Quality indicators for reproducing fine details of digital images with threshold limiting of spectral components

Olena Osharovska
Department of Television and Sound Broadcasting
O.S. Popov Odessa National Academy of Telecommunications
Odessa, Ukraine
osharovskaya@gmail.com

Mikola Patlayenko
Department of Television and Sound Broadcasting
O.S. Popov Odessa National Academy of Telecommunications
Odessa, Ukraine
nick_msa@ukr.net

Natalia Samus
Department of Television and Sound Broadcasting
O.S. Popov Odessa National Academy of Telecommunications
Odessa, Ukraine
natalia_samus@ukr.net

***Abstract*—The article considers the effect of compression of the information stream on high-definition image quality image parameters on the receiving side. In order to assess objectively the quality of the reconstructed images, two parameters are selected: the peak-to-noise ratio and the ratio of the peak signal-to-noise value at the boundaries of the image objects. The boundaries of objects are proposed to be selected by the gradient method, after that to make a binary matrix whose area will depend on the set threshold of significant signal values. With the use of spectral transformations and the rejection of small spectral components, the compression of the digital stream is carried out with losses. Several quantization matrices of frequency components are considered and it is shown that, in some cases, image textures that have low contrast are distorted or disappeared. Compression coefficients and corresponding signal-to-noise ratios for frequency-dependent quantization matrices are calculated. In this work, we have also resulted in the dependence of the decoded image quality parameters depending on the length of words that define the brightness and color difference signals.**

Keywords*—*image, compression, frequency-dependent quantization, quality indicators

I. INTRODUCTION

Advances in digital image processing and image processing technologies have revolutionized our way of life. The acquisition of images, storage, transmission, viewing, and processing technologies have undergone incredible achievements in recent years.

In our daily life, we use a number of applications for image processing with or without our knowledge. For example, when someone captures a scene using a mobile phone, the image captured by the sensor after appropriate correction is compressed in JPEG format and stored in memory. The image can then be transferred to the social media network via the communication channel. [1-7]. The user on the computer screen may later view the image; the pixel size is smaller than the actual image size. In this case, the image must be changed in order to fit on the display screen.

During these operations, the original image undergoes changes that may affect image quality. Therefore, it is necessary to evaluate the suitability of the received (extracted) image for use for its intended purpose. Since most images are ultimately intended for viewing by observers, the only reliable test for evaluating image quality is a subjective examination that allows you to visually evaluate an image by a group of observers and derive a statistically reliable quality assessment [4-6]. Subjective image estimation not only takes a long time, but also a very expensive. The procedure is not practical in real-time applications. In addition, there may be individual factors that can affect perceived image quality. Therefore, it is necessary to evaluate the image quality objectively, taking into account the properties of the human visual system (HVS) as the basis for such an assessment. Any objective algorithm for assessing the quality of IQA images must meet the following requirements: (1) it must have a close connection with visual perception; (2) it must work in a wide range of types of distortion; (3) it must be computationally simple and efficient, and (4) it can be embedded in imaging systems or allow real-time evaluation.

II. ALGORITHM FOR ASSESSING THE QUALITY

Accordingly, IQA algorithms can be broadly classified into three categories, namely, without comparison with a reference image, without reference IQA (No-Reference IQA), algorithms using partial references to the reference image (Reduced Reference IQA), and algorithms determining image quality by a full comparison with the standard, called the Full Reference IQA.

The method for predicting image quality using the Full-Reference IQA (FR-IQA) algorithm uses a reference image to estimate the quality of the distorted image. Since this method has complete information about the reference image, the FR-IQA results must be higher than other prediction algorithms IQA: Some approaches to FR-IQA are based on the accuracy of image representation, accumulated errors, RGB or HVS color characteristics, image structure, content, image statistics and machine capabilities etc.

A. Mean Squared Error

This algorithm calculates the root-mean-square error between the examined image and the original image pixel-by-pixel. Usually the MSE is calculated according to formula (1):

$$MSE = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left[B_r(x,y) - B_d(x,y)\right]^2 \quad (1)$$

$B_r(x,y)$ the reference image;

$B_d(x,y)$ the distorted image;

$M \times N$ the image dimension.

The advantage of this metric is its simplicity, but the MSE has a poor correlation with the subjective test results.

B. The Ratio of the Peak Value of the Signal to the Average Value of the Noise

The ratio of the peak value of the signal to the average noise (PSNR). This method also compares the reference image and the distorted image for each pixel and calculates the PSNR [9-11] as follows (2):

$$PSNR = 10\log_{10}\frac{2^P - 1}{MSE}, dB \quad (2)$$

Parameter P is the bit depth of the image representation of the pixel brightness. The main disadvantage of PSNR is a weak correlation with HVS.

Expression (2) can be represented in the form (3):

$$PSNR = 20\log_{10}\left(\frac{B_{pic}(x,y)}{\sqrt{MSE}}\right), dB \quad (3)$$

where $B_{pic}(x,y)$ is the maximum value of image signal.

PSNR is most commonly used to measure the quality of reconstruction of lossy compression codecs (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs, PSNR is an approximation to human perception of reconstruction quality.

Typical values for the PSNR in loss image and video compression are between 30 and 50 dB, provided the bit depth is 8 bits, where higher is better. For 16-bit data, typical values for the PSNR are between 60 and 80 dB, [9, 10]. Acceptable values for wireless transmission quality loss are considered to be about 20 dB to 25 dB, [11, 12].

C. The PSNR of the edge areas

The PSNR of the edge areas (EPNSR) is computed as follows:

$$EPSNR = 10\log_{10}\left(\frac{P^2}{MSE_{edge}}\right), dB \quad (4)$$

where:

P : peak pixel value;

MSE_{edge} average noise near bounders.

In the model, this EPSNR is used as a basic objective video quality score.

In the model, an edge detection operator is first applied, producing edge images. It is noted that this edge detection algorithm is applied to the source image (fig.1). Then, a mask image (binary edge image) is produced by applying thresholding to the edge image. In other words, pixels of the edge image whose value is smaller than threshold, t_e, are set to zero and pixels whose value is equal to or larger than the threshold are set to a non-zero value. Fig. 2 and fig. 3 show examples of mask images. Although one may apply the edge detection algorithm to processed images, it is more accurate to apply it to the source images. Next, differences between the source image and processed image, corresponding to non-zero pixels of the mask image are computed. In other words, the squared error of edge areas of the frame is computed as follows to (1).

To estimate EPSNR for separation boundaries and their surroundings, the gradient of the image brightness distribution function is used as the sum of the gradient projection modules $\mathrm{grad}_{\mathrm{horizontal}}(x,y)$ and $\mathrm{grad}_{\mathrm{vertical}}(x,y)$ on horizontal and vertical axes (4):

$$\mathrm{grad}(x,y) = \left|\mathrm{grad}_{\mathrm{horizontal}}(x,y)\right| + \left|\mathrm{grad}_{\mathrm{vertical}}(x,y)\right|, \quad (5)$$

where x and y are the image coordinates horizontally and vertically.

Fig. 1. Example of a image "Vehicles truck"

Fig. 2. Example of an edge image "Vehicles truck"

Fig. 3. Example of a mask image "Vehicles truck"

The resulting mask is multiply per pixel by the brightness component of the image, highlighting the values on the contours. In some cases, it is advisable to subject the luminance signal to low-frequency filtering.

III. FREQUENCY-DEPENDENT QUANTIZATION OF SPECTRAL COMPONENTS

The progress of compression methods for still images and intra-frame compression of video sequences is a key element in the construction of broadband image transmission systems for various purposes. In this respect, the main criterion for the degree of compression is the estimates obtained by measuring distortions of the sharp image boundaries arising as a result of nonlinear processing of the spectral components.

In this work, the main attention is paid to the problem of choosing a threshold level on the basis of a compromise between the achievable compression ratio and the possible preservation of the image texture and the corresponding estimates are given. This compromise can be of great importance both for broadcast applications and for a wide range of video applications, in which a large number of stages of production, processing, storage, transmission and reproduction of video information where texture transmission can play a significant role are realized.

When you try to save a digital image in a smaller volume, you often have to decide what "quality settings" (compression level) to use. The JPEG file format allows you to choose the appropriate trade-off between file size and image quality. It is important to understand that JPEG (and almost all lossy file formats) are not suitable for intermediate editing because repetitive saving usually reduce the quality of the working file. In addition to the cumulative introduction of visual artifacts, repeated decompression also leads to destructive color changes. It is for these reasons that the preferred choice for intermediate processing is lossless compression of file formats (such as TIFF, PSD, BMP, etc.). JPEG should only be used to store the final image.

Suppose that there is a color image with dimensions N × N = 2048 × 2048 pixels, each element is represented by m = 10-bit code in each of the three components. In this case, the required amount of memory V needed to store one still image RGB or Y Cb Cr is (5):

$$M=3 \times N \times N \times m = \tag{6}$$

$$M= 3 \times 2048 \times 2048 \times 10 = 120 \text{ Mbit}.$$

To reduce the required amount of memory, lossy compression is used based on spectral transformations. The most common is the Fourier transform. Compression of information about color images consists of several stages, including both lossless compression based on statistical properties, and lossy compression based on the generalized discrete Fourier transform $F(v_H, v_V)$. T(Tabhe spectrum of the two-dimensional image signal consists of horizontal ν_H, vertical ν_V and diagonal ν_D spatial frequencies (7):

$$F(v_H, v_V) = \sum_{x=0}^{N_x-1} \sum_{y=0}^{N_y-1} B(x,y) e^{-j2\pi\left(\frac{xv_H}{N_x}+\frac{yv_V}{N_y}\right)}, \tag{7}$$

Thanks to the numerous algorithms for fast Fourier transforms, it is possible to obtain the spectrum of the entire image in an acceptable time. The main energy of the spectrum is concentrated at low frequencies, and the high-frequency components are usually very small in amplitude and. in a number of cases, they can be neglected, equating them to zero. The greater the number of high-frequency components will be equated to zero, the greater the compression ratio can be obtained. However, from the point of view of storage and transmission of the image spectrum obtained by the formula (6), a significant drawback is the need to work with two arrays representing the real and imaginary part of the complex numbers of the spectrum.

In practice, special cases of Fourier transforms are used, such as discrete cosine transform (DCT) or discrete sine transformation (DSP). The JPEG and MPEG-2 standards use DCT. A feature is segmentation in blocks - raw image data is divided into blocks of 8 × 8 pixels (these blocks are the minimum coded block). This means that the JPEG compression algorithm largely depends on the position and alignment of the boundaries of these blocks. In MPEG-4, the block sizes vary from 4 × 4 to 16 × 16 pixels.

Taking into account the obtained values of the spectral coefficients at the DCT step, they are sorted in order of increasing numbers from the low-frequency components (changes that occur at a greater distance from the image block) to high-frequency components (changes that can occur with each pixel). It is widely known that people are more critical of errors in low-frequency information than high-frequency information. The JPEG algorithm discards many of these high-frequency (noise-like) details and saves slowly changing information about the image. This is done by dividing all the spectral coefficients by the corresponding value in the quantization table and rounding the result to the nearest integer. Components that either had a small coefficient or a large divisor in the quantization table are likely to be rounded to zero. The lower the quality setting, the greater the divisor, which gives a greater chance of getting a zero result. On the other hand, the setting for the highest quality would have the values of the quantization table of all frequencies equal to one, which means that all the original data of the discrete cosine transform (or sinus) is stored.

If the quantization tables correspond to the standard trend of limited compression in low-frequency components that increase to moderate compression in high-frequency components, then the approximate quality factor can really give an idea of how the overall quality can be displayed.

IV. RESULTS OF COMPUTATIONAL EXPERIMENTS

At the first stage, the values of the spectral components were limited, depending on the peak value of the luminance signal in the selected area with a low-contrast texture. Limit thresholds were 0.5%: 1%; 2% and 5%. The results are given in Table I for test images Tab II.

TABLE I. QUALITY INDICATORS FOR FOURIER TRANSFORMATION

Test Image	Threshold of restriction			
	0.5%	1 %	2 %	5 %
	EPSNR, dB			
Vehicles truck	25.1	22	19.8	18.4
Flowers	28.3	26	24.6	21.9
Sea	22.4	21	19.2	18.6
Music Box	23.7	21	20.2	18.7

TABLE II. TEST IMAGES

IMAGE NAME	TEST IMAGE
Flowers	
SEA	
MUSIC BOX	

Subjective evaluation of reconstructed images corresponds to a satisfactory estimate, there is blurring of boundaries and fading of details of low-contrast textures.

In the next step, we calculated signal-to-noise ratios PSNR, EPSNRS and the DCT compression ratio (CR) for the four signal-word lengths. (Table III).

TABLE III. DEPENDENCE OF EPSNR AND CR ON THE WORD LENGTHS

Bit	EPSNR	PSNR	CR	EPSNR	PSNR	CR
	Flowersr			*Sea*		
6	30.5	39.2	27	33.5	42.8	29
8	31.3	40.1	24	34.5	44.1	27
10	31.9	40.9	22	35.3	45.1	26
12	32.5	41.7	18	36.1	46.0	23
	Music Box			*Vehicles truck*		
6	36.2	46.1	27	36.5	46.5	33
8	36.2	46.2	31	37.1	47.2	29
10	36.4	46.3	22	37.5	47.8	30
12	36.7	46.0	18	37.9	48.3	24

We carried out frequency-dependent quantization using the quantization matrices Q of the cameras of known manufacturers. The results of calculating the PSNR EPSNR, CR are summarized in Table IV for one of the test images "Sea"

TABLE IV. DEPENDENCE OF EPSNR ON THE Q-MATRIX

Quantization Table Q-matrix	Sea	
	DCT EPSNR, dB	***DST EPSNR, dB***
SONY - DSC-N2 (fine)	45.1	44.8
NIKON - E8800 (EXTRA)	45.5	46.3
Canon EOS 10D (fine)	46.1	45.2
NIKON - COOLPIX S10 (FINE)	46.8	46.5
NIKON D80 (FINE)	47.4	46.6
Canon PowerShot A700 (superfine)	48.2	49.1
JPEG standard	48.4	45.8

Analysis of the results shows that all these frequency-dependent quantization matrices of the spectral coefficients give an acceptable quality of the reconstructed image, but not a very high compression ratio. Conclusion

CONCLUSION

The use of such quality indicators of recovered images as PSNR, or even better, EPSNR allows you to objectively predict the reproduction of boundaries and textures.

REFERENCES

[1] Recommendation ITU-R BT.1683:2004, “Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference”.

[2] Recommendation ITU-R BT.1867-0:2010, “Objective perceptual visual quality measurement techniques for broadcasting applications using low definition television in the presence of a reduced bandwidth reference”.

[3] Recommendation ITU-R BT.1908:2012, Objective video quality measurement techniques for broadcasting applications using HDTV in the presence of a reduced reference signal.

[4] Recommendation ITU-T J.144:2004, Recommendation BT,1908:2012, “Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference”.

[5] Recommendation ITU-T J.246:2008, Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference.

[6] Recommendation ITU-T J.247: 2008, Objective perceptual multimedia video quality measurement in the presence of a full reference.

[7] Oleg Gofaizen, Olena Osharovska, Mikola Patlayenko, Volodymyr Pyliavskyi, “Test signals for assessment image quality in HD and UHD TV video path”, 2016 8th International Conference on Ultrawideband and Ultrashort Impulse Signals (UWBUSIS), pp. 42–46, 5–11 Sept., 2016, DOI: 10.1109/UWBUSIS.2016.7724147

[8] N. Thomos, N. V. Boulgouris, M. G. Strintzis. “Optimized Transmission of JPEG2000 Streams Over Wireless Channels.” IEEE Transactions on Image Processing, 15 (1) January, 2006.

[9] L. Xiangjun, C. Jianfei, “Robust transmission of JPEG2000 encoded images over packet loss channels.” ICME 2007 (pp. 947-950). School of Computer Engineering, Nanyang Technological University.

[10] David Salomon, Data Compression: The Complete Reference (4 ed.). Springer. p. 281. ISBN 978-1846286025. 2007.

[11] Huynh-Thu, Q.; Ghanbari, M. "Scope of validity of PSNR in image/video quality assessment". Electronics Letters. 44 (13): 800. 2008. doi:10.1049/el:20080522.