2 \quad (5)$$

$$\sigma_{k,min}^2 = \min(\sigma_{k,1}^2, \sigma_{k,2}^2, \dots, \sigma_{k,n}^2) \quad , k = 1, 2, \dots, g \quad (6)$$

where g is the number of image groups, n is the number of image per group, W, H represent the width and height of the input image, $I_G(x,y) * N$ is the output image, and $*$ represents convolution.

When the blurred image's edge is smoothed, the σ_{kmin}^2 is very small, usually below 1. In contrast, when the noise contained in the image increases, the σ_{kmin}^2 will be increased, generally more than 1. However, this method does not apply if there is only one image. The details of the method can be referred in [14].

2.3. Image Quality Measure (IQM)

Noise and blur have a paradoxical effect on image attributes. In the frequency domain, noise affects the entire frequency band and primarily increases the higher frequencies, while blurring primarily suppresses higher frequencies. For this reason, the image quality measurement of noise and blurred images is different.

For a blurry image, the higher the number of HF components, the better the image quality. For noisy images, the HF_c value is inversely proportional to the image quality. Therefore, the final $HFIV$ of the blurred image and noisy image can be expressed as Eq. (7), respectively.

$$HFIV = \begin{cases} 1 - HF_c & ; \text{if } \sigma_{kmin}^2 \geq 1 \\ HF_c & ; \text{if } \sigma_{kmin}^2 < 1 \end{cases} \quad (7)$$

3. Experimental Results

In order to evaluate the performance of the proposed IQM, our experiments used three subjective IQA databases, CSIQ [15], TID2008 [16] and TID2013 [17]. In this paper, only a few types of distortion are included. In the CSIQ database, it includes additive white Gaussian noise (AWGN), Gaussian blur (Blur) and additive pink Gaussian noise (Fnoise). For the TID2008 and TID2013 databases, the types of distortion include additive Gaussian noise, additive noise in colour components, spatially correlated noise, masked noise, HF noise, impulse noise, quantization noise, Gaussian blur, and image denoising. Each of the distorted images in TID2008, TID2013, and CSIQ include a corresponding MOS and a DMOS, respectively, which can be used as a ground truth. Table 1 summarizes the database information used in our experiments.

To illustrate the performance of IQA metrics, two evaluation criteria are used, the Pearson linear ($PLCC$) and the Spearman's rank order correlation coefficients ($SRCC$). High $PLCC$ scores and $SRCC$ scores are associated with high accuracy, monotonicity, and consistency of the metrics in the test. The closer the value of $SRCC$ and $PLCC$ to 1, the better the metric performance.

Table 1. Database with selected distortion for performance evaluation

| Database | Reference Image | Selected distortion types | Distortion level | Total distorted image used | Subjective score |
|----------|-----------------|---------------------------|------------------|----------------------------|------------------|
| CSIQ | 30 | 3 | 5 | 450 | DMOS |
| TID2008 | 25 | 9 | 4 | 900 | MOS |
| TID2013 | 25 | 9 | 5 | 1125 | MOS |

The performance of this method is also compared to PSNR and SSIM. The evaluation metrics *SRCC* and *PLCC* are then calculated to present their accuracy and consistency. Tables 2, 3 and 4 show the results of PSNR, SSIM and proposed methods in the CSIQ, TID2008 and TID2013 databases, respectively. A total of 2475 images were used in this evaluation.

As can be seen from Table 2, in CSIQ database, our proposed method can compete with the FR method. Even though the result of our algorithm is slightly lower than PSNR and SSIM, but, the difference between *SRCC* and *PLCC* is less than 0.04. Our result shows a strong negative *SRCC* of 0.9978 and a strong negative *PLCC* of 0.9465 between our calculated score and the DMOS. The relationship between our objective score and DMOS is a strong negative relationship, which means high prediction accuracy and monotonically.

Table 2. Average *SRCC* and *PLCC* with DMOS for different types of distortion in the CSIQ database.

| | <i>SRCC</i> | | | <i>PLCC</i> | | |
|---------|-------------|------|-----------------|-------------|---------|-----------------|
| | PSNR | SSIM | Proposed method | PSNR | SSIM | Proposed method |
| Average | -1 | -1 | -0.9978 | -0.9828 | -0.9689 | -0.9465 |

From Table 3 and Table 4, it is clear that our proposed method significantly outperforms PSNR and SSIM in the TID2008 and TID2013 databases. Our approach achieves a very strong positive *SRCC* and *PLCC* correlation between subjective ratings MOS, which are very close to 1. This demonstrates that our algorithm can achieve stable and good performance across different noise and blur distortions for the entire three databases.

Table 3. *SRCC* and *PLCC* with MOS for different types of distortion in the TID2008 database.

| | <i>SRCC</i> | | | <i>PLCC</i> | | |
|---------|-------------|--------|-----------------|-------------|--------|-----------------|
| | PSNR | SSIM | Proposed method | PSNR | SSIM | Proposed method |
| Average | 0.8404 | 0.8404 | 0.9758 | 0.8351 | 0.8352 | 0.9668 |

Table 4. *SRCC* and *PLCC* with MOS for different types of distortion in the TID2013 database.

| | <i>SRCC</i> | | | <i>PLCC</i> | | |
|---------|-------------|--------|-----------------|-------------|--------|-----------------|
| | PSNR | SSIM | Proposed method | PSNR | SSIM | Proposed method |
| Average | 0.8680 | 0.8680 | 0.9675 | 0.8656 | 0.8661 | 0.9634 |

In addition, the average execution time of each IQA algorithms was measured to assess the complexity. The operating system used in this experimental environment was Windows 10 64-bit operating system with an i7-6500U CPU @ 2.60GHz CPU and 8.00GB RAM. Table 5 summarizes all simulation results. It shows that PSNR has the fastest execution speed, followed by our proposed method and SSIM. Although PSNR has the fastest speed, it requires a reference image that is not always available. Therefore, our method is still the best way to deal with noise and blurred images. It can process images in 0.1s without reference, so it is suitable for real-time applications.

Table 5. Average execution time (s / img) of PSNR, SSIM and proposed method.

| Method | PSNR | SSIM | Proposed method |
|---------|--------|--------|-----------------|
| Average | 0.0328 | 0.3811 | 0.0941 |

4. Conclusion

This paper introduces a hybrid of *HF component measurement and image variance estimation (HFIV)* for NR-IQA. This method is dedicated to blur and noise distortion, which is the most common type of distortion in digital imaging applications. In addition, we compared our approach to the performance of PSNR and SSIM in the CSIQ, TID2008 and TID2013 databases. Experiments showed that our proposed method has strong *SRCC* and *PLCC* correlation with human visual perception, able to accurately evaluate the image without any reference. In addition, the experiment also demonstrated that our proposed algorithm has a short processing time of 0.0941s/img, so, it can meet the real-time requirement of most current applications. Blocking artifacts caused by compression are also very common in digital imaging. Therefore, for future work, this method can be improved to accommodate JPEG and JPEG2000 distorted images.

References

- [1] Yi Y, Yu X, Wang L and Yang Z 2008 Image Quality Assessment Based on Structural Distortion and Image Definition 2008 *International Conference on Computer Science and Software Engineering* (IEEE) pp 253–6
- [2] Choi M G, Jung J H and Jeon J W 2009 No-Reference Image Quality Assessment using Blur and Noise *Int. J. Electr. Comput. Energ. Electron. Commun. Eng.* **3** 184–8
- [3] Abdel-Hamid L, El-Rafei A and Michelson G 2017 No-reference quality index for color retinal images *Comput. Biol. Med.* **90** 68–75
- [4] Feichtenhofer C 2003 No-reference Sharpness Metric Based on Local Gradient Analysis *Proceedings International Conference on Image Processing* (IEEE Xplore) pp 1–34
- [5] De K and Masilamani V 2013 Image Sharpness Measure for Blurred Images in Frequency Domain *Procedia Eng.* **64** 149–58
- [6] Zhu X and Milanfar P 2009 A no-reference sharpness metric sensitive to blur and noise 2009 *International Workshop on Quality of Multimedia Experience* (IEEE) pp 64–9
- [7] Zhu X and Milanfar P 2010 Automatic Parameter Selection for Denoising Algorithms Using a No-Reference Measure of Image Content *IEEE Trans. Image Process.* **19** 3116–32
- [8] Prajapati P, Narmawala Z, Darji N P, Moorthi S M and Ramakrishnan R 2015 Evaluation of Perceptual Contrast and Sharpness Measures for Meteorological Satellite Images *Procedia Comput. Sci.* **57** 17–24
- [9] Cohen E and Yitzhaky Y 2010 No-reference assessment of blur and noise impacts on image quality *Signal, Image Video Process.* **4** 289–302
- [10] Wang Q, Chu J, Xu L and Chen Q 2016 A new blind image quality framework based on natural color statistic *Neurocomputing* **173** 1798–810
- [11] Gonzalez R C and Woods R E 2009 *Digital Image Processing (3rd Edition)* vol 56
- [12] Yalman Y and Ertürk I 2013 A new color image quality measure based on YUV transformation and PSNR for human vision system *Turkish J. Electr. Eng. Comput. Sci.* **21** 603–12
- [13] Yalman Y 2014 Histogram based perceptual quality assessment method for color images *Comput. Stand. Interfaces* **36** 899–908
- [14] Immerkaer J 1996 Fast Noise Variance Estimation *Comput. Vis. Image Underst.* **64** 300–2
- [15] E. C. Larson and Chandler D M 2010 Most Apparent Distortion: Full-Reference Image Quality Assessment and the Role of Strategy *J. Electron. Imaging* **19**
- [16] Ponomarenko N, Lukin V, Zelensky A, Egiazarian K, Astola J, Carli M and Battisti F 2009 TID2008-A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics *Adv. Mod. Radioelectron.* **10** 30–45
- [17] Ponomarenko N, Jin L, Ieremeiev O, Lukin V, Egiazarian K, Astola J, Vozel B, Chehdi K, Carli M, Battisti F and Jay Kuo C-C 2015 Image database TID2013: Peculiarities, results and perspectives *Signal Process. Image Commun.* **30** 57–77