Ukrainian Journal of Remote Sensing

journal homepage: www.ujrs.org.ua

https://doi.org/10/36023/ujrs.2022.9.3.216

UDC 621.3:004.932

Usage of different Chroma Subsampling Modes in Image Compression by BPG Coder

B. V. Kovalenko*, V. V. Lukin

Department of Information and Communication Technologies, National Aerospace University, 61070 Kharkiv, Ukraine

A BPG (better portable graphics) coder is a novel approach that aims to replace common standards of compression such as JPEG, JPEG2000 and so on. That is why, the BPG coder needs a detailed analysis of its basic characteristics from the viewpoint of visual quality and compression ratio. The BPG coder can use different modes of chroma subsampling for color and three-channel images and it is worth analyzing and comparing them. In practice, images to be compressed are often noisy. Then, lossy compression of such images has a specific noise filtering effect. In particular, optimal operation point (OOP) might exist where compressed image quality is closer to the corresponding noise-free (true) image than uncompressed (original, noisy) image quality according to certain criterion (metric). It is also needed to analyze the coder performance from compression ratio point of view. In this paper, we pay attention on impact of different chroma subsampling modes on image quality and compression ratio. Based on simulation results obtained for a set of color images, the best possible ways of compression are recommended.

Keywords: color image, lossy image compression, chroma subsampling, BPG coder, visual quality, YCbCr.

1. Introduction

Image compression stays an important part of image processing in modern world. This is due to the fact that amount and size of images increase rapidly in remote sensing (RS), everyday life, advertising, medicine and so on (Nan et al., 2022; Singh et al., 2022; Spasova et al., 2021). A huge amount of image data needs to be stored and transferred via communication lines. A common way to solve this problem is to apply some image compression algorithm.

Image compression algorithms can be divided into two groups: lossy and lossless ones (Prasanna et al., 2021; Manga et al., 2021; Ponomarenko et al., 2005). The lossy compression results in a loss of data and, in lossless compression, the decompressed data are exactly the same as the corresponding original data. In other words: lossless compression preserves all information contained in data, but the compression ratio (CR) attained for the used methods can be not appropriate in practice (Ponomarenko et al., 2005). Then, lossy compression occurs to be preferable since CR can be varied and controlled. Meanwhile, CR increasing usually leads to worse quality of a compressed image and a reasonable compromise has to be found in each particular case. Priority of requirements depend upon an application at hand (Zabala et al., 2006).

In this paper, we concentrate on lossy compression of color and three-channel remote sensing images using BPG (Better Portable Graphics) coder. The problem of finding a compromise has been partly considered in our previous studies (Kovalenko et al., 2022) with application to single-channel (grayscale) image compression. Since most modern images are color or multichannel, in this article we would like to focus on lossy compression of such images supposing that they are noisy. Recall that there are different chroma subsampling modes in BPG coder and our goal is to study how these modes affect the coder performance in different senses. In particular, we analyze the coder performance by estimating the compression ratio as well as using visual quality metrics such as PSNR-HA (Ponomarenko et al., 2011) and MDSI (Ziaei Nafchi et al., 2016).

2. Used methods and criteria

2.1. Image and noise properties

It is a well-known fact that compression characteristics depend on image properties sufficiently. Depending on image complexity, such characteristic as compression ratio or values of quality metrics can vary in very wide limits. Complexity of image can be described in many different ways: by a percentage of pixels that belong to image homogeneous regions or entropy for noise-free case and so on (Kovalenko et al., 2022). In this paper, we use two RS images from SIPI Image Database which has long been used for performance estimation of different coders and filters. We also employ three standard color test images.

In fact, in practice, it is almost impossible to get noise free images (except computer generated ones). Noise can appear in images due of many factors; it can be visible or invisible (Chatterjee et al., 2010). Visible noise has an essential impact on image quality and

E-mail: b.kovalenko@khai.edu

classification accuracy. That is why, we concentrate here on the case of visible noise typical for images acquired in bad illumination conditions.

In this work, we consider additive white Gaussian noise (AWGN) which is known to be the simplest noise model. We suppose that in each image component one has

$$I_{ij}^{noisy} = I_{ij}^{true} + n_{ij}, \quad (1)$$

where I_{ij}^{true}, $i = 1,\ldots,I_{Im}$, $j = 1,\ldots,J_{Im}$ is the true or noise-free image, n_{ij} denotes AWGN in the ij-th pixel, I_{Im} and J_{Im} define the considered image size. Noise in image components is supposed uncorrelated.

2.2. Used coder

This paper focuses on Better Portable Graphics encoder which is a quite novel image compression method that aims to replace the old JPEG format due to considerably better performance in the sense of higher quality and/or lower size of compressed data. The BPG coder is based on the High Efficiency Video Coding (HEVC) technique, which was proposed by Fabrice Bellard in the form of the open-source code (Bellard, 2018). Overall, BPG has the following attributes:

- lossless and lossy compression;
- it is supported by most Web browsers;
- it is open-source, royalty-free, and patent-free;
- provides native support of 8 to 14 bits per channel typical for most applications;
- supports the same chroma formats as JPEG (grayscale, YCbCr 4:2:0, 4:2:2, 4:4:4) to reduce the losses during the conversion. An alpha channel is supported. The RGB, YCgCo and CMYK color spaces are also supported.

As many other coders, Better Portable Graphics also has compression controlling parameter (CCP). The parameter Q, used internally to control the compression ratio and image quality in the BPG encoder, serves as the CCP where Q can vary in the range of 1 to 51. A larger Q results in a higher CR corresponding to a lower visual quality (in the case of compressing noise-free images).

2.3. Chroma subsampling and metrics

As it has been mentioned above, the BPG coder supports the YCbCr 4:2:0, 4:2:2, 4:4:4 chroma formats. YCbCr is used for digital encoding of color information suited for video and still-image compression and transmission. The YCbCr color space conversion strongly decorrelates the color channels, so they can be coded independently without loss of efficiency (Doutre et al., 2007). Conversion of images in RGB and other color spaces to YCbCr is not a problem.

Since the human visual system is less sensitive to chrominance than luminance, the Cb and Cr channels can be down-sampled by a factor of two in both the horizontal and vertical directions almost without loss of perceived image quality. This down sampling produces the YCbCr 4:2:0 sampling, where there are four Y samples for every Cb and Cr sample. The 4:2:2 scheme of subsampling means that Cb and Cr components are each subsampled by a factor of 2 horizontally (Dumic et al., 2009). Here we take into account the fact that color and three-channel remote sensing images are often subject to visualization and analysis by experts.

To estimate and compare different modes of chroma subsampling in BPG-based lossy compression, we use several metrics. First of all, its compression ratio (CR) is commonly used to estimate efficiency of compression by different coders. Compression ratio is estimated by comparing the size of two images: initial images and ones after compression.

As mentioned above, lossy compression parameters are always a compromise between compression ratio and image quality. To estimate image quality, we use PSNR-HA (modification of peak signal-to-noise ratio taking into account human vision system (HVS)) and MDSI (Mean Deviation Similarity Index).

PSNR-HA is defined as

$$\mathrm{PSNR-HA} = 10\log_{10}\left(\frac{255^2}{\mathrm{MSE-HA}}\right), \quad (2)$$

where MSE-HA is determined as a sum of MSE-HVS-M values in blocks. The approach to calculating them takes into account the masking effect and less sensitivity of a human eye to distortions in high spatial frequencies than distortions in low spatial frequencies (Ponomarenko et al., 2011). Some other peculiarities of HVS are also incorporated. PSNR-HA is expressed in dB and its larger values relate to a better visual quality. PSNR-HA is considered to be one of the most adequate HVS metrics.

Currently, a commonly accepted perfect metric does not exist, and, in our case, to ensure of truthfulness of results, we use another HVS metric – MDSI – that produces effective and reliable full reference image quality assessment (IQA) model based on the gradient and chromaticity similarities. The gradient similarity is used to measure local structural distortions. In a complementary way, a chromaticity similarity is used to measure color distortions. The proposed metric also employs a deviation pooling to compute the quality score from the two similarity maps (Ziaei Nafchi et al., 2016). Recall here that, in opposite to PSNR-HA, smaller values of MDSI relate to a better visual quality.

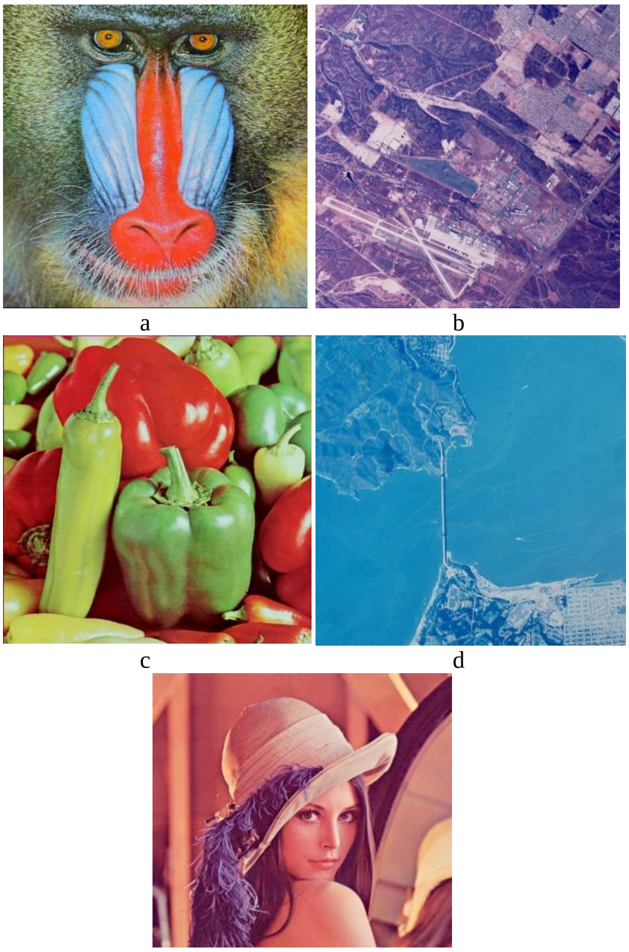
3. Result of using different chroma subsampling modes

3.1. PSNR-HA and MDSI

For our purposes of making general conclusions, we pick images (both traditional color images and remote sensing ones) for three levels of complexity: complex (Fig. 1, a, b), simple (Fig. 1, c, d) and medium (Fig. 1, e) structure images. Images that were used in our research are presented below (Fig. 1).

In simulations, to each RGB original image of size 512x512 pixels we added noise with zero mean and variances equal to 25, 64, 100 and 196 to consider the cases from almost invisible to annoying noise. After this, compression with different chroma subsampling mode of BPG has been applied.

A starting point of our analysis is noise variance equal to 100; in this case, distortions caused by the noise are easy to notice. Since we artificially add noise, we can calculate metrics in two ways: 1) for the noisy and compressed images; 2) for the compressed and noise-free images. In our study, we are more interested in the second option.



a b c d e

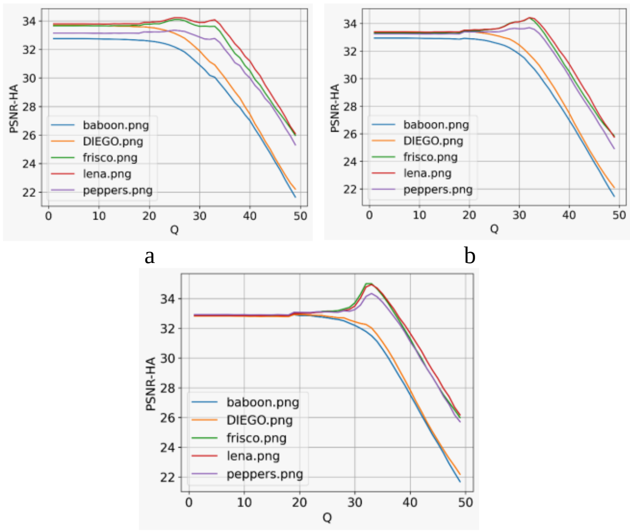
Fig. 1. Used test images Baboon (a), Diego (b), Peppers (c), Frisco (d) and Lenna (c)

Thus, let us "compare" the noise-free image and the corresponding compressed one using the PSNR-HA metric applied to color images previously converted into YCbCr color space. The obtained dependencies (rate-distortion curves) are presented in Fig. 2.

For the 4:2:0 mode (Fig. 2, a), for the complex structure images, the parameter Q increasing results in a flat region with further monotonous PSNR-HA decreasing for Q equal to 22 and larger. For middle and simple structure images, the situation is different in the way that the point of decreasing start is shifted to larger values of Q, also local maximums can be observed.

Switching to the mode 4:2:2 (Fig. 2, b), similar results are observed except that the starting point of quality decreasing is shifted to slightly larger values for complex structure images. For other images, the area of quality increasing appears before reaching the corresponding local maximum. This point is called Optimal Operational Point (OOP), we have already done some research on this theme and proved possibility of OOP occurrence for grayscale noisy images compressed by BPG for several metrics including visual quality ones (Kovalenko et al., 2022).

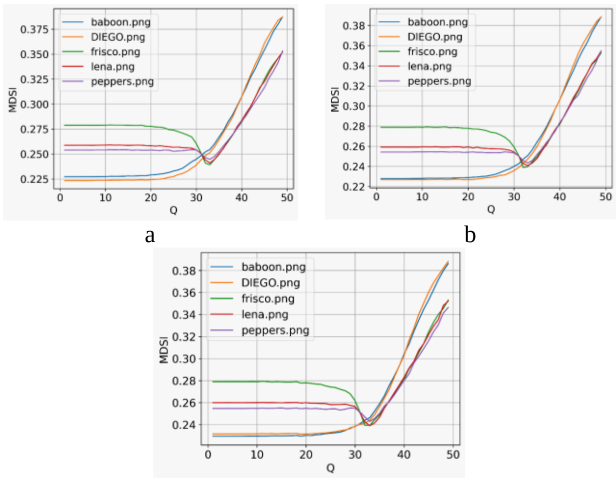
The last possible mode is 4:4:4 (Fig. 2, c). This time, OOP for simple and middle structure images is more obvious, the starting point of quality decreasing for complex structure images is almost in the same place (for the same Q) as OOP for simple structure images (Q = 32).



a b c

Fig. 2. Dependencies of PSNR-HA calculated between the compressed noisy images with noise variance equal to 100 and noise-free images for chroma subsampling modes: 4:2:0 (a), 4:2:2 (b), 4:4:4 (c)

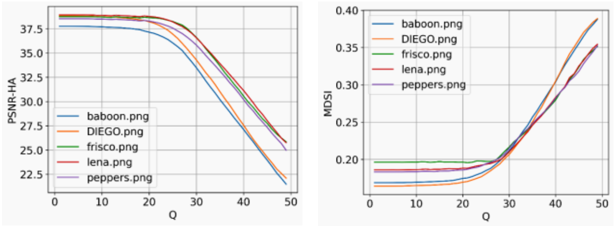
The conclusions given above based on analysis of dependences for the PSNR-HA metric are supported by analysis results for the MDSI metric (Fig. 3). Again, there is no OOP for complex structure images and OOPs exist for simple and middle structure images where for all those images the quality improvement in OOP can be noticeable in all modes of chroma subsampling. The complex structure images are compressed with better quality for the mode 4:4:4 for Q slightly smaller than $Q_{OOP} = 32$ for middle and simple structure images.



a b c

Fig. 3. Dependencies of MDSI calculated between the compressed noisy images with variance equal to 100 and noise-free images for the chroma subsampling modes: 4:2:0 (a), 4:2:2 (b), 4:4:4 (c)

Additional analysis using different noise variances shows that for small values of the noise variance (25) all images demonstrate similar behavior of rate-distortion curves: the dependencies of PSNR-HA for all images are monotonously decreasing (Fig. 4, a) whilst for the metric MDSI they are all monotonously increasing (Fig. 4, b). Fig. 4 shows the data for the mode 4:2:2, but the aforementioned properties are the same for other subsampling modes as well.



a b

Fig. 4. Dependencies of PSNR-HA (a) and MDSI (b) calculated between the compressed noisy images with variance equal to 25 and noise-free images for the 4:2:2 chroma subsampling mode

Another tendency that remains the same for all BPG subsampling modes is the increase of the Q parameter in possible OOP as the noise intensity increases. Figure 5 represents the case of two values of variance of additive white Gaussian noise equal to 64 and 196, respectively.

According to the obtained dependencies, when noise variance is 64 (Fig. 5, a), OOP is present for simple and middle structure images in Q equal to 31. For complex images, OOP is not observed.

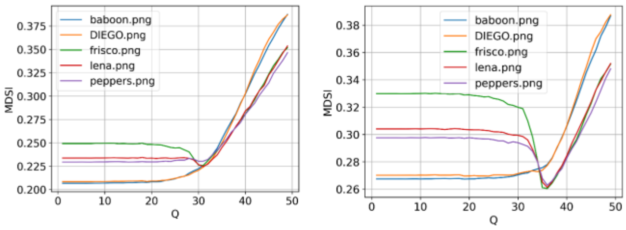
Observations for more intensive noise, i.e. for AWGN variance equal to 196 (Fig. 5, b), shows that this time the OOP has shifted to Q equal to 36 for the same set of images.

For the metric PSNR-HA, similar tendencies are observed and analysis of the obtained plots shows shifting of the optimal operation point towards larger values in the case of increased intensity of the noise.

Summarize the obtained results, the next recommendations can be given. First, it is desirable to use Q in the range from 31 to 37 depending a noise variance, for more accurate selection of the operating point it is recommended to use the following formula:

$$Q_{OOP} = 12.9 + 20 \cdot \lg(\sigma). \quad (3)$$

According to the metric values in OOP, the mode 4:4:4 seems preferable. Based on this statement, the second recommendation can be obtained: to get the maximum benefit from the operating point in the sense of image quality it is desirable to use the 4:4:4 mode of chroma subsampling.



a b

Fig. 5. Dependencies of MDSI calculated between the compressed noise images with variances equal to 64 (a) or 196 (b) and the corresponding noise-free images for the 4:4:4 chroma subsampling mode

3.2. Compression ratio

Another goal of this paper is to provide information about the impact of different chroma subsampling on compression ratio.

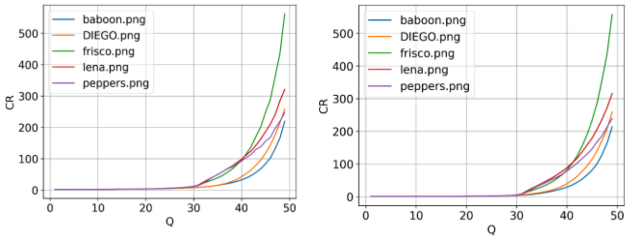
Five test images with initial file size of 768 kb were corrupted by AWGN and compressed by the BPG coder with initial value of compression control parameter equal to 1. The results for Q = 1 are presented in Table 1. AWGN variance equals to 100.

Table 1. Initial compression ratios for different chroma sampling modes for Q = 1 (almost lossless compression)

Image name	Q	Compression rate		
		4:4:4	4:2:2	4:2:0
baboon	1	1,158966	1,766764	2,321141
DIEGO	1	1,208332	1,834994	2,383039
frisco	1	1,288324	2,013722	2,681773
lena	1	1,296848	2,03216	2,718279
peppers	1	1,279801	2,007273	2,689392

It is clear from data in this Table that the 4:2:0 mode outperforms the 4:4:4 and 4:2:2 modes in the sense of larger CR values, the mode 4:2:2 occupies the intermediate place between the modes 4:4:4 and 4:2:0.

The next step is compare two dependencies of CR on Q for the modes 4:2:0 (Fig. 6, a) and 4:4:4 (Fig. 6, b). The dependencies behave in similar manner but the curves for the 4:2:0 mode start it rapidly increase “earlier”, for Q about 25.



a b

Fig. 6. Dependencies of compression ratio for noisy images with variances equal to 100 for the 4:2:0 (a) and 4:4:4 (b) chroma subsampling modes

A special attention deserves the area of optimal operation point and the values of CR in its neighborhood. Table 2 contains information (CR values) for all three modes of the BPG coder.

Table 2. Compression ratio in OOP for different chroma sampling modes

Image name	Q	Compression rate		
		4:4:4	4:2:2	4:2:0
baboon	33	8,775278	12,59846	13,8833
DIEGO	33	12,10651	14,08063	14,68142
frisco	33	32,75609	42,07843	46,19015
lena	33	42,25247	51,71074	55,60385
peppers	33	35,29761	45,26773	48,78874

The first observation is that, in the sense of CR, the mode 4:2:2 outperforms the mode 4:4:4. In turn, the mode 4:2:0 outperforms the mode 4:2:2 but this time the difference no so large as for small Q. Based on this observation, it better to use 4:2:2 in OOP compression scenario due in two reasons. First, the OOP is more obvious than for the mode 4:2:0 and almost the same as in the mode 4:4:4. Second, CR is almost the same as in the 4:2:0 mode.

Additional attention deserves dependencies of visual quality metric to compression ratio. The results of comparison are presented in Figure 7.

According to Figure 7, the plots’ analysis shows that similar compression ratio for different subsampling modes bring different quality of the image in the range of compression ratios from 1 to 250 (in case of simple structure images). In other words, for the same compression ratio 4:4:4 mode can offer improvements in terms of quality compare to 4:2:2 and also for 4:2:0 in case of simple structure images. For complex structure images, this effect is not observed. Moreover, PSNR-HA can be even slightly lager for small CR.

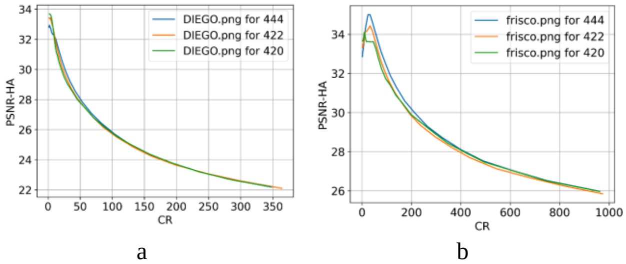


Fig. 7. Comparison of dependencies of PSNR-HA on CR for the test images: DIEGO (a) and frisco (b) corrupted by AWGN with variances equal to 100

4. Conclusions

The task of lossy compression of noisy images using different modes of chroma subsampling of images corrupted by AWGN by the BPG coder is considered. The use of different modes brings specific features. In particular, the obtained results demonstrate that for simple, and middle structure images the OOP is observed with a larger probability and it might exist for both considered metrics PSNR-HA and MDSI. In the case of using the mode 4:2:0, OOP takes place less often, several local maxima of PSNR-HA and minima of MDSI can be observed. For the 4:2:2 mode, OOP looks more obvious as well as for the 4:4:4 mode. For complex structure images, better quality is provided by the 4:4:4 mode if CR is large. This allows using larger values of Q that leads to a larger CR.

It also demonstrated that OOP existence and its position depend on noise variance. For low intensive noise, the probability of OOP existence is small. However, with increasing the noise intensity, the probability also increases but the OOP position shifts to larger values of Q. That statement is valid for all chroma subsampling modes of the BPG coder.

From compression ratio perspective, there are several important observations. First of all, behavior for all modes are the same. It means that with increasing of the parameter Q the compression ratio slowly increases till Q is equal to 25–30, after this the CR starts to rapidly increase. The final observation is that if the three subsampling modes are directly compared in terms of dependencies of image quality on compression ratio, then the 4:4:4 chroma subsampling mode can, in general, provide better quality at the same compression ratio.

References

Bellard, F. (2018). BPG Image format. Retrieved from: https://bellard.org/bpg/.

Chatterjee, P., Milanfar, P. (2010). Is Denoising Dead? *In IEEE Transactions on Image Processing*, 19(4), 895–911. DOI: 10.1109/TIP.2009.2037087.

Doutre, C., Nasiopoulos, P., Plataniotis, K. N. (2007). A Fast Demosaicking Method Directly Producing YCbCr 4:2:0 Output. *IEEE Transactions on Consumer Electronics*, 53(2), 499–505. DOI: 10.1109/TCE.2007.381721.

Dumic, E., Mustra, M., Grgic, S., Gvozden, G. (2009). Image quality of 4:2:2 and 4:2:0 chroma subsampling formats. 2009 *International Symposiu ELMAR*, 19–24.

Kovalenko, B., Lukin, V., Kryvenko, S., Naumenko, V., Vozel B. (2022). BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters. *Appl. Sci.*, 12(15) (7555). DOI: 10.3390/app12157555.

Kovalenko, B., Lukin, V., Kryvenko, S., Naumenko, V., Vozel, B. (2022). Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images. *Ukrainian Journal of Remote Sensing*, 9(2), 4–12. DOI: 10.36023/ujrs.2022.9.2.212.

Manga, I., Garba, E. J., Ahmadu, A. S. (2021). Lossless Image Compression Schemes: A Review. *Journal of Scientific Research and Reports*, 27(6), 14–22. DOI: 10.9734/jsrr/2021/v27i630398.

Nan, S., Feng, X., Wu, Y., Zhang, H. (2022). Remote sensing image compression and encryption based on block compressive sensing and 2D-LCCCM. *Springer*, 108, 2705–2729. DOI: 10.1007/s11071-022-07335-4.

Ponomarenko, N., Ieremeiev, O., Lukin, V., Egiazarian, K., Carli, M. (2011). Modified Image Visual Quality Metrics for Contrast Change and Mean Shift Accounting. Proceedings of CADSM, 305–311.

Ponomarenko, N., Lukin, V., Zriakhov, M., Egiazarian, K. (2005). Lossy compression of images with additive noise. *Proceedings of International Conference on Advanced Concepts for Intelligent Vision Systems*, 3708, 381–386.

Prasanna, Y. L., Tarakaram, Y., Mounika, Y., Subramani, R. (2021). Comparison of Different Lossy Image Compression Techniques. *2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 1–7. DOI: 10.1109/ICSES52305.2021.9633800.

Singh, B. K., Sinha, G. R. (2022). Medical Image Processing. *In book: Machine Learning in Healthcare*. DOI: 10.1201/9781003097808-4.

Spasova, G., Boyachev, I. (2022). A Method of Color Images Compression. *2021 International Conference on Biomedical Innovations and Applications (BIA)*, 111–114. DOI: 10.1109/BIA52594.2022.9831403.

Zabala, A., Pons, X., Diaz-Delgado, R., Garcia, F., Auli-Llinas F., & Serra-Sagrista, J. (2006). Effects of JPEG and JPEG2000 Lossy Compression on Remote Sensing Image Classification for Mapping Crops and Forest Areas. *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 790–793. DOI: 10.1109/IGARSS.2006.203.

Ziaei Nafchi, H., Shahkolaei, A., Hedjam, R., Cheriet, M. (2016). Mean Deviation Similarity Index: Efficient and Reliable Full-Reference Image Quality Evaluator. *IEEE Access*, 4, 5579–5590. DOI: 10.1109/ACCESS.2016.2604042

ВИКОРИСТАННЯ РІЗНИХ РЕЖИМІВ КОЛІРНОЇ СУБДИСКРЕТИЗАЦІЇ ПРИ СТИСНЕННІ ЗОБРАЖЕННЯ ЗА ДОПОМОГОЮ BPG КОДЕРА

Б. В. Коваленко, В. В. Лукін

Національний аерокосмічний університет, кафедра інформаційно-комунікаційних технологій, Харків 61070, Україна

Кодер BPG (better portable graphics) – це новий підхід, який спрямований на заміну звичайних стандартів стиснення, таких як JPEG, JPEG2000 тощо. Тому такий кодер потребує детального аналізу його основних характеристик з погляду візуальної якості та ступеня стиснення. BPG-кодер може використовувати різні режими кольорової субдискретизації для кольорових і триканальних зображень, тому їх варто проаналізувати та порівняти. На практиці зображення, які потрібно стиснути, часто містять шум. Тоді стиснення з втратами таких зображень має специфічний ефект фільтрації шуму. Зокрема, оптимальна робоча точка (OPT) може існувати, коли якість стисненого зображення ближче до відповідного безшумного (справжнього) зображення, ніж якість стисненого (вихідного, зашумленого) зображення за певним критерієм (метрикою). Також необхідно проаналізувати ефективність кодера з погляду ступеня стиснення. У цій статті автори звертають увагу на вплив різних режимів субдискретизації кольоровості на якість зображення та ступінь стиснення. За результатами моделювання, отриманими для набору кольорових зображень, рекомендовано оптимальні способи стиснення.

Ключові слова: кольорове зображення, стиснення зображення з втратами, колірна субдискретизація, BPG-кодер, візуальна якість, YCbCr.

Рукопис статті отримано 27.08.2022