Advances in Natural and Applied Sciences, 8(19) Special 2014, Pages: 23-30

AENSI Journals

Advances in Natural and Applied Sciences

ISSN:1995-0772 EISSN: 1998-1090

Journal home page: www.aensiweb.com/ANAS

Impact of Image Size on Performance of IQA Algorithms on Large Images

[1]K.R. Joy and [2]E. Gopalakrishna Sarma

[1]*Research Scholar, Electronics & Communication Engineering Department, Karpagam University, Coimbatore 641021, India*
[2]*Principal, Sree Buddha College of Engineering for women, Pathanamthitta, Kerala, India*

ARTICLE INFO

Article history:
Received 3 September 2014
Received in revised form 30 October 2014
Accepted 4 November 2014

Keywords:
Image Quality Assessment, IQA, FR-IQA, IQA Runtime, Large Images

ABSTRACT

Background: Image Quality Assessment (IQA) tools are widely used in image processing applications to estimate the quality of processed image. The demand for accurate, consistent, computationally simple and easy-to-use IQA tools is increasing day by day. The sizes of digital images used in most of our applications are increasing at very fast rate. But, most of the currently available IQA algorithms were evaluated on images of relatively smaller sizes. The performance of these algorithms on large images (HD, UHD and above) are not fully studied yet. **Objective:** To study the performance of selected Full Reference IQA algorithms on large images in terms of their runtime, nature of variation of runtime with image size and the impact of down-sampling on the estimated quality. **Results:** Algorithms having very good performance with small images took long runtime and inflated quality figures with large images. **Conclusion:** Many IQA algorithms in their existing form may not work accurately and efficiently for large images and need to be adapted before deploying them in applications.

To Cite This Article: K.R. Joy and E. Gopalakrishna Sarma, Impact of Image Size on Performance of IQA Algorithms on Large Images. **Adv. in Nat. Appl. Sci.,** *8(19): 23-30, 2014*

INTRODUCTION

Applications of digital image processing are expanding at very fast rate in various fields of human life including security and surveillance, medical imaging, remote sensing, astronomy, multimedia communications, entertainment etc. The significance of Image Quality Assessment is also increasing, as most of the above applications have to monitor and ensure the quality of the processed image. IQA is the process and methodology of assessing the quality of an image for the intended purpose. As most of the images are ultimately viewed by human beings, the best method to judge the quality of the image is by visually evaluating the images by human observers. However, such subjective evaluation is time consuming, expensive and not practical in real-time or embedded systems. Therefore, it is well accepted that objective assessment methods, which mimic the human visual system (HVS) can do this task very conveniently. In image processing IQA has a number of applications which include, monitoring the image quality for controlling quality of processing systems, benchmarking image processing systems, optimizing algorithms and parameter settings for image processing systems etc. (Wang Z., 2011).

Existing work:

In the earlier stages of IQA research, the main focus of the researchers was to predict the quality of the image consistently as close as to how human observers assess the quality. Little importance was given to the runtime efficiency of these algorithms. The major performance indicators for the IQA algorithms were Spearman Rank Order Correlation Coefficient (SROCC), Kendall Rank Order Correlation Coefficient (KROCC), Pearson Linear Correlation Coefficient (PLCC) and Root Mean Squared Error (RMSE) (VQEG 2003, Sheik *et al* 2006, Wang Z. and Li Q., 2011). Later, researchers started realizing the importance of runtime efficiency and started evaluating and optimizing this parameter also (Silva *et al* 2007, Hofbauer Heinz and Andreas Uhl 2011, Xue Wufeng 2013). At present, a number of IQA algorithms are available in the literature (K.R. Joy and E. Gopalakrishna Sarma, 2014). These algorithms work well with most of the public image databases because most of these algorithms were validated on these images, where the image sizes are relatively small (W. Lin and C.-C. Jay Kuo 2011). Their performances on large images have not yet been studied. A comprehensive evaluation of FR-IQA algorithms was performed on public image databases by Lin Zhang *et al.* in 2012 and the performance indicators including runtime were published. Runtime comparisons of thirteen FR-

Corresponding Author: K.R. Joy, Research Scholar, Department of Electronics & Communication Engineering, Karpagam University, Coimbatore 641021, India.
Tel: +91 8281918668 E-mail: joykarotte@gmail.com

IQA algorithms have been performed by Wufeng Xue *et al*. in 2014 using public image databases while validating their new FR-IQA algorithm GMSD. The results of these experiments show that many algorithms which performed very well in terms of SROCC, KROCC, PLCC and RMSE, have very long execution time. Even though these works have given some insight to the runtime efficiency of IQA algorithms for small images, they could not provide information about the behavior of these algorithms on large images. Moreover, the nature of variation of runtime with image size was not covered in those studies. Algorithms having a linear increase of time and memory cost, with image size are becoming more and more attractive.

Image Down-sampling:

Many currently available IQA algorithms deploy automatic down-sampling of the input images while performing the IQA in order to adjust for the viewing distance (W. Lin and C.-C. Jay Kuo 2011, Wang Z. *et al* online:). An empirical relation for selecting the down-sampling scale f can be obtained from the following MATLAB statements.

```
[M N]=size(im1);
f = max(1,round(min(M,N)/256));
```

where im1 is the reference image. The scale f as obtained above is referred as the down-sampling factor throughout this article. By down-sampling, the size of the input images will be reduced by a factor f, which depends upon the original image size. For example, the value of f is 14 for an input image of size 5202*3465 pixels. It is clear that down-sampling the images as above would result in loss of information and that may cause unexpected values for the estimated quality.

Proposed work:

The performances of IQA algorithms on large images have not yet been studied in detail. One reason for this may be the unavailability of large test images in public image databases. In the recent years, image acquisition and processing have undergone rapid advancements. Processing and use of large digital images (HD, UHD and above) have become very common. In this paper, we evaluate the performance of selected FR-IQA algorithms on large images, for the runtime, nature of the variation of runtime with image size and the impact of image down-sampling on the estimated quality. In the following sections, the classification of IQA algorithms, a brief description of FR-IQA algorithms used in this study, details of the experiments conducted, the results and discussions and the conclusion are given. At the end of this article, the list of references used in this study is given.

MATERIALS AND METHODS

Classification of IQA Algorithms:

IQA algorithms can be broadly classified in to three categories depending upon a reference image is available or not. They are No-Reference IQA (NR-IQA or Blind IQA), Full Reference IQA (FR-IQA) and Reduced Reference IQA (RR-IQA) (Chandler DM., 2013, Tsung-Jung Liu *et al* 2013).

NR-IQA refers to quality assessment of an image without a reference image (Wang Z and Bovik AC., 2011). Our visual system can easily distinguish high-quality images and low-quality images with little effort and without seeing the original image. There are three basic approaches towards NR-IQA based on how the objective algorithm derives the quality score (Saad *et al* 2010). They are: 1. Distortion-Specific approach, 2. Feature extraction and learning based approach: and 3. Natural Scene Statistics based approach (NSS).

FR-IQA uses a reference image for the assessment of quality of the distorted image. Since this method has the complete information about the reference image, the results of FR-IQA are supposed to be superior to other IQA algorithms. Some of the approaches for FR-IQA are those based on image fidelity and pixel wise differences, those based on HVS, those based on image structures, those based on information content, those based on image statistics and machine learning etc. (Chandler DM., 2013). Some popular FR-IQA algorithms are explained in the following section.

In Reduced Reference IQA model, the quality of the distorted image is assessed with partial information from the reference image (Wang Z. and Bovik A.C., 2011, Li Q. and Wang Z., 2009). The partial information is the features extracted from the reference image. It is a compromise between FR and NR approaches to IQA in terms of quality prediction accuracy and the amount of information required describing the reference image. In the case of FR-IQA, the reference image is always required to estimate the quality of the distorted image, but the results are reliable. NR-IQA does not require a reference image, but results are not proven to be consistently reliable in performance.

Algorithms used in this study:

In this section, we explain some important full reference IQA algorithms used in this study.

Mean Squared Error (MSE):

MSE computes the mean of the squares of the errors of the distorted image with reference to the reference image on a pixel by pixel basis as given below (Rafael. C. Gonzalez and Richard E. Woods, 2008, Wang Z. and Bovik A.C., 2009).

$$\mathrm{MSE} = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}[f(x,y) - g(x,y)]^2 \quad (1)$$

where $f(x,y)$ and $g(x, y)$ are the reference and distorted images respectively of size M*N pixels. The advantage of this metric is its simplicity, but has poor correlation with human perception of quality.

Structural Similarity Index (SSIM):

Natural images are highly structured and their pixel values exhibit strong dependencies. SSIM is an IQA algorithm based on these structural dependencies with in an image (Wang Z. *et al*, 2004). The human visual system is highly adapted to extract structural information from the viewing field. The SSIM algorithm separates the luminance component l(f,g), contrast component c(f,g) and the structural component s(f,g) from the reference image (f) and the distorted image (g) and compares these components. SSIM index is calculated as

$$\mathrm{SSIM}(f,g) = [l(f,g)]^{\alpha}.[c(f,g)]^{\beta}.[s(f,g)]^{\gamma} \quad (2)$$

where $\alpha > 0, \beta >$and $\gamma > 0$ are constants used to adjust the relative importance of the three comparisons. The luminous, contrast and structural components are computed as follows.

$$l(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1} \quad (3)$$

$$c(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \quad (4)$$

$$s(f,g) = \frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3} \quad (5)$$

The parameters μ_f, σ_f and μ_g, σ_g are the mean and standard deviations of images f and g respectively. The parameter σ_{fg} is calculated as

$$\sigma_{fg} = \frac{1}{N-1}\sum_{i=1}^{N}(f_i - \mu_f)(g_i - \mu_g) \quad (6)$$

The constants C_1, C_2 and C_3 are included to avoid instability when when $(\mu_f^2 + \mu_g^2)$, $(\sigma_f^2 + \sigma_g^2)$ or $\sigma_f\sigma_g$ are very close to zero. These values are selected in such a way that $C_1 = (K_1L)^2$ and $C_2 = (K_2L)^2$. K_1 and K_2 are constants such that $K_1 << 1$, $K_2 << 1$and L is the dynamic range of the pixel values (L=255 for 8-bit gray scale images). The equation for SSIM can be simplified by putting $\alpha = \beta = \gamma = 1$. The value of C_3 is normally taken as $C_2/2$. SSIM in its simplified form can be written as

$$\mathrm{SSIM}(f,g) = \frac{(2\mu_f\mu_g + C_1)(2\sigma_{fg} + C_2)}{(\mu_f^2 + \mu_g^2 + C_1)(\sigma_f^2 + \sigma_g^2 + C_2)} \quad (7)$$

In practice, SSIM is calculated for local windows as per equation (7) and the mean of SSIM (MSSIM) is used as the single overall quality measure. The influences of various parameters and window size on the quality index have been studied in detail by Javier Silvestre-Blane, 2011. The main advantage of SSIM is that it has a good correlation with the subjective test results over a wide range of distortion types. An improved version of SSIM known as Multi Scale SSIM (MS-SSIM) has still better performance compared to single scale SSIM (Wang Z. at al, 2003). SSIM and its variants are superior to other algorithms such as MSE, SNR and PSNR. This is because SSIM treats image degradations as structural changes and it mimics the HVS to certain extend.

Feature Similarity Index (FSIM):

FSIM is based on the theory that HVS understands an image based on its low level features such as edges, and a good IQA metric could be obtained by comparing these low level features of the reference image and the distorted image (Lin Zhang *et al*, 2011). At points of high phase congruency of the Fourier waves of different frequencies of the image, highly informative features can be extracted. FSIM utilizes this property of the Fourier transform of the reference and distorted images for the quality assessment. In FSIM, the phase congruency (PC) and the image gradient magnitude (GM) are computed for the quality assessment of the distorted image with respect to the reference image. The PC of the image is computed using the response of a 2-D log-Gabor function. The GM is calculated as $G = \sqrt{G_x^2 + G_y^2}$ where G_x and G_y are the partial derivatives of the image along the x and y directions. The gradient operator used was Scharr which gave better performance compared to Sobel or Prewitt operators. The performances of the FSIM and FSIMc (for color images) were superior over many other IQA algorithms for a variety of image databases.

IQA based on Spectral Residuals (SR-SIM):

The spectral residual visual saliency (SRVS) is used in the computation of this metric. The hypothesis behind this approach is that an image's perceived quality is related to its visual saliency map. In this method, the Visual Saliency (VS) is calculated for the reference and distorted images based on Spectral Residual. The Gradient Modulus (GM) is calculated using the Scharr operator. The local values for SR-SIM is calculated using the two components namely SRVS and GM using the relation $S(x) = S_V(x).[G_V(x)]^{\alpha}$ where $S_V(x)$ and $G_V(x)$ are the local values for SRVS and GM and α is a parameter used to adjust the relative strength of these components. After obtaining the local values for S(x), the global value is calculated by applying suitable pooling mechanisms (Lin Zhang and Hongyu Li, 2012). The overall performance of this algorithm was superior to many existing FR-IQA algorithms.

Edge Strength Similarity Index (ESSIM):

It is based on the hypothesis that HVS is more sensitive to the direction showing stronger edge strength (Xuande Zhang *et al* 2013). Any directional high pass filters can be used to define the edge strength. Different gradient operators such as Sobel, Prewitt or Scharr can be used to extract the edge strength. The edge strengths are calculated in the horizontal-vertical direction and in the diagonal direction. The maximum of these two values is taken as the edge strength at any point. The ESSIM index is defined as

$$\mathrm{ESSIM}(f,g) = \frac{1}{N}\sum_{i=1}^{N}\frac{2E(f,i)E(g,i)+C}{(E(f,i))^2+(E(g,i))^2+C} \quad (8)$$

where f and g are the reference and distorted images, N is the total no. of pixels in f or g, E(f,i) and E(g,i) are the edge strength at pixel "i" of images f and g respectively. C is a scaling parameter such that $C= (BL)^2$ where B is a constant and L is the dynamic range of edge strength. It has been shown that the ESSIM has good correlation with HVS.

Details of Experiment:

The experiments were conducted on a personal computer with Intel Core i5-2430 M processor with 4 GB RAM and Windows 7 operating system. For measuring the run time, the algorithm was executed a number of times and the average runtime was recorded. The MATLAB source codes by the authors of the respective algorithms were used in the experiment for MSSIM, FSIM, SRSIM and ESSIM.

The images from public image databases were not used in this experiment due to their relatively smaller size. Large images were captured using a digital camera in RAW (.CR2) format and then converted into bitmap image format (.BMP) for this study. The original images are 8-bit RGB colour images of size 5202*3465*3 pixels. Gray images are generated using rgb2gray function. Images are scaled down using 'imresize' function of MATLAB to get smaller images. The distorted images are obtained by using the MATLAB function 'imnoise' with type 'gaussian'. JPEG compressed images are generated using MATLAB function 'imwrite'.

The algorithms for this study were selected in such a way that they cover the different approaches to the FR-IQA and also they represent the important stages in the evolution of FR-IQA. They include MSE, SSIM, FSIM, SRSIM and ESSIM. The important performance parameters of these algorithms for the LIVE2 database (Sheik H.R. *et al*, 2005) are given in Table 1.

Table 1: Algorithms used in the study.

Algorithm	SROCC	KROCC	PLCC	RMSE
MSE	0.8756	0.6865	0.8723	13.3597
MSSIM	0.9479	0.7963	0.9449	8.6188
FSIM	0.9634	0.8337	0.9597	7.678
SRSIM	0.9618	0.8299	0.9553	8.0811
ESSIM	0.9622	0.8397	-	-

The experiments were conducted in two parts. The first part was to find the dependence of runtime on image size. For this purpose images of five different sizes are used. The IQA algorithm was executed on a reference and distorted image pair for a number of times and the average runtime was recorded. The experiment was repeated for the next image size and the same procedure was repeated. The above procedure is done for distortion types 'gaussian noise' and JPEG compression. A similar procedure was done for grayscale images also. The results are summarized in Table 2 and Table 3 and the graphical representation are given in Figures 1 and 2. Since there is no noticeable difference between the runtimes of 'gaussian noise' images and JPEG compressed images, the latter is omitted from the tables and figures.

The second part of experiments was to study the impact of down-sampling on estimated quality. For this purpose, IQA experiments were performed on a particular image pair for different values of down-sampling factor f as explained above. The down-sampling factor f was passed to the program as an argument and the estimated quality index and the runtime were recorded. The results are given in Figures 3 and 4.

RESULTS AND DISCUSSIONS

Runtime and image size:

MSE was included in this study as a reference for a comparison for other algorithms. The average runtime for MSE shows a linear relation with the image size as shown in figure 1. This is expected because in MSE, the calculations have to be performed as many times as the number of pixels. MSE was the fastest out of the five algorithms tested. It took an average runtime of 0.014 and 0.039 second respectively for the smallest gray scale and colour images of size 1041*693 pixels. Similarly, the runtime for the largest images of 5202*3465 pixels are 0.329 and 0.983 second respectively. The performance of this algorithm is excellent from the runtime point of view and also due to its linear relation with image size.

Table 2: Average Runtime for grayscale images.

Image Size	Runtime (Sec)				
	MSE	MSSIM	ESSIM	SRSIM	FSIM
5202*3465	0.329	1.067	1.073	2.534	8.497
4162*2772	0.208	0.626	0.624	1.116	4.274
3122*2079	0.118	0.352	0.358	0.432	2.155
2081*1386	0.053	0.160	0.174	0.159	1.853
1041*693	0.014	0.054	0.063	0.055	0.956

However, its correlation with the perceived quality is very poor. Modified versions of MSE algorithms yield better correlation with perceived quality with very high execution speed (Wufeng Xue *et al*, 2013).

MSSIM also performed excellently on gray scale and colour images with a near-linear variation with the image size and performed close to MSE with better correlation with the perceived quality. The runtime for the smallest and the largest gray scale images are 0.054 and 1.067 seconds respectively and 0.074 and 1.407 seconds respectively for colour images.

SRSIM is faster than ESSIM for small images. But due to its nonlinear nature, it becomes slower than ESSIM for large images. Hence ESSIM is the better choice for large images compared to SRSIM.

Out of the five algorithms tested, FSIM is the slowest and showed a nonlinear variation with image size. It took an average of 0.956 and 1.034 seconds for the smallest grayscale and colour images respectively and 8.496 and 10.308 seconds for the largest images. FSIM has very good correlation with perceived quality and prediction accuracy as can be seen from Table 1. But a runtime of nearly 10 seconds may not be acceptable for many image processing applications.

The Edge Strength Similarity Index ESSIM is highly promising due to its speed, quality prediction accuracy and linear variation of runtime with the image size. For grayscale images, the curve almost coincides with MSSIM indicating its excellent performance. In summary, ESSIM performs well compared to other algorithms in terms of execution speed, quality prediction accuracy and linearity for large images.

Out of the five algorithms tested, MSE was the fastest, but has poor correlation with human perception of quality. ESSIM and MSSIM performed almost equally for gray scale images in terms of execution speed and linearity. However, ESSIM has better quality prediction accuracy.

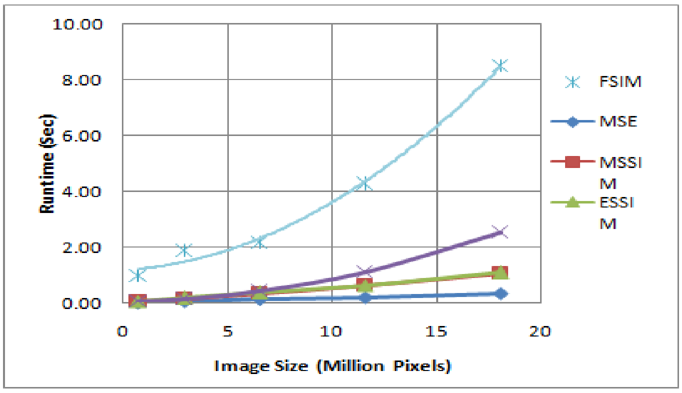

Fig. 1: Variation of runtime with image size for grayscale images.

Table 3: Average Runtime for colour images.

Image Size	Runtime (Sec)				
	MSE	MSSIM	ESSIM	SRSIM	FSIM
5202*3465	0.983	1.407	2.419	3.050	10.308
4162*2772	0.620	0.848	1.482	1.451	5.429
3122*2079	0.346	0.480	0.847	0.625	2.833
2081*1386	0.154	0.215	0.390	0.233	2.161
1041*693	0.039	0.074	0.119	0.078	1.034

Advances in Natural and Applied Sciences, 8(19) Special 2014, Pages: 23-30

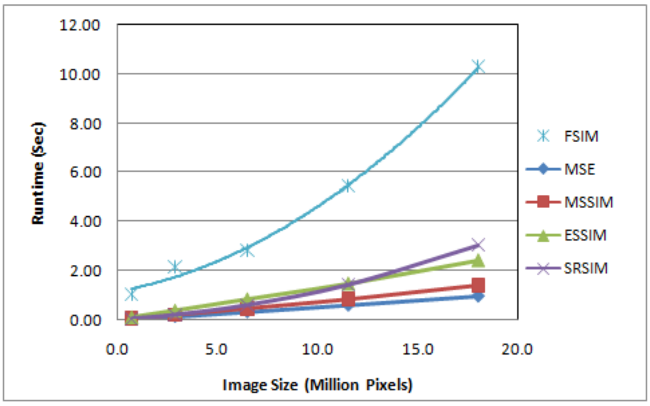

Fig. 2: Variation of runtime with image size for colour images.

Impact of down-sampling on estimated quality index:

All the algorithms in this study except MSE deploy automatic down-sampling the input images by a factor f as explained earlier. For small images, down-sampling may not contribute much difference in the estimated quality. But as the image size increases, this becomes very significant. The variation of estimated quality with the down-sampling factor and the variation of runtime with down-sampling factor are shown in two illustrative cases in Figures 3 and 4.

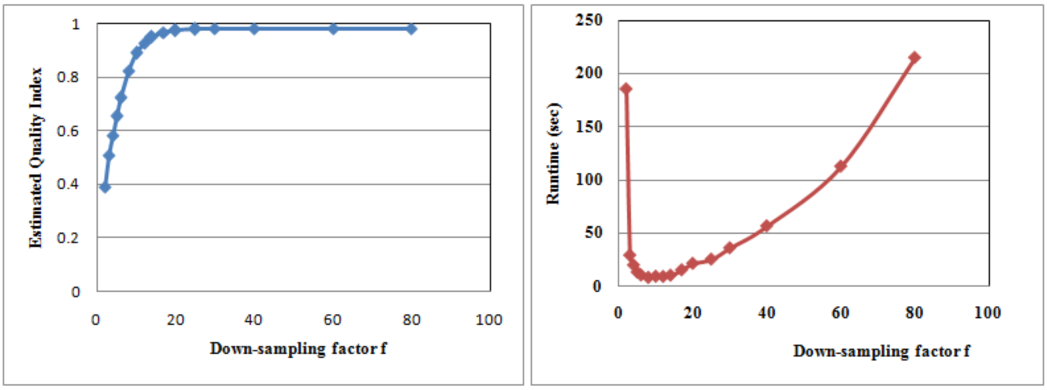

Fig. 3: Variation of estimated quality index and runtime with down-sampling factor f for FSIM algorithm for colour image of 5202*3465 Pixels.

From the above figures, it can be seen that the value of the estimated quality index varies almost linearly for small values of f. As f increases, it varies in a non-linear manner and with further increase of f, the quality index saturates in to its maximum value. This is due to the fact that down-sampling the input images causes loss of information (loss of details) from the reference and distorted images. This loss of details reduces the difference between the two images thereby increase in the estimated quality. As the value of f is increased further, the loss of information would be to such an extent that the difference between the reference and the distorted images becomes almost zero and the quality index saturates to its maximum value of 1. The variation of runtime with f can be explained as follows. For small values of f, the size of the down-sampled image reduces, which makes the runtime low. But for higher values of f, even though the size of down-sampled images reduced, the computational cost for the down-sampling procedure become very high, as a result the overall runtime increases.

It may be noted that the general pattern of variation of the quality index and runtime with reference to down-sampling factor f are the same as above, the values depend on the particular algorithm as well as the images to be compared. As an empirical rule, for better prediction accuracy, the value of **f** should lie in the lower side of the linear region of the curve.

Conclusion:

In this paper, we have evaluated the performance of five popular FR-IQA algorithms in terms of runtime and the nature of variation of runtime with image size for very large images. The absolute values of runtime for various algorithms may vary with the processor configurations and the experimental setup. However, it gives a relative comparison about the performance of various algorithms when applied to large images. It is important to note that many algorithms in their existing form may not work accurately and efficiently for large images and need to be adapted before deploying them in applications. This study can facilitate the IQA research community to understand the importance of runtime and linear variation of runtime with image size as important performance indicators. This may inspire researchers to conduct such tests as part of the evaluation for their algorithms. Even though we have included only five FR-IQA algorithms in this study, similar studies for other IQA algorithms are also very relevant. The relation between the down-sampling factor and the quality index gives an idea for selecting appropriate values for down-sampling. It is also worth to mention here that one of the major constraints in conducting such evaluation in a systematic manner is the lack of standard test images of large size (HD, UHD and above) in public image databases. It is hoped that the public image databases will be updated with images of large sizes.

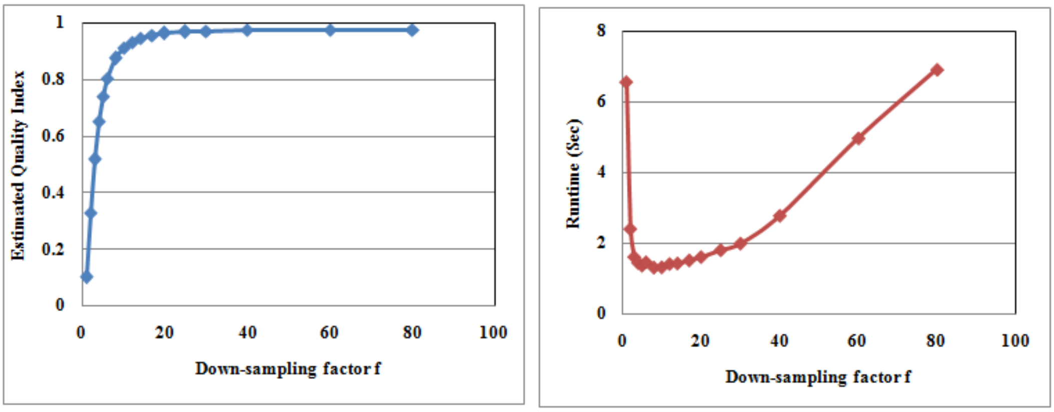

Fig. 4: Variation of estimated quality index and runtime with down-sampling factor f for MSSIM algorithm for colour image of 5202*3465 Pixels

REFERENCES

Chandler, D.M., 2013. Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Processing*.

Hofbauer Heinz and Andreas Uhl, 2011. An effective and efficient visual quality index based on local edge gradients. IEEE 3rd European Workshop on Visual Information Processing (EUVIP): 162-167.

Javier Silvestre-Blane, 2011. Structural Similarity Image Quality Reliability Determining Parameters and Window Size. Signal Processing, 91(4): 1012-1020

Joy, K.R. and E. Gopalakrishna Sarma, 2014. Recent Developments in Image Quality Assessment Algorithms: A Review. Journal of Theoretical and Applied Information Technology, 65(1): 192-201

Li, Q. and Z. Wang, 2009. Reduced-reference image quality assessment using divisive normalization-based image representation. *IEEE Journal of Selected Topics in Signal Processing,* 3(2): 202-211.

Lin Zhang and Li. Hongyu, 2012. SR-SIM: A Fast and High Performance IQA Index Based on Spectral Residual. IEEE International Conference on Image Processing, (ICIP): 1473-1476.

Lin Zhang, Lei Zhang, Xuanqin Mou and David Zhang, 2011. FSIM: Feature Similarity Index for Image Quality Assessment. IEEE Transactions on Image Processing, 20(8): 2378-2386.

Lin Zhang, Lei Zhang, Xuanqin Mou and David Zhang, 2012. A comprehensive evaluation of full reference image quality assessment algorithms. *IEEE International Conference on Image Processing, (ICIP):* 1477-1480.

Lin, W. and C.C. Jay Kuo, 2011. Perceptual visual quality metrics: A survey. Journal of Visual Communication and Image Representation, 22(4): 297–312.

Rafael, C. Gonzalez and Richard E. Woods, 2008. Digital Image Processing, 3rd edition, Pearson Education Inc., pp: 354-355.

Saad, M.A., A.C. Bovik and C. Charrier, 2010. A DCT statistics-based blind image quality index. *IEEE Signal Processing Letters*, 17(6): 583-586.

Sheikh, H.R., M.F. Sabir and A.C. Bovik, 2003. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. IP,* 15(11): 3440-3451.

Sheikh, H.R., Z. Wang, L. Cormack, A.C. Bovik, 2005. LIVE image quality assessment database release 2, http://live.ece.utexas.edu/research/quality

Silva, A. Eric, Karen Panetta and S. Sos Agaian, 2007. Quantifying image similarity using measure of enhancement by entropy. *Defense and Security Symposium*. International Society for Optics and Photonics: 65790U-65790U

Tsung-Jung Liu, Yu-Chieh, Weisi Lin and C.C. Jay Kuo, 2013. Visual Quality Assessment: recent developments, coding, applications and future trends. APSIPA Transactions on Signal and Information Processing, 2, e4

VQEG, 2003. Final Report From the Video Quality Experts Group on the validation of Objective Models of Video Quality Assessment, Phase II, VQEG, August 2003. Online: http://www.vqeg.org

Wang, Z. and A.C. Bovik, 2009. Mean squared error: love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine,* 26(1): 98-117.

Wang, Z. and A.C. Bovik, 2011. Reduced-and no-reference image quality assessment. *IEEE Signal Processing Magazine,* 28(6): 29-40.

Wang, Z. and Q. Li, 2011. Information content weighting for perceptual image quality assessment. IEEE Trans. IP, 20(5): 1185-1198.

Wang, Z., 2011. Applications of objective image quality assessment methods. *IEEE Signal Processing Magazine, 28*(6): 137-142.

Wang, Z., A.C. Bovik, Hamid Rahim Sheik and Eero P. Simocelli, 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Processing, 13(4): 600-612.

Wang, Z., A.C. Bovik, Hamid Rahim Sheik and P. Eero Simoncelli, The SSIM Index for Image Quality Assessment; Online: https://ece.uwaterloo.ca/~z70wang/research/ssim/

Wang, Z., E.P. Simoncelli, A.C. Bovik, 2003. Multiscale structural similarity for image quality assessment. IEEE *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers,* 2: 1398-1402.

Wufeng Xue, Lei Zhang, Xuanqin Mou and A.C. Bovik, 2014. Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index. IEEE Transactions on Image Processing, 23(2): 684-695.

Wufeng Xue, Xuanqin Mou, Lei Zhang and Xiangchu Feng, 2013. Perceptual Fidelity Aware Mean Squared Error, ICCV, Computer Vision Foundation, pp: 705-712.

Xuande Zhang, Xiangchu Feng, Weiwei Wang, Wufen Xue, 2013. Edge Strength Similarity for Image Quality Assessment. IEEE Signal Processing Letters, 20(4): 319-322.