

National Aerospace University. M.E. Zhukovsky
«Kharkov Aviation Institute »
Ministry of Education and Science of Ukraine

National Aerospace University. M.E. Zhukovsky
«Kharkov Aviation Institute »
Ministry of Education and Science of Ukraine

Qualifying scientific
work on the rights of the manuscript

LI FANGFANG

UDC 004.627

DISSERTATION

**DESIGN AND ANALYSIS OF EFFICIENT METHODS FOR PROVIDING A
DESIRED QUALITY IN IMAGE LOSSY COMPRESSION**

172 Telecommunications and radio engineering

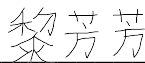
(шифр і назва спеціальності)

Electronics and telecommunications

(галузь знань)

Applied for the degree of Doctor of Philosophy

The dissertation contains the results of own research. The use of ideas, results,
and texts of other authors has references to the relevant sources



F. Li

(signature)

Supervisor

Lukin Volodymyr Vasylyovych, Doctor of Technical Sciences, Professor.

Kharkiv – 2023

ABSTRACT

Li Fangfang. Design and analysis of efficient methods for providing a desired quality in lossy image compression. – Qualifying scientific work on the rights of the manuscript.

The thesis for a degree of Doctor of Philosophy (PhD) in the field of knowledge 17 Electronics and Telecommunications in specialty 172 Telecommunications and radio engineering. – National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, 2022.

Lossy image compression has been developed to be an essential tool in the past several decades. This is due to the rapid development and broad application of imaging technology that has resulted in a sharp increase in the number of images and an increase in images size. Compared to lossless compression, lossy compression can achieve a higher compression ratio. However, inevitable distortions are introduced, which determine the visual quality of decompressed images. Consequently, the visual quality needs to be evaluated and distortions need to be controlled in practical applications.

The thesis is devoted to solving the scientific and applied problem of increasing the distortion controlling efficiency in lossy compression with controlling parameter determination. **The object of the study is** the controlling of image distortions introduced by lossy compression. **The subject of the study is** the method of providing a desired visual quality in lossy compression.

An analytical review of existing popular lossy compression coder’s performance in terms of compression ratio and several visual quality metrics, purpose and requirements of visual quality controlling in typical applications, as well as types and characteristics of different methods of distortion control are performed in the thesis. It is clear that the visual quality of decompressed images impacts further image processing and should be controlled carefully. The trade-off between distortion and compression ratio is often looked for in lossy compression, and compression control parameters are calculated according to this purpose. Based on the various requirements of users, the perceptual lossless effect can be demanded, and more flexible desired visual qualities could be set by users. It has been shown that the visual quality of decompressed images depends not only

on the compression control parameter but also on the complexity of an original image, and the encoder adopted. The main problem of existing image lossy compression methods, which take account into visual quality, is either unsatisfactory accuracy with an acceptable time computation or appropriate accuracy but inappropriate or uncertain time efficiency.

In the thesis, the actual scientific and applied task of designing lossy compression method for providing a desired visual quality, which considers control accuracy of different coders combined with analysis of the method efficiency is set and solved. The method of predicting visual quality of decompressed image, the two-step lossy compression method and the improved adaptive method which groups images based on their complexity, the method of multi-channel image lossy compression have been designed and analyzed with several visual quality metrics.

The thesis aims to design efficient method of lossy compression to provide a desired visual quality, and analyze the accuracy of distortion control, in particular the lossy compression method for the multi-channel image in remote sensing, and to analyze the effect of visual quality on the accuracy of classification.

The scientific results are:

1) For the first time, a two-step method is proposed to provide the required visual quality in gray-scale lossy compression. The average rate-distortion curve obtained offline is used to calculate the initial parameters, and the visual quality in the first-step compression is corrected by feedback. The results show that, in terms of human visual system (HVS)-based metrics, the accuracy of the second-step compression of general images is good enough, and the residual error is acceptable.

2) The method of predicting the visual quality of decompressed images for given parameters has been improved and extended to lossy compression encoders based on DWT.

3) The primary two-step method has been modified to enhance its robustness, and two methods are used. The first is to correct the constraints of the scheme to reduce the errors for high-texture images; the second is to pre-classify images according to their complexity and then use adaptively the proper average rate-distortion curve.

4) The two-step method has been extended to color and three-channel images, and its application in remote sensing has been discussed. The results show that the classification accuracy of compressed images is roughly the same as for the original (uncompressed) data if there are no visually noticeable distortions. In addition, it is possible to control the quality of 3D compression. The compression ratio it produces is twice that of wise-component compression, and the probability of correct classification is slightly higher.

The study has been carried out for DCT-based coders (including AGU and ADCTC), DWT-based SPIHT coder and novel BPG coder, and the universality of the method for different images has been discussed. It is shown that the method works well for various coders considered. Specifically, the results of DCT-based coders and BPG are slightly better than the DWT-based SPIHT coder. Good results can be obtained for normal images, including highly textured images, and the method application for strange images, which has a non-monotonic rate-distortion curve, should be cautious.

The results of the work have been used in research studies of the National Aerospace university named after M. E. Zhukovsky “Kharkiv Aviation Institute” and Nanchang Hangkong University.

According to the thesis materials, 17 papers were published. 3 of them are articles in journals of Ukraine, 2 articles have been published in foreign country periodical journal (Q1, Switzerland), among them 4 articles are included in the international scientific metric database SCOPUS, 12 – publications are in conference proceedings (12 in international conferences included in the international scientific metric database SCOPUS).

Key words: image lossy compression, visual quality metric, average rate-distortion curve, two-step method, grayscale, three-channel image, complexity.

List of publications of the applicant on the topic of the dissertation:

Articles in scientific periodicals:

1. F. Li, V. Lukin, O. Ieremeiev, and K. Okarma, “Quality control for the BPG lossy compression of three-channel remote sensing images,” *Remote Sensing*, vol.14, no. 8, pp.1824, 2022, doi: 10.3390/rs14081824.

2. F. Li, S. Krivenko, and V. Lukin, "Two-step providing of desired quality in lossy image compression by SPIHT," *Radioelectronic and computer systems*, vol. 2, no. 2020, pp. 22-32,2020, doi: 10.32620/reks.2020.2.02.
3. F. Li, "Adaptive two-step method for providing the desired visual quality for SPIHT," *Radioelectronic and computer systems*, vol.1, no. 2022, pp. 195-205,2022, doi: 10.32620/reks.2022.1.15.
4. V. Lukin, I. Vasilyeva, S. Krivenko, F. Li, S. Abramov, O. Rubel, B. Vozel, K. Chehdi, and K. Egiazarian, "Lossy Compression of Multichannel Remote Sensing Images with Quality Control," *Remote Sensing*, vol. 12, no. 22: 3840. doi: 10.3390/rs12223840.
5. F. Li, S. Krivenko, and V. Lukin, "Analysis of two-step approach for compressing texture images with desired quality," *Aerospace technic and technology*, vol. 1, no. 2020, pp. 50-58, 2020,doi:10.32620/aktt.2020.1.08.

Scientific works certifying the approbation of the dissertation materials:

6. F. Li, S. Krivenko and V. Lukin, "A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality," *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine, 2020, pp. 307-312, doi: 10.1109/DESSERT50317.2020.9125000.
7. S. Krivenko, F. Li, V. Lukin, B. Vozel and O. Krylova, "Prediction of Visual Quality Metrics in Lossy Image Compression," *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine,2020, pp. 478-483, doi: 10.1109/ELNANO50318.2020.9088819.
8. F. Li, S. Krivenko and V. Lukin, "A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT," In: Nechyporuk M., Pavlikov V., Kritskiy D. (eds) *Integrated Computer Technologies in Mechanical Engineering- ICTM 2020. Lecture Notes in Networks and Systems*, vol 188. Springer, Cham. doi:10.1007/978-3-030-66717-7_2.
9. F. Li, S. Krivenko and V. Lukin, "A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression," *2020 IEEE 15th International*

Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Slavske, Ukraine, 2020, pp. 502-506, doi: 10.1109/TCSET49122.2020.235483.

10. F. Li, S. Krivenko and V. Lukin, “An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality,” *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine, 2020, pp. 13-17, doi: 10.1109/ATIT50783.2020.9349289.

11. F. Li, S. Krivenko and V. Lukin, “Adaptive two-step procedure of providing desired visual quality of compressed image,” *2020 4th International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, Xiamen, China, 2020, pp. 407-414, doi:10.1145/3443467.3443791.

12. G. Proskura, I. Vasilyeva, F. Li and V. Lukin, “Classification of Compressed Multichannel Images and Its Improvement,” *2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA)*, Bratislava, Slovakia, 2020, pp. 1-6, doi: 10.1109/RADIOELEKTRONIKA49387.2020.9092371.

13. I. Vasilyeva, F. Li, S. Abramov, V. Lukin, B. Vozel, and K. Chehdi, “Lossy compression of three-channel remote sensing images with controllable quality,” *Proc. SPIE 11862, Image and Signal Processing for Remote Sensing XXVII*, pp. 118620R, doi: 10.1117/12.2599902.

14. F. Li, V. Lukin, K. Okarma and Y. Fu, “Providing a Desired Quality of BPG Compressed Images for FSIM Metric,” *2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine, 2021, pp. 10-14, doi: 10.1109/ATIT54053.2021.9678522.

15. F. Li, S. Krivenko and V. Lukin, “An Automatic Optimization Method for BPG Compression Based on Visual Perception,” *2021 Mathematical Modeling and Simulation of Systems, Selected Papers of 16th International Scientific-practical Conference (MODS)*. Chernihiv, Ukraine, 2021, pp. 213-225. doi:10.1007/978-3-030-89902-8_17.

16. F. Li, V. Lukin, K. Okarma, Y. Fu, and Duan J. “Intelligent lossy compression method of providing a desired visual quality for images of different

complexity,” *Applied Mathematics, Modeling and Computer Simulation (AMMCS)*, Wuhan, China, 2021, pp.500-505. doi:10.3233/ATDE220050.

17. F. Li, S. Krivenko and V. Lukin, “A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images,” *2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, Odesa, Ukraine, 2021, pp. 16-21, doi: 10.1109/UkrMiCo52950.2021.9716684.

CONTENTS

ABSTRACT	2
LIST OF CONDITIONAL ABBREVIATIONS	12
INTRODUCTION	13
CHAPTER 1. PROVIDING A DESIRED QUALITY IN LOSSY COMPRESSION	22
1.1 Application and technologies of digital imaging	22
1.1.1 Applications of digital images	22
1.1.2 Related technologies of digital image processing	24
1.2 Image compression	26
1.2.1 A review of image compression	26
1.2.2 Lossless compression	27
1.2.3 Lossy compression	29
1.3 Existing lossy compression coders	31
1.3.1 DCT-based coders	32
1.3.2 DWT-based coders	35
1.3.3 Learning-based coders	38
1.3.4 Other coders	39
1.4 Requirements for lossy compression techniques	41
1.4.1 The requirements of human visual perception	41
1.4.2 The requirements to image processing algorithm	42
1.4.3 Image quality metrics	44
1.5 Analysis of methods of image quality controlling in lossy compression	50
1.5.1 Visually lossless lossy compression	51
1.5.2 JND point	52

1.5.3 Iterative method	53
1.5.4 Prediction method	54
1.6 Conclusions and formulation of research objectives	54
CHAPTER 2. RESEARCH AND DEVELOPMENT OF PRIMARY METHOD OF PROVIDING THE DESIRED VISUAL QUALITY IN LOSSY COMPRESSION	57
2.1 Selection of appropriate visual quality metrics	57
2.2 Comparison of various lossy compression coders	59
2.3 Prediction of visual quality in lossy compression.....	62
2.3.1 Prediction of visual quality in lossy compression for medical images.....	63
2.3.2 Prediction of visual quality in lossy compression by SPIHT	68
2.4 The primary two-step compression method for providing the desired visual quality	75
2.4.1 The primary two-step approach to providing a desired visual quality in AGU lossy compression	75
2.4.2 A two-step approach to providing a desired visual quality in BPG lossy compression.....	80
2.5 Conclusions	85
CHAPTER 3. IMPROVEMENTS OF TWO-STEP COMPRESSION METHOD FOR GRAY-SCALE IMAGE.....	87
3.1 Modification of correction scheme of CCP in the second step.....	87
3.1.1 Analysis of two-step approach for compression of texture images with desired quality.....	88
3.1.2 A two-step procedure for image lossy compression by ADCTC with a desired quality.....	96
3.1.3 Two-step providing of desired quality in lossy image compression by SPIHT	102
3.2 Modification in curve model.....	107

3.2.1 Adaptive two-step procedure of providing desired visual quality of compressed image	108
3.2.2 Adaptive two-step method for providing the desired visual quality for SPIHT	117
3.2.3 Intelligent lossy compression method of providing a desired visual quality for images of different complexity	123
3.3 Hybrid method of correction scheme of CCP in second step	127
3.4 Conclusions	136
CHAPTER 4.TWO-STEP COMPRESSION METHOD FOR COLOR AND THREE-CHANNEL IMAGES AND ITS USE IN REMOTE SENSING	137
4.1 Two-step method for color and three-channel images.....	137
4.1.1 An automatic optimization method for BPG compression based on visual perception	137
4.1.2 Providing a desired quality of BPG compressed images for FSIM metric	144
4.2 Quality control for the BPG lossy compression of three-channel remote sensing images.....	150
4.2.1 Metrics for the Assessment of the Visual Quality of Three-Channel RS Images	151
4.2.2 The two-step method for lossy compression	154
4.2.3 Implementation of the Two-Step Method for the BPG Coder	156
4.2.4 The experimental results	159
4.2.5 Discussion	162
4.3 Analysis of the effect of controlling visual quality on remote sensing classification	164
4.3.1 Classification of compressed multichannel images and its improvement	164
4.3.2 Lossy compression of multichannel remote sensing images with quality control	170

4.3.3 Lossy compression of the three-channel remote sensing images with controllable quality	183
4.4 Conclusion.....	193
CONCLUSION	195
APPENDIX A.LIST OF APPLICANT PUBLICATIONS.....	220
APPENDIX B.IMPLEMENTATION ACTS	223

LIST OF CONDITIONAL ABBREVIATIONS

CR- compression ratio;

CCP- compression control parameter;

RS- remote sensing;

QS- quantization step;

DCT - discrete cosine transformation;

DWT - discrete Wavelet Transform;

SPIHT - Set partitioning in hierarchical trees

JPEG - Joint Photographic Experts Group

BPG- Better Portable Graphics

PSNR - Peak Signal to Noise Ratio;

MSE - Mean Squared Error;

HVS - Human Vision System;

PSNR-HVS-M - PSNR based on HVS Masking effect;

MDSI- mean deviation similarity index

FSIM- feature similarity

JND - Just-Noticeable-Difference;

MLP - multilayer perception;

SROCC - Spearman's rank order correlation coefficient;

KROCC - Kendall's rank order correlation coefficient;

RMSE - root-mean-square error;

IQA - Image quality Assessment;

FR IQA - full-reference IQA;

NR IQA - no-reference (blind) IQA;

RR IQA - reduced-reference IQA.

INTRODUCTION

Rationale for choosing a research topic. In this modern information age, images have been widely used in various fields since they can record richer content than text data forms, e.g., in security[1], intelligent medical[2, 3], remote sensing[4-6], and other technical fields, as well as daily life[7, 8]. Simultaneously, image technology has developed rapidly with the research and improvement of various image sensors, such as optic, ultrasound, and radar. Nevertheless, a general tendency of high technical image devices is to acquire data with higher resolution and/or more frequently.

In the past two decades, the mobile phone camera has developed from the initial 110 thousand pixels to hundreds of millions of pixels now, allowing users to take high-resolution images conveniently[9]. Moreover, combining with social media applications leads to millions of image data being uploaded and downloaded worldwide every second[7, 10]. With mobile Internet development, images and video files have become important social networking and information transmission carriers. The imaging quality and effects of Smartphone cameras have become the competition for Smartphone manufacturers in the mobile phone market. This also occurs in remote sensing systems, which have been developed to have an extremely high spatial resolution of up to 0.25m[11], as well as nanometer-level spectral resolutions[12]. Based on this, remote sensing is one of the critical investment and priority fields of many countries around the world. This is evidenced by the scientific and technical Ukraine's space program for 2018-2022[13], China's High Resolution Earth Observation System[12, 14, 15], and world system observation of the "Global Earth Observation System of Systems"[16].

Images of high definition help a user to make decision or analysis easier and give a viewer a better visual perceptive experience. However, it also results in a larger size and massive volume of images. It runs into typical big data problems when there are difficulties with image processing, transferring, storage, and dissemination[17].

Limitations on a storage volume, communication line bandwidth and available data transferring time lead to necessity to compress images[18]. Consequently, alongside the development of imaging technology, many scientists focus on image compression research,

which has gradually become a hot topic[19-21]. In general, image compression can be divided into two categories, namely lossless compression and lossy compression, respectively. Lossless compression is conducted by removing redundancy in an image, and the image can be restored after compression/decompression without distortion. It is often used in intelligent medical[3], security[1], and other applications where any distortion can result in an improper diagnosis. However, the small compression ratio (CR) is achieved in lossless compression, about 1,2~5[3, 22, 23], which does not meet the expectation of the user in most applications. On the contrary, the lossy compression can easily break the upper limit and achieve a higher CR, with the expense of a certain amount of distortion. For the significant size reduction, lossy compression is widely used in many applications, where distortion is accepted to a certain degree and does not affect further image processing tasks or user perception[24-27].

The goal of the current lossy compression technique is to design advanced compression techniques with higher CR and better visual quality of the decompressed image. However, an inevitable law is that CR increasing leads to worse quality for any image and any lossy compression coder but at different degree. Consequently, two important factors should be considered in lossy compression implementation, CR and visual quality, respectively. An appropriate compromise is needed to be provided for lossy compression in real applications or engineering practices, that the visual quality of the decompressed image does not have negative influence on further image processing or viewing, and the CR is as high as possible for an adopted coder. If the terminal is human, the bad visual quality will cause the difference between decompressed image and the original one easy to distinguish. This means the degree of distortion is noticeable and seriously affects visual perception. Suppose an image needs to be processed by a computer, for example, for object classification and detection in remote sensing. In that case, severe distortion will cause a significant decrease in classification accuracy or misjudgment of target detection. In these two cases, a better visual quality should be provided in lossy compression, even if the cost is a reduction in CR. This principle is also valid for other applications because compression is an intermediate process, and use is the ultimate goal.

Therefore, it is extremely necessary to provide high CR and reliable image quality in lossy compression.

Despite the often quite high performance of existing lossy compression coders, the distortion is unavoidable and affect the quality of the decompressed images. The distortion produced by lossy compression is diverse, which is related to the image compressed and the encoder adopted. Lots of work are devoted to the study of distortion in lossy compression, the popular method is to preset a metric value describing the desired image quality with respect to the original image[24, 27-33]. Scientists conducted intensive research in image quality evaluation, and a variety of excellent metrics have been proposed[33-38]. The distortion level can be characterized by quality metric value; therefore, it can be inferred from this value whether the compression can be reasonably applied to the corresponding application and does not have negative influence.

For the better operation, the compression should be controlled in terms of visual quality determined by the specific applications and, simultaneously, the CR to be maximized in order to obtain both efficiency and reliability. Besides, such a method also needs to balance the accuracy and cost, including time-efficient and computation complexity, in particular, for real-time communication and application platforms with limited resources. As a general rule, there are some parameters in the lossy compression, which can be utilized to control the compression. However, the visual quality of decompressed images depends not only on the compression control parameter (CCP) but also on a coder adopted and an image to be compressed. Additionally, compression efficiency depends on the choice of visual quality metric to evaluate image distortion, ways of image information extraction, CCP calculation, modifications to reduce the error, and so on. Besides, it is necessary to analyze the efficiency of such a method in a particular application, whether the provided quality is accurate enough not to have a negative impact.

In this regard, the current scientific and applied problem, which is solved in this dissertation, is the development of efficient methods of lossy compression, which aim to provide a desired quality for decompressed image to guarantee the distortion introduced by

lossy compression does not influence the further image processing or terminal user's perception.

The object of study is the controlling of images distortions introduced by lossy compression.

The subject of the study - the method of providing a desired visual quality in lossy compression.

The purpose and objectives of the study. The purpose of the dissertation is to design efficient methods of lossy compression to provide a desired visual quality, and analyze the accuracy of distortion control, in particular the lossy compression methods for the multi-channel image in remote sensing, and analyze the effect of visual quality on the accuracy of classification.

Under the goal in the dissertation, the following main tasks are formulated and solved:

- analysis of the requirements to lossy compression techniques;
- study and improvement of the method of predicting the visual quality of decompressed images for lossy compression;
- design and improvement of the two-step method of grayscale lossy compression for providing a desired visual quality;
- extension of the two-step method to color and three-channel RS images;
- analysis of the distortion controlling in the lossy compression for the classification processing of remote sensing three-channel images.

Research methods. To solve the tasks in the dissertation, the following methods were used: the methods of numerical modeling to build a model curve of rate-distortion, mathematical to predict the decompressed image quality corresponding to the CCP, function interpolation to correct the parameter, machine learning to build multilayer neural networks and analyze the effect of image quality error on the probability of remote sensing classification accuracy, probability theory and mathematical statistics to analyze the accuracy of providing a desired quality.

The scientific novelty of the results obtained by the author is as follows:

1) For the first time, a two-step method is proposed to provide the required visual quality in gray-scale lossy compression. The average rate-distortion curve obtained offline

is used to calculate the initial parameters, and the visual quality in the first-step compression is corrected by feedback. The results show that, in terms of human visual system(HVS)-based metrics, the accuracy of the second-step compression of general images is good enough, and the residual error is acceptable.

2) The method of predicting the visual quality of decompressed images for given parameters has been improved and extended to lossy compression encoders based on DWT.

3) The primary two-step method has been modified to enhance its robustness, and two methods are used. The first is to correct the constraints of the scheme to reduce the errors for high-texture images; the second is to pre-classify images according to their complexity and then use adaptively the proper average rate-distortion curve.

4) The two-step method has been extended to color and three-channel images, and its application in remote sensing has been discussed. The results show that the classification accuracy of compressed images is roughly the same as for the original(uncompressed) data if there are no visually noticeable distortions. In addition, it is possible to control the quality of 3D compression. The compression ratio it produces is twice that of wise-component compression, and the probability of correct classification is slightly higher.

Personal contribution of the applicant. All the main results of the dissertation were obtained by the author herself. Article [39] was published without co-authors. The following results belong to the application in the works published in co-authorship.

In [40], the applicant proposed an approach to predict the visual quality value of decompressed image for the coder AGU based on discrete cosine transform(DCT) using statistical information extracted from a limited number of 8×8 pixel blocks. In [41], the applicant exploited the prediction method to the coder SPIHT based on discrete wavelet transform (DWT). In [42, 43], the applicant proposed a two-step method to provide a desired visual quality in lossy compression of gray-scale images, the coders AGU and BPG were adopted, average rate-distortion curve model was obtained off-line and utilized to calculate the appropriate CCP, the classic metric PSNR and HVS-based metrics were employed to characterize the quality of decompressed images. In[44], the applicant presented a simple modification of correction scheme of CCP in the second step in AGU

coder to reduce the errors in terms of PSNR. In [18, 45], the applicant exploited the simple modification scheme for ADCT and SPIHT coders, respectively. In [46], the applicant proposed an optimization scheme to choose the proper average rate-distortion curve model adaptively according to the image complexity, which was determined by the prediction procedure. In [39], this adaptive method was exploited to DWT-based coder SPIHT, and the prediction algorithm was utilized to pre-grouping in terms of images complexity.

In [47], the applicant improved the optimization scheme by adopting entropy to characterize the image complexity, which helps speeding up the algorithm. Besides, it is possible to find “strange” images which produce the largest errors in providing a desired quality of compression. In [48], the applicant chose special basic images as basic image set to obtain the average rate-distortion curve for remote sensing image, and proposed a combination scheme of derivative and tangent to reduce the final error in two-step method. In [49-52], the applicant exploited the method of providing a desired visual quality in lossy compression to color images and three-channel remote sensing images, and adopted an automatic optimization method for BPG coder, 3D approach for AGU coder, employed the metrics PSNR-HMA, MDSI and FSIM to evaluate the decompressed color images quality. In [53], the applicant considered three widely used classifiers, namely, support vector machine, multilayer perception, and logistic regression, studied the effectiveness of multilayer classification on a real satellite image compressed with distortions characterized by PSNR-HVS-M, performed the analysis of classification accuracy for overall probability of correct recognition of classes in the images. In [54], the applicant studied the dependence between the classification accuracy of maximum likelihood and neural network classifiers applied to three-channel test and real-life images and quality of compressed images characterized by standard and visual quality metrics. In [55], the applicant proposed an approach to lossy compression, which is based on providing quality of compressed images not worse than desired according to quality metrics, and investigated the effect of distortion of decompressed image on classification accuracy in remote sensing images.

The list of the applicant's publications on the topic of the dissertation is given in Appendix A.

Approbation of dissertation materials. The main provisions of the dissertation were reported and discussed at all-Ukrainian Scientific and Technical Conference "Integrated Computer Technology in Mechanical Engineering" (2020, Kharkiv, Ukraine), International Conference "Dependable Systems, Services and Technologies" (2020, Kyiv, Ukraine), International Conference "Electronics and Nanotechnology" (2020, Kyiv, Ukraine), International Conference "Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering" (2020, Lviv-Slavske, Ukraine), International Conference "Advanced Trends in Information Theory" (2020, 2021, Kyiv, Ukraine), International Conference "Electronic Information Technology and Computer Engineering" (2020, Xiamen, China), International Conference "Radioelektronika" (2020, Bratislava, Slovakia), International Conference "Image and Signal Processing for Remote Sensing" (2021, Online Only, Spain), International Conference "Mathematical Modeling and Simulation of Systems" (2021, Chernihiv, Ukraine), International Conference "Applied Mathematics, Modeling and Computer Simulation" (2021, Wuhan, China), International Conference "Information and Telecommunication Technologies and Radio Electronics" (2021, Kyiv, Ukraine).

Connection of work with scientific programs, plans, themes, grants. The research presented in this dissertation was conducted at the Department of Information and Communication Technologies named after O.O. Zelensky National Aerospace University. ME Zhukovsky "KHAI" and were reflected in the following reports on research: the Project M/94-2021 (504-22/2021, (state registration No 0120U103595) «Methods of intellectual image and video processing based on visual quality metrics for emerging applications», 2021; the Project Д504-2/2018-Ф (state registration No 0018U003020) «Methods of intellectual computer processing of big data in remote sensing, multimedia and telecommunications», 2020; the Project Д504-2/2021-Ф (state registration No 0121U112176) «Methods of multichannel image processing in mobile systems based on prediction and machine learning», 2021.

The practical significance of the results:

- developed a very fast method for predicting decompressed image quality by calculating information from a certain number of random blocks of an image in compression, the accuracy of which is 0,88-0,95 for different quality metrics according to criterion R^2 (coefficient of determination);

- proposed the universal two-step image compression approach to provide the desired image quality, where the average rate-distortion curves obtained offline are utilized to calculate the initial CCP value, and the quality feedback after the first step compression is used to effectively correct the parameters;

- in the proposed compression method, such quality metrics are applied that they are characterized by the highest values of the Spearman rank correlation with the average opinion of humans, and it is in line with real-life application requirements;

- It is proved that the basic correction scheme of CCP allowed improving the accuracy for different types of images in AGU and BPG coders (reducing variance at least by 10times compared to the initial CCP in terms of HVS-based metrics), which contributes to provide a desired quality without noticeable distortions;

- For complex images (in particular, high-texture images or remote sensing images) and different coders, the optimization correction scheme and the adaptive curve model have been proposed, which allowed solving the over-correction problem (reduce variance by 2-10 times compared to the initial CCP in terms of PSNR and HVS-based metrics), make it possible to guarantee that the correction of CCP is performed in the right way.

- It is proved that the proposed lossy compression method of providing a desired quality also applies to color images and three-channel remote sensing images, and the reasonable ranges of images quality are given in terms of different metrics, which contribute to set the desired image quality.

- Analysis of the effect of the controlling image quality on remote sensing classification, where different classifiers are considered, and two compression method are employed, has proved that classification accuracy of images compressed without visually noticeable distortion is approximately the same as original one (the threshold is about 40-42dB in terms of PSNR-HVS-M).

Publications. The main results of the dissertation are published in 17 works, including: one articles in scientific journals of Ukraine, which are included in international scientific metric databases (Index Copernicus, Google Scholar), two articles in scientific journals of Ukraine, which are included in international scientific metric databases SCOPUS, two articles in a scientific periodical of foreign country (MDPI, Switzerland), which are indexed in SCOPUS (Q1), 12 papers in proceedings of International scientific conferences, the materials of which are included in the scientific metric database SCOPUS.

The structure and scope of the dissertation. The dissertation consists of an abstract, table of contents, list of abbreviations, introduction, four sections, conclusions, a list of used sources and appendices. The full volume of the work is 224 pages of printed text, of which the abstract - by 3 pages, the table of contents - by 4 pages, the list of abbreviations - by 1 page, the main text - by 184 pages, the list of 225 used sources - by 23 pages, appendices - 6 pages. The dissertation contains 78 figures (7 of them on 6 separate pages), 76 tables (30 of them on 10 separate pages).

CHAPTER 1

PROVIDING A DESIRED QUALITY IN LOSSY COMPRESSION

The section considers the main types and characteristics of image compression, the features of popular lossy compression coders, the requirements to lossy compression techniques. In addition, lossy compression methods that take into account visual quality, and approaches to assessing their effectiveness are considered.

1.1 Application and technologies of digital imaging

1.1.1 Applications of digital images

As a visualized and intuitive description of objective objects, the image can convey rich information to humans through vision. It has also become the most crucial source of human visual information. People share their life, study, and work by transmitting multimedia information such as images and videos through the Internet. With the continuous development of the Internet, particularly the fifth mobile communication network 5G, people lifestyles have changed significantly. Various Smartphone applications such as Instagram, Facebook, Webchat, Twitter make people's socializing methods more and more diversified and convenient.

On the other hand, in science and technology, images are also playing a more important role. For example, this happens in remote sensing systems, Smart medical, security, industrial testing, and detection. It can be said that the current image techniques' applications are in various fields. Some examples are shown in Fig 1.1. Let us consider some types of images more in detail.

1) Remote sensing images. The development of low-cost airborne hyper spectral sensors and other imaging devices has facilitated the wider application of remote sensing, including monitoring of crop growth status and assessment of crop yield[56, 57]; ecological environment monitoring such as forests and wetlands[58-60]; mapping and architectural planning[61]; emergency disaster prevention and control[62, 63], and other civilian and military applications[64].

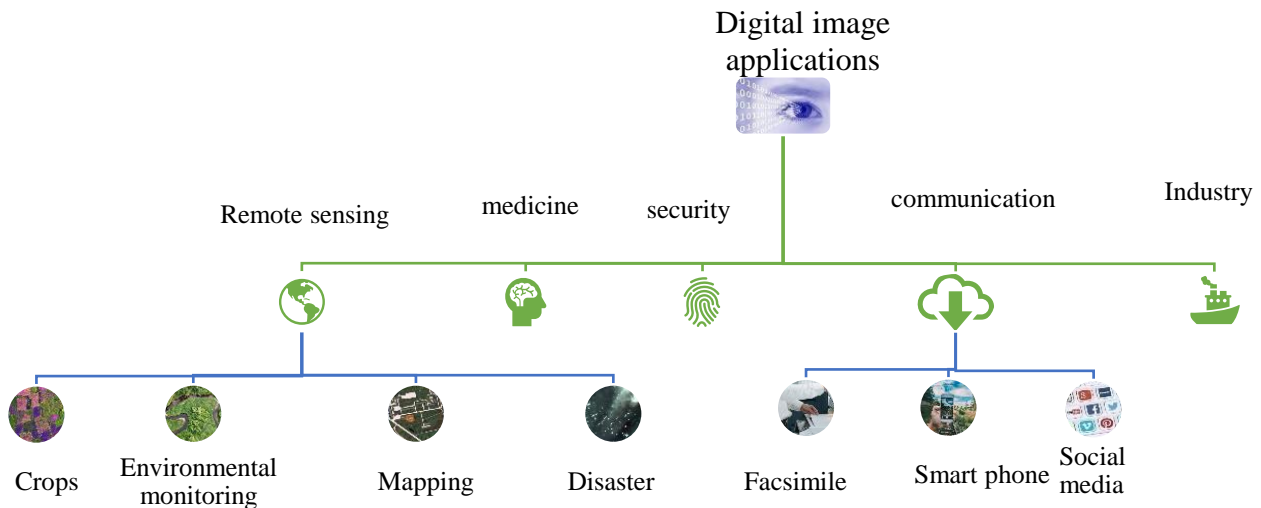


Figure 1.1–Examples of digital image applications

2) Medical images. Magnetic resonance imaging (MRI) and Computed tomography (CT) scans are used to quantify tumors, Positron emission tomography (PET) and functional MRI scans are utilized to detect regional metabolic brain activity[2, 3, 65]. Medical imaging has also become a great method to combat the COVID-19 outbreak and it helps controlling the pandemic crisis[66].

3) Security. Biometric images are being used increasingly in security technology, where fingerprint and face are utilized for identification. Large-scale touch-based fingerprint recognition systems are not only used worldwide by law enforcement and forensic agencies but are also deployed in Smartphone applications[1, 67]. Driven by computer hardware and imaging technology development, Face recognition has been applied widely to daily lives, such as access control, video surveillance, etc. [68, 69].

4) Communication. Nowadays, People use the self-help shooting of Smartphone cameras to record videos of family and friends. This results in the fact that quality of Smartphone camera photos is becoming one of the dominant considerations when consumers purchase Smartphone[70]. Furthermore, these images are shared on the social media platforms, uploaded, downloaded, and viewed via different types of communication lines, including wireless ones[18]. In addition, through photocopying of important paper handwritten materials, information exchange and preservation in commercial activities can be better achieved.

5) Industry. Under the background of industry 4.0, Quality inspection is an essential component of smart manufacturing systems. The powerful and affordable digital cameras

and machine learning algorithms have elevated the potential of computer vision in quality inspection applications to a new level[71]. In Non-Destructive Testing (NDT), ultrasound imaging is a helpful tool for detecting flaws or measuring corrosion in metal pipes, and other equipment flaw detection applications[72].

1.1.2 Related technologies of digital image processing

In these applications, images need to be acquired, subsequent processing and analysis are required; in addition, these images need to be transmitted, stored, and displayed. The entire digital image processing system is shown in Fig.1.2. Therefore, with the development and widespread use of digital images and imaging technology, many related technologies have also been derived.

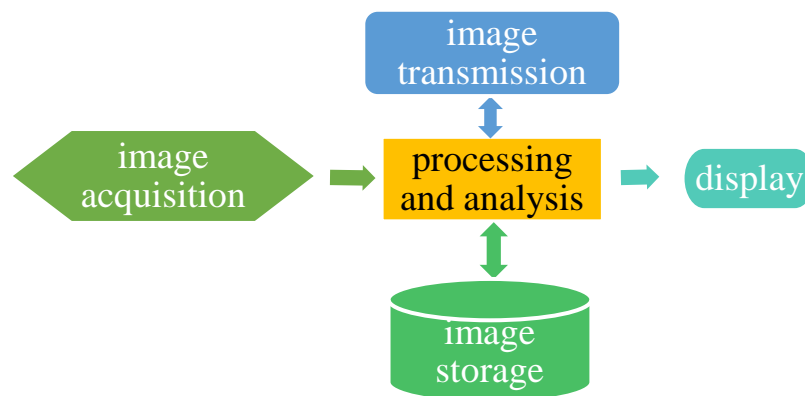


Figure 1.2–Digital image processing system

First, an analog image obtained by the imaging sensor array often needs to be discretized to obtain a digitally represented image. This process includes discretizing the spatial coordinates (sampling) and the discretization of the grayscale (quantization). An analog image is thus converted to the digital image (Fig.1.3), which can be expressed and processed in a computer through a matrix, such as a formula (1). Each gray value in the matrix corresponds to a pixel in the image, and the color image is composed of three RGB channels, respectively represented by three gray value matrices (or in some other color system). The properties of digital images depend to a large extent on the number of samples and gray levels used in sampling and quantization.

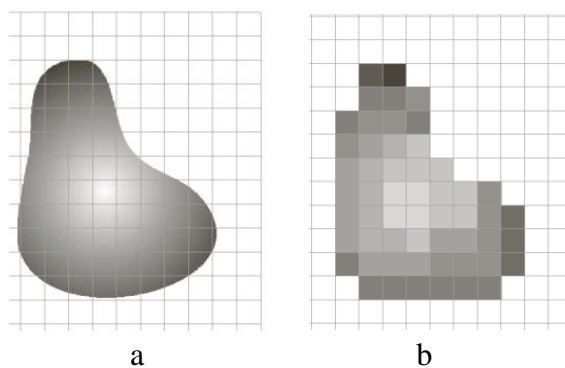


Figure 1.3–The process of digitizing images, a) analog image, b) results after sampling and quantification.

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \dots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \dots & a_{1,N-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{M-1,0} & a_{M-1,1} & \dots & a_{M-1,N-1} \end{bmatrix} \quad (1.1)$$

Second, a digital image is usually processed or analyzed by serial algorithms. Image processing usually outputs images, including image enhancement[73-75], image fusion[4], image compression[19, 20, 23], etc.; Image analysis often output images properties, including image segmentation[76-78], pattern recognition[1, 79], etc. Consider these operations more in detail.

1) **Image enhancement.** The purpose is to enhance the valuable information in the image through smoothing and/or sharpening and other methods to suppress noise enhance edge and texture information to improve the image visual effect. An example is restoration of images captured in harsh conditions such as low-light [73-75].

2) **Image fusion.** The goal of image fusion is to integrate the complementary or redundant information obtained by different imaging equipment or different imaging modes for the same scene or the same target into an image, to facilitate human observation and understanding or further computer processing[4, 6, 80].

3) **Image compression.** It is also known as image coding that stores and transmits digital images with the smallest possible amount of data by removing statistical redundancy, spatial redundancy, irrelevant information, and coding redundancy[21, 29, 81-85].

4) **Image segmentation.** An image is divided into several disjoint regions according to features such as grayscale, color, spatial texture, geometric shape, etc. These features

show consistency or similarity in the same region but apparent differences in different regions. Its goal is to separate the target from the background in an image to achieve tasks such as classification[76, 78, 86, 87].

5) **Pattern recognition.** This direction relates to study how to use computers to simulate or realize human learning activities, through the extraction of image features, to establish recognition and classification models for things, in face recognition, fingerprint recognition, optical character recognition, natural language processing, and many Information management systems fields are widely used[79, 88].

Finally, digital image information is transferred via a communication line, saved to a storage medium, and viewed on display.

1.2 Image compression

Digital image data have brought convenience to the transmission of information and decision-making. However, at the same time, the massive amount of image data has continuously increased the requirements for platform storage resources and network transmission bandwidth. Original image requires considerable storage capacity and transmission bandwidth. Despite rapid development in mass-storage density and digital communication system performance, demand for image storage capacity and transmission bandwidth still outpace the capabilities of available technologies[89].Therefore, the effective image compression techniques are essential and becoming to be the most significant in the image processing field.

1.2.1 A review of image compression

Image compression aims to express as much image information as possible with as little data as possible. In this way, the image can be transmitted at limited bandwidth. The transmission efficiency can be improved; many images can be stored in the limited capacity, and the space utilization rate can be improved.

Before image compression, algorithms such as filtering and denoising may be used for image enhancement and image fusion to improve the quality of an image. After the image is compressed, it can be stored locally or transmitted remotely. When the user needs to browse or complete object recognition, classification, and other image understanding tasks through image analysis algorithms, the compressed image needs to be decompressed first. Image compression is closely related to all aspects of digital image systems. With the vigorous development of digital images, image compression technology has also attracted more attention and has become a hot research topic.

Since the digitization of television signals was proposed in 1948, people have begun to study image compression coding. In 1952, Huffman proposed the "minimum redundancy construction method"[90] and described his algorithm for producing optimal codes of integer length. In 1975, arithmetic coding was introduced by Rissanen, and this coding technique may be regarded as successfully approaching the limit of information entropy[91]. In 1984, Chen and Pratt applied the discrete cosine transform (DCT) technique to image compression[92]. In 1986, the Joint Photographic Experts Group (JPEG) team was created to develop a compression standard for still images[93]. In 1992, Lewis and Knowles used the 2-D wavelet transform in image compression[94], the most widely used coding method currently, and it has been maintained and optimized since then. More versions were released, such as JPEG 2000[85], JPEG XS[95],etc. With the continuous development of deep learning, it has become one of the hot research topics how to apply deep learning in image compression to obtain a higher compression ratio and higher quality[96, 97].

A range of compression methods have been developed to address main challenges faced by digital imaging over the past few decades. In general, these image compression technologies are divided into two categories, lossless compression and lossy compression, it concerns the quality of an image after decompression. The two groups of methods use different implementation algorithms and are suitable for different applications.

1.2.2 Lossless compression

Lossless compression is generally implemented in two steps: first, the original images are transformed to some other format, and the inter-pixel redundancy is minimized during this process; second, the symbol encoder is utilized to remove the coding redundancy. An inverse process is conducted while the lossless compressed image be decompressing (Fig. 1.4)[89].

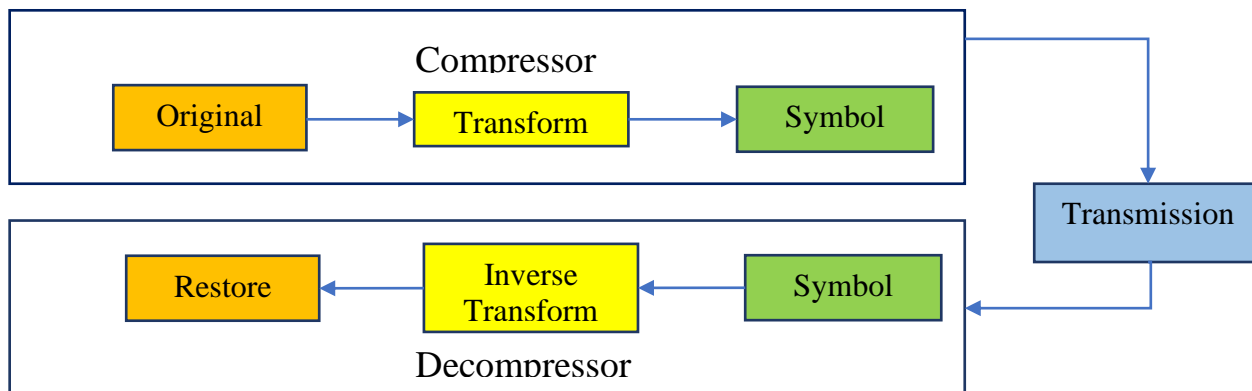


Figure 1.4 –Lossless image compression process

Transform: Removal in image redundancy is typically achieved by transforming the original image data from one form of representation to another[22].Some of the popular transformation techniques are Discrete Fourier Transform (DFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and others, each has advantages and disadvantages[89].

Symbol coding: The encoding process generates fixed-length or variable-length codes, minimizing code redundancy without any information loss. There are some outstanding entropy coding algorithms, such as Huffman coding, arithmetic coding, and their variations as context-adaptive variable length coding, context-adaptive binary arithmetic coding, and dictionary-based coding[89, 98]. The symbol decoding is the exact inverse process of the symbol encoding, which restores the original image from the encoded binary codeword[99, 100].

The information is completely restored in lossless compression, and the decompressed image is identical to the original image. For lossless compression, an important measure is compression ratio (CR), which is calculated as follows,

$$CR = \frac{B_o}{B_c} \quad (1.2)$$

where, B_o is the size of the original image in bytes, and B_c is the compressed file size in bytes.

Besides, the lossless compression algorithms are also compared in view of another parameter, time complexity, which is the total time required to compress and decompress the images.

The implementation of most compression algorithms is often not based on only one specific encoding method, but a combination of multiple encoding methods is used. This approach helps to improve the compression ratio but also increases the time complexity.

Lossless compression techniques are commonly used to compress medical images, archive scans, biometric images, and other high-value images where the data loss is intolerable. The theoretical limit of lossless compression is information entropy. It is required that the amount of information is not lost during the encoding process. Therefore, the compression ratio is limited, and the average value is about 1.2~5[3, 22, 23]. However, the image can be restored without loss in this compression method.

1.2.3 Lossy compression

If the decompressed image is not precisely the same as the original image, the image compression technique is a lossy one. Unlike the lossless compression, almost all lossy compression algorithms are conducted in three stages: first, a transform is used to remove the inter-pixel redundancy; second, a quantizer is utilized to eliminate the psycho-visual redundancy; third, a symbol coder is applied to get extra compression from the coding redundancy (Fig. 1.5). Decompression is the reverse process.

Transform: Similar to the algorithm in lossless compression, the image gray value matrix is mapped by DCT or DWT to a coefficient matrix.

Quantization: This step is the key to the difference between lossy compression and lossless compression. Quantization is several to one mapping that replaces a set of values with only one value, and the process is non-reversible. The loss occurs due to some insignificant information being discarded, which is treated as psycho-visual redundancy[98]. Nevertheless, of course, this discarding also results in a higher

compression ratio and dictates the quality of the reconstructed image. An example of DCT quantizer code is given in Fig.1.5.

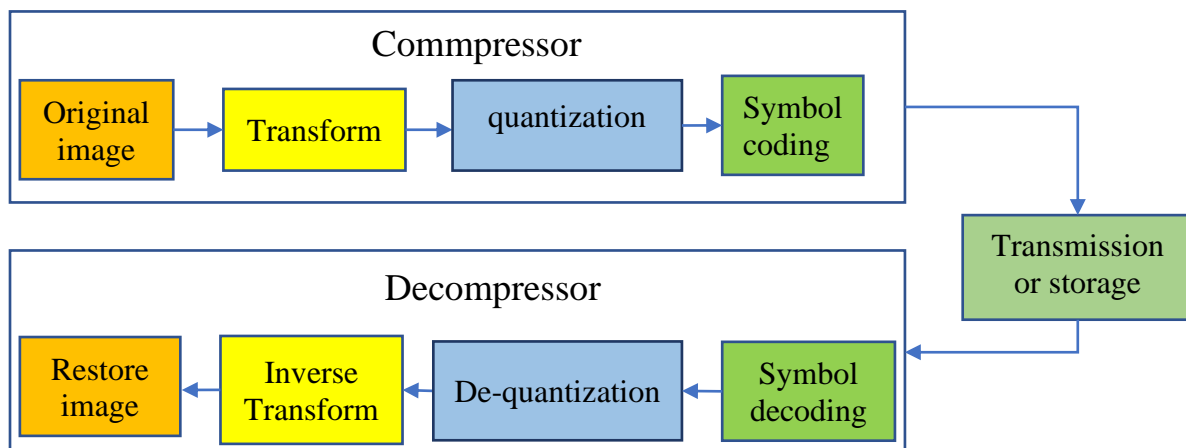


Figure 1.5–Lossy image compression process

1) a block of 8×8 DCT coefficient matrix is:

$$B = \begin{bmatrix} -415 & -33 & -58 & 35 & 58 & -51 & -15 & -12 \\ 5 & -34 & 49 & 18 & 27 & -5 & 1 & 3 \\ -46 & 14 & 80 & -35 & -50 & 7 & 19 & -18 \\ -53 & 21 & 34 & -20 & 2 & 36 & 34 & 12 \\ 9 & -2 & 9 & -5 & -32 & 45 & -15 & 37 \\ -8 & 15 & -16 & 7 & -8 & 4 & 11 & 7 \\ 19 & -28 & -2 & -26 & -2 & -44 & 7 & -21 \\ 18 & 25 & -12 & -44 & 35 & -37 & 48 & -3 \end{bmatrix} \quad (1.3)$$

Low-frequency components (low-frequency signals) represent slowly changing areas in the image, describing the central part of the image. High-frequency components (high-frequency signals) correspond to the parts of the image that change drastically: the edges or noise and details of the image. Discrete cosine transformation is performed on the original image, and in the DCT coefficient matrix after transformation, the low-frequency components are mainly concentrated in the upper left corner[85].

2) A common quantization matrix:

$$Q = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 201 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix} \quad (1.4)$$

Where each element in the matrix is the quantization step, and the quantization step of each coefficient is inconsistent. In practical applications, the quantization step can be multiplied by the quantization matrix Q by a scale factor to obtain an appropriate quantization step[101].

3) the DCT coefficient matrix is divided element-wise by the quantization matrix and then rounded to integers to obtain the quantization result.

$$Y = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (1.5)$$

For example, using -415(the DC coefficient) and rounding to the nearest integer

$$Y_B^Q(m, n) = \text{round}\left(\frac{B(m, n)}{Q(m, n)}\right) = \text{round}\left(\frac{-415}{16}\right) = -26 \quad (1.6)$$

Symbol coding: The new data generated after quantization is minimal, and many are zeroes. Such a matrix can significantly reduce the amount of data by symbol encoding.

In this type of compression, data degradation and loss of quality are expected while it produces higher compression rates and smaller compressed data size than the lossless methods. The basic principle is that the coefficients at different positions are quantized with different quantization step sizes. The human eye is less sensitive to distortions in high frequencies. The low-frequency components are quantized in small steps, and the high-frequency components are quantized in larger steps. This approach improves compression efficiency while maintaining subjective quality.

The lossy compression techniques are mainly applied in natural images, such as images shared in social media[102], civil remote sensing images[103, 104], and other applications where the tiny and, sometimes, imperceptible loss is acceptable.

1.3 Existing lossy compression coders

There are various lossy compression algorithms, and many compression coders are designed based on these algorithms. Different encoders have their specific characteristics and are suitable for different applications. Concerning lossy compression algorithm performance comparison, CR and image quality are the main two measures, and time complexity should also be properly considered in some cases. In [105], four popular lossy compression algorithms are analyzed and compared: DCT, DWT, vector quantizer(VQ) compression, and Fractal compression, respectively. Peak Signal to Noise ratio (PSNR) was utilized to evaluate the decompressed image quality referring to the original image. The DWT-based coder provides high compression ratio and is strongly recommended for lower bit rates; The VQ compression has low time complexity but is not suitable for low bit rate; DCT coder and fractal coder are conditionally applicable to low bit rates. In[106], a comparison was taken between Singular Value Decomposition (SVD) and DCT compression techniques. The DCT compression technique can effectively reduce the mammogram image size by 65% from the original size without affecting the suspicious regions such as micro calcifications. In comparison, the SVD compression technique can reduce the image by 33% from the original image size. In [107], the study on major IoT applications indicates that Discrete Cosine Transform (DCT) and Fast Walsh-Hadamard Transform (FWHT) generate higher compression ratios than others. Lossy Delta Encoding (LDE) significantly outperforms others in terms of information loss. The impact of introduced error is much more severe in DCT and FWHT, while LDE maintained a relatively lower error rate than other methods.

To sum up, it can be seen that different compression algorithms have their applicable occasions. Different coders are developed in the continuous optimization, such as AGU[108],ADCTC[109],SPIHT[5, 110], JPEG[93], JPEG2000[25, 85], BPG[111-113], etc.

1.3.1 DCT-based coders

DCT is an invertible linear transform and is widely used in the field of image compression, where it is used to separate the image into different important parts relative

to the visual quality of the image. The transformation rule of DCT is similar to the Discrete Fourier Transform, but using only real numbers, it transforms a signal or an image from the spatial domain to the frequency domain[114]. For the convenience of operation, the image is divided into many 8×8 (or other size) blocks, where each element falls in the range $[0,255]$ for an 8-bit image, then this range is modified by shifting from $[0,255]$ to $[-128,127]$. Finally, the images are separated into parts of different frequencies by the DCT[115], where the following equations are employed.

The forward 2D-DCT transformation:

$$F(u, v) = \frac{2}{N} C(u)C(v) \sum_{x=0}^{n-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \cos \left[\frac{\pi(2y+1)v}{2N} \right]$$

$u=0, \dots, N-1$ and $v=0, \dots, N-1$ where $N=8$ and $C(k) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{otherwise} \end{cases}$ (1.7)

the inverse 2D_DCT transformation:

$$f(x, y) = \frac{2}{N} C(u)C(v)F(u, v) \cos \left[\frac{\pi(2x+1)u}{2N} \right] \cos \left[\frac{\pi(2y+1)v}{2N} \right]$$

$x=0, \dots, N-1$ and $y=0, \dots, N-1$ where $N=8$ (1.8)

After DCT transformation, the first frequency in the set is the most meaningful; the latter, depending on the allowable loss of resolution, the minor significant frequencies can be removed (see equation (1.3)). Generally, an 8×8 transform coefficient array in which the (0,0) element (top-left) is the DC (zero-frequency) component and entries with increasing vertical and horizontal index values represent higher vertical and horizontal spatial frequencies.

Some coders are developed based on DCT, but with different optimizations, the representatives are JPEG, AGU, and ADCTC.

JPEG: The most popular DCT-based lossy compression coder is JPEG. The JPEG coding architecture is optimized for compression efficiency at even very low bit-rates and optimized for scalability and interoperability in the networks and noisy mobile environments[93, 101]. The coefficient matrix of image blocks is quantized with a uniform or non-uniform scalar quantizer, Zig-zag scanned and entropy coded with Huffman coding[116]. Quality Factor (QF) is one of the most critical parameters utilized to control

the compression. It is an integer between 0 and 100 and is used to scale the values in a quantization matrix. The Zig-zag scanned method is shown in Fig.1.6. The non-zero data is concentrated in the front end (recalling equation (1.5)), and this data arrangement is beneficial to improve the Huffman coding efficiency, thereby reducing the data size.

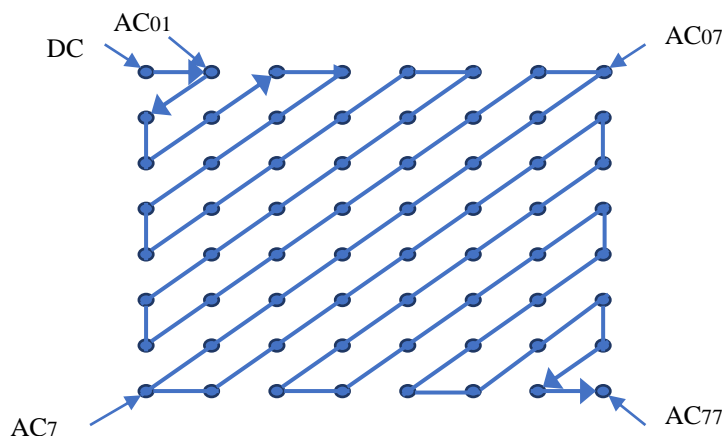


Figure 1.6 –Zigzag scanning in JPEG coder

AGU serial coder: AGU is a DCT-based coder, but it provides decompressed image quality better than DWT-based coder JPEG2000 by up to 1,9dB. The compression efficiency is improved by the following measures: an image is divided into 32x32 pixel blocks; the quantized DCT coefficients are divided into bit-planes, and DCT based filtering[117] is used as post-processing for removal of blocking artifacts from decoded images, thereby improving decoded image quality[108], the block-diagram of AGU coder is shown in Fig.1.7.

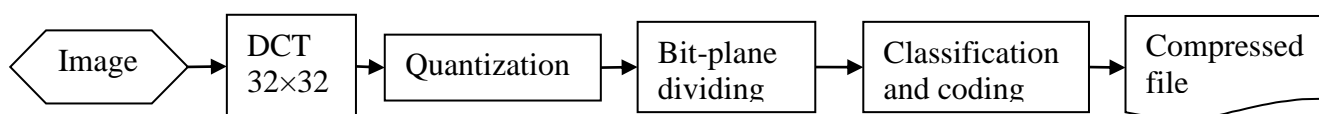


Figure 1.7 – The block-diagram of AGU image coding

In the quantization stage, a uniform quantizer is used to ensure the best results within the structure of the considered method. In the bit-plane stage, the obtained bit-planes are coded in the order starting from higher bits to lower ones. While coding each next plane, the values of bits of earlier coded planes are taken into account. A coded bit is referred to one or another group of bits according to the values of already coded bits. For each group of bits, an individual probability model is used for dynamic arithmetic coding[108].

AGU-MHV is a new high-quality DCT-based method. Its modification is that an image is divided into blocks of different sizes by a rate-distortion-based modified horizontal-vertical partition scheme(PS), which produces higher CR than JPEG, SPIHT, and JPEG2000[118]. However, one of the bottlenecks of the AGU-MHV coder is a relatively high computational complexity of both partitioning an image into blocks and bit-plane coding of transform coefficients[119].

ADCTC uses a cost function to considerably reduce a partition scheme (PS) optimization time. Besides, coding of numbers of significant bits (NOSB), context modeling, and coding of signs of DCT coefficients are utilized to improve the compression efficiency[109].

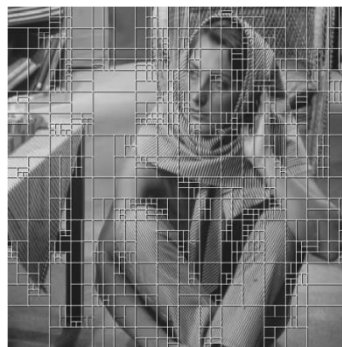


Figure 1.8– Modified horizontal-vertical partition scheme for ADCTC for image Barbara, $\text{bpp}=1$

Fig.1.8 shows the PS obtained for the image Barbara, $\text{bpp}=1$. As can be seen, large size blocks for ADCTC correspond to either homogeneous image regions or fragments with a regular texture.

In addition, some other excellent coders are proposed. In [120], an approximate multiplication-free of discrete cosine transform (DCT) for still image compression was presented. The introduction of null elements into a specified integer DCT leads to a new low complexity, faster, and more efficient transform. In[121], replacing the bit shift elements of a variant of the Signed DCT transform by zeros is involved, it eliminates the bit shift operations. A small amount of arithmetical computation is well developed as no multiplications and bit-shift operations are required, with only 16 additions involved.

1.3.2 DWT-based coders

DCT uses image blocks to transform images and cannot eliminate the correlation between block edges. Therefore, some block effects affect visual appearance, especially at low bit rates. In recent decades, the wavelet transform has been developed to solve the blocking effect caused by discrete cosine transform with its inherent multi-scale analysis structure[94, 122]. The wavelet transform can compress most of the energy into the low-frequency sub-band while separating its high and low-frequency information to reducing spatial redundancy. The definition of discrete wavelet transform[4, 74] is as follows:

$$DW_{m,n} = \langle f, \Psi_{m,n} \rangle = \int_{-\infty}^{+\infty} f(t) \Psi_{m,n}(t) dt = a_0^{-\frac{m}{2}} \int_{-\infty}^{+\infty} f(t) \Psi(a_0^{-m}t - nb_0) dt \quad (1.9)$$

where $f(t)$ is the discrete signal, Ψ is the discrete wavelet basis function, it is expressed as:

$$\Psi_{m,n}(t) = \frac{1}{\sqrt{a_0^m}} \Psi\left(\frac{t - nb_0 a_0^m}{a_0^m}\right) = a_0^{-m/2} \Psi(a_0^{-m}t - nb_0), \quad a_0 > 1, b_0 \in R \quad (1.20)$$

where R is the set of real numbers, let us give $a = a_0^m$, $b = nb_0 a_0^m$, then a is the scale factor, b is the shift factor of the basis function.

Based on the wavelet transform, Mallat[123] proposed a tower multi-resolution decomposition and reconstruction algorithm for signals. Mallat's algorithm is a fast algorithm of the wavelet transform and can be extended to a two-dimensional case, thus to multi-scale analysis (multi-resolution).

Two-dimensional DWT is implemented as follows: first, performs the one-dimensional DWT row-wise to produce an intermediate result; second, performs the same one-dimensional DWT column-wise on this intermediate result to produce the final result[124]. The row-column computation diagram for two-dimensional DWT is shown in Figure 1.9.

After the first level of decomposition, it generates four subbands, LL1, HL1, LH1, and HH1, as shown in Figure 1.9(a). The LL1 subband can be considered a 2:1 subsampled (horizontally and vertically) version of the original image. The other three subbands, HL1, LH1, and HH1, contain higher frequency detail information. The LL1 subband can be further decomposed into four subbands LL2, HL2, LH2, and HH2, as shown in Figure 1.9(b). The same computation can continue to further decompose LL3 into higher levels according to the principle of multi-scale analysis[124].

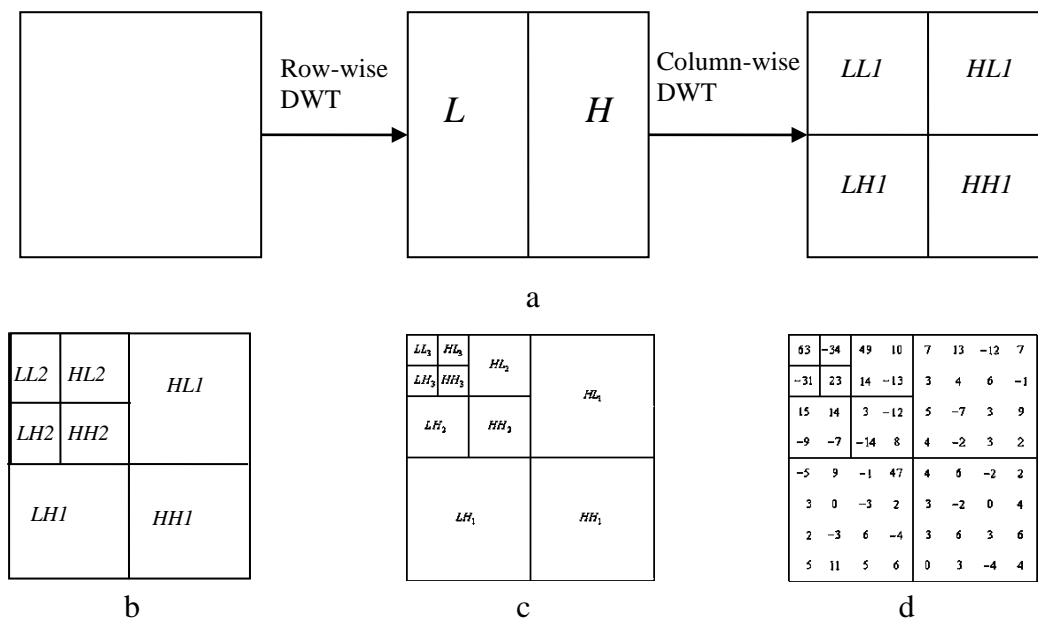


Figure 1.9—Row-Column computation of two-dimensional DWT, a) first level of decomposition, b) second level decomposition, c) third level decomposition, d) three-level wavelet decomposition coefficients of 8×8 image block

Generally, DWT has two advantages for lossy compression. First, wavelet transform is a global transformation, and the "block effect" inherent in block orthogonal transform coding can be avoided in reconstructed images. Second, the wavelet transform adopts the data structure of tower decomposition. Consistent with the visual physiological characteristics of the human eye from coarse to fine, from the whole to the details, better visual quality can be obtained[4, 74].

At present, based on the two-dimensional discrete wavelet transform, excellent lossy image compression algorithms are developed, such as the EZW algorithm, SPIHT algorithm, and EBCOT algorithm, and some popular image coders are designed.

JPEG2000: JPEG 2000 achieves higher CR for equivalent visual quality compared to the baseline JPEG, particularly at very low bit-rates. the image quality can also be adjusted depending on the degree of interest in each region of interest due to the adopted region of interest (ROI) coding[125, 126], which is a unique feature of the JPEG2000 standard. ROI allows different regions of an image to be coded with different fidelity criteria. For lossy compression, the default wavelet filter used in the JPEG2000 standard is the Daubechies (9,7) biorthogonal spline filter. The quantized subbands are then divided

into a number of smaller code-blocks of equal size, except for the code-blocks at the boundary of each subband.

The main drawback of the JPEG2000 standard compared to the current JPEG is that the coding algorithm is much more complex, and the computational needs are much higher. Moreover, bit-plane-wise computing may restrict good computational performance with a general-purpose computing platform[124].

SPIHT: Set partitioning in hierarchical trees (SPIHT) is a wavelet-based algorithm that offers good compression ratios, image quality, and fast execution time[127]. SPIHT coding involves the coding of the position of significant wavelet coefficients and the coding of the position of zero trees in the wavelet subbands. SPIHT introduces three lists of wavelet coefficients[128]:

1) List of insignificant sets (LIS): contains the set of wavelet coefficients defined by tree structures, magnitudes < thresholds.

2) List of insignificant pixels (LIP): contains the individual coefficients, magnitudes < thresholds.

3) List of significant pixels (LSP): contains the pixels, magnitudes < thresholds.

The significance function is defined as follows:

$$S_n(t) = \begin{cases} 1 & \max_{(e,f) \in t_m \{C_{e,f}\}} \geq 2^n \\ 0 & \text{otherwise} \end{cases} \quad (1.21)$$

where $S_n(t)$ represents the significance of the set of coordinates, t and $C_{e,f}$ are the coefficient values of the (e,f).

By sequentially encoding the records in the LIP, LIS, and LSP at different bit levels, the image can be approximated progressively. The DWT-SPIHT was proved to achieve good image quality with high PSNR, and is a very computationally simple algorithm and easy to implement in comparison with other coding methods[128].

1.3.3 Learning-based coders

With the development of deep learning techniques, the use of deep learning for lossy compression can be seen in recent research, greatly improving image compression performance.

In [129], a lossy hyper spectral image compression algorithm was proposed. It uses a combination of the convolution layer and max-pooling layer to reduce the dimensions of the input image then generate a compressed image. In [97], a novel standard-compliant convolutional neural network-based Multiple description coding (MDC) framework was proposed, efficiently leveraging the image context information to compress the image. In [130], an extended hybrid image compression scheme based on soft-to-hard quantification was proposed with only two layers. The base layer determines the performance of the rough reconstructed image. The enhancement layer is to further encode the residual image, thereby further improving the performance of the final reconstructed image.

In addition, most end-to-end lossy image compression schemes based on deep learning methods[131-133] were proposed to solve the blocking, blurring, or ringing artifacts at low bit rates, which are common issues in existing compression coding. Generally, applying deep learning techniques to lossy compression can help improve compression performance and achieve better visual quality. However, the increase in computational complexity is inevitable. It is difficult to implement them in some applications, such as mobile embedded application platforms and platforms with limited computing resources such as remote sensing satellites.

1.3.4 Other coders

JPEG-XR: This standard originated from HD Photo technology developed by Microsoft Corporation, which was developed for digital imaging applications[134]. JPEG-XR image compression algorithm consists of color space conversion, Lapped Bi-orthogonal Transform (LBT) to transform the image from spatial domain to frequency domain, quantization, prediction, and adaptive entropy coding[135].

BPG: Better Portable Graphics (BPG) is a novel compression algorithm, which is based on High-Efficiency Video Coding (HEVC)[113]. Compared to the existing lossy compression coder, BPG has some distinct attributes:

- 1) The open-source and royalty-free and patent-free nature;
- 2) With advanced quality features, chroma formats supported include grayscale, RGB, YCgCo, YCbCr, Non-premultiplied alpha, and Premultiplied alpha.
- 3) BPG is capable of cross-platform use through its JavaScript interpreter.

It can be seen that BPG is a light lossy compression coder, meeting modern display requirements (high quality and smaller size). These characteristics make it suitable for embedded systems or real-time systems, and thin-client browsers. Based on the same bit-per-pixel (BPP), BPG can provide visual quality comparable to compression algorithms based on convolutional neural networks and performs significantly better than the traditional algorithms such as JPEG2000[130, 136, 137], particular in human evaluation according to the MS-SSIM scores.

Quantization parameter Q is the main parameter to control the compression in BPG, where $Q \in \{1, \dots, 51\}$. Smaller Q results in less quantization and better quality, vice versa[138].

From the analysis in this section, it can be concluded that generally speaking, DWT-based lossy coding can achieve better coding efficiency than DCT-based coding, but it also increases computational complexity. In the current technology, AGU series encoders based on DCT encoding can provide visual quality comparable to DWT encoders through some simple optimizations; DWT-based SPIHT, through the significant list of wavelet coefficients, reduces the computational complexity and accelerate coding speed; Among emerging coding technologies, BPG coding can provide visual quality comparable to coding based on deep learning technology, while its portability allows it to be applied to a wider range of applications.

Anyway, it can be considered that the visual quality of the decompressed image (error or distortion degree) is an indicator for choosing a lossy compression algorithm in an application. Thus, it is worth paying special attention to SPIHT, AGU, and BPG encoders.

1.4 Requirements for lossy compression techniques

Lossy compression achieves high CR by discarding some high-frequency components, resulting in distortion and degradation of decompressed image quality. Severe distortions or bad image quality affect the visual perception of human being and cause a negative influence on further image processing and/or analysis.

1.4.1 The requirements of human visual perception

In many applications, the decompressed images are used for viewing by users, and the quality of the images affects the user's visual perception. Lossy compression discards high-frequency information, and the human visual system is not very sensitive to slight distortions of this information. Then if a higher CR needs to be achieved, more information needs to be discarded, and more distortions have to be introduced, which can be noticed when the visual quality drops to a certain value. A lossy compression example of image Lenna is shown in Fig 1.10.

Compared to the original image, the image in Fig.1.10(b) has the appropriate quality with the considerable CR (20,5), which is much larger than for lossless compression. It proves that discarding certain high-frequency information can effectively improve the compression efficiency and has a limited impact on the image visual quality. It is also the advantage of lossy compression and the reason that it is widely used.



a. original image

b.CR=20.5, PSNR=36.4dB

c.CR=128.8, PSNR=28.7dB

Figure 1.10 - Lenna image and its decompressed images

However, when CR continues to increase, the image quality further decreases (Fig.1.10.c), then the difference between the original image and the decompressed image is easy to notice. This means the distortions have exceeded a certain threshold the image quality can be not acceptable for a user.

In lossy compression, in addition to CR, image quality is another important metric. Performing an image quality assessment (IQA) for the decompressed image is necessary. One reason is that under the same CR, the compression efficiency of different lossy compression algorithms can be compared through IQA. Another reason is that it is required to control distortions in many applications, and image quality is a priority in some cases.

For a fixed lossy compression coder, higher CR is in the cost of introduction of distortion, which means that the decrease of quality accompanies the increase of CR. However, the bad image quality can lead to poor visual perception by human beings or errors in further image processing and analysis. An example is shown in Fig.1.10.c, where the distortion can be easily noticed, then this compression is not satisfactory. Therefore, it is essential to control the distortion in a reasonable range[30], and it is desired to provide a visual quality particular value with an acceptable error[51, 139].

1.4.2 The requirements to image processing algorithm

Image compression is not a final operation, and it has to be treated as a method used to save storage space and improve transmission efficiency. The compressed image needs to be decompressed before further image processing or analysis, and the quality of the decompressed image directly affects accuracy of subsequent image processing or analysis.

Classification is a common image analysis task, which is demanded widely in remote sensing, IoT, and other digital image fields. In these applications, the data acquisition device often has limited resources. Images need to be transmitted to a base or processing center with restricted bandwidth for further processing or analysis. Thus, it is necessary to use lossy compression to reduce the size of the image. However, the classification accuracy is in correlation with image quality. Distortions introduced by lossy compression

can lead to degradation of image quality, which can impact on the accuracy of classification.

Some researchers have made efforts to study the impact of lossy compression on the classification accuracy of remote sensing images. In [125], a study was carried out on the implications of JPEG (JPG) and JPEG 2000 (J2K) lossy compression for image classification of forests in Mediterranean areas. The result proved that even at high compression levels, the resulting visual effects do not necessarily severely impact on the classification results. However, if the optimal compression level is exceeded, the classification accuracy can be adversely affected. Depending on the image fragments, the inflection points are located at different CRs. In [140], the study demonstrates that the effects were almost linear when the CR was less than 35 on their data set. However, once the CR was over 35, the image quality would significantly decrease, leading to a sharp drop in classification accuracy. An interesting finding in [141] shows that in some cases, a higher compression ratio may yield better overall classification accuracy due to the local spatial-spectral smoothing effect introduced by specific compression schemes.

Unlike the previous studies that focus on qualitative research on the impact of lossy compression on the accuracy of classification algorithms, other researchers focus on quantitative analysis of these effects. In [142], a quantitative study has analyzed the effects of compression on remote sensing image classification and proposed a method to estimate the remote sensing image classification accuracy based on fractal analysis. It investigated the effects of compression ratio on classification accuracy of both overall and particular classified types. Their results demonstrate that the extent of the effects depends on both the compression levels and the image contents. For quantitative research, metrics such as PSNR, SSIM and MS-SSIM, and VIQ were used to evaluate the quality of images after compression, and the correlation between these metrics and overall classification accuracy and particular class classification accuracy was analyzed by SROCC value, among them, MS-SSIM correlation is higher than other full parameters.

In [125], the final task is to recognize and classify shapes in a processed 2D image. These images are acquired in distributed, low-power IoT sensors then offloaded on a remote server where more computational power is available. Hence, lossy compression is

essential to save bandwidth and power during transmission. The results show that the Convolutional Neural Networks (CNNs) classification performance was correlated with the quality of the images in terms of the Structural SIMilarity (SSIM) metric. Moreover, each image can be associated with an SSIM threshold below its classification rank drops. However, no clear correlation or threshold was found in the experiments according to Peak Signal to Noise Ratio (PSNR) metric.

It can be seen from the above analysis that the image quality in lossy compression needs to be constrained not only in applications based on human perception but also in other image processing or analysis tasks. Moreover, selecting evaluation indicators for image quality is also very important, and the correlation between different metrics and task execution accuracy is different. The current popular visual quality evaluation metrics will be reviewed in the next section.

1.4.3 Image quality metrics

From the above analysis, it can be concluded that if it is desired to compare the performance of different compression coding algorithms or to control the distortions in lossy compression, it is necessary to evaluate the quality of the decompressed image. Currently, the research on image quality of lossy compression has become an important branch. Image quality assessment (IQA) methods can be categorized into subjective and objective.

Subjective assessment is the most reliable way to evaluate the quality of images since humans are the main consumers of images[143]. The mean opinion scores (MOS) are collected in the subjective tests, where many observers are required for assessing the visual quality of a huge number of distorted images in image databases, e.g., LIVE image quality assessment database[144], Tampere image database 2013 (TID2013)[145], MICT image quality evaluation database [146]. Finally, the MOS of image quality can be determined. However, subjective image quality assessment is a costly process that requires many observers and takes a lot of time. Therefore, it cannot be used in automatic

evaluation programs or in real-time applications[147]. A typical example is RS image lossy compression, which is expected to be performed quickly.

Contrary to subjective evaluation, Objective quality assessment methods, usually designed to automatically predict the visual quality, are implemented by training using the supervisor evaluation database. Numerous metrics have been designed and proposed to evaluate image quality in recent years. According to the reference to the original image, these metrics can be divided into three categories, full-reference (FR), reduced-reference (RR), and no-reference (NR) assessment. This is according to the availability of the distortion-free reference image. FR-IQA requires an image as the original reference, while RR methods require access to certain features that have been extracted from the reference image. On the other hand, NR methods evaluate the quality of the distorted image in the absence of the reference image[148]. Since the reference images are available at the encoder side, the FR-IQA metrics are applicable to evaluate the distortions introduced by lossy compression. Therefore, this section follows with a brief review of popular FR visual quality metrics.

1) Conventional error pixel-based metrics

The most classical metrics are mean square error (MSE) and peak signal-to-noise ratio (PSNR), both focusing on pixel-level signal fidelity. These measures are calculated according to the equations (1.22) and (1.23). Here R denotes the reference (original) image, T denotes the test (distorted) image.

$$MSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [R(i,j) - T(i,j)]^2 \quad (1.22)$$

where i and j are pixel indices, and M, N are the height and width of an image. PSNR is calculated for the test image and reference images as

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (1.23)$$

where MAX_I denotes the maximum possible pixel value of the image. If it is an 8-bit unsigned integer data type, then MAX_I equals 255. PSNR indicates the ratio between the maximum possible signal power and the power of the distorting noise, which affects the image quality. It is expressed in dB, and a larger value relates to fewer distortions.

Image quality assessment metrics such as MSE and PSNR do not consider the human visual perception features, and then the assessment result is often inconsistent with the human subjective perception. Nevertheless, they are still the most widely used objective metric since they are simple for mathematical calculation and are clear in the physical meaning. While many new objective metrics have been proposed, PSNR is usually treated as one basic reference metric to be analyzed.

2) HVS-inspired metrics

In order to give the evaluation, which is more consistent with the quality evaluation results of the human visual system on the image, some researchers combined the characteristics of HVS with pure mathematical algorithms. The following metrics are the typical HVS-inspired metrics.

VSNR: visual signal-to-noise ratio (VSNR)[149], operates based on physical luminance and visual angle, wavelet-based models of visual masking and visual summation are utilized for distortion analysis.

PSNR-HVS:HVS-based peak SNR (PSNR-HVS)[150] combines PSNR with HVS characteristics by considering the contrast sensitivity function (CSF).

PSNR-HVS-M: a modified version based on the PSNR-HVS, it is defined in a similar way to the PSNRHVS, but the difference between the DCT coefficients is further multiplied by a contrast masking metric (CM) for every 8x8 block (Fig. 1.11). Here MSE_H is the MSE taking into account CSF, each DCT coefficient of an image block in some degree masks any other block coefficients, except DC coefficient with the index 0,0 that corresponds to the block mean luminance. The basis of the model calculation is masking the degree of each coefficient depending upon its square value (power) and human eye sensitivity to this DCT basis function determined by means of CSF[38].

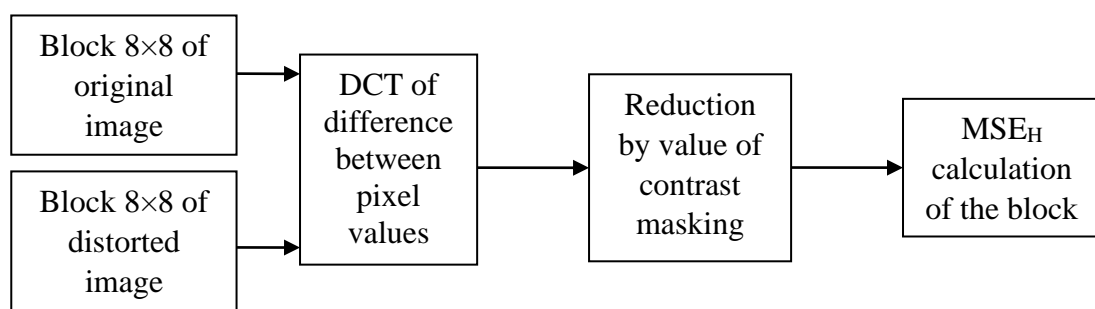


Figure 1.11 – Flow-chart of PSNR-HVS-M calculation for each 8×8 pixel block

Both PSNR-HVS and PSNR-HVS-M take into account the effect of different sensitivity of HVS to distortions in different spatial frequencies[32]. The considered metrics are calculated as

$$PSNR - HVS = 10 \log_{10}(255^2 / MSE_{HVS}) \quad (1.24)$$

$$PSNR - HVS - M = 10 \log_{10} \left(\frac{255^2}{MSE_{HVS-M}} \right) \quad (1.25)$$

where MSE_{HVS} and MSE_{HVS-M} are mean square errors (MSEs) weighted in DCT domain with specific weights.

PSNR-HMA: In order to improve the performance of PSNR-HVS-M on the Exotic and Exotic2 subsets of TID2008, the PSNR-HMA metric was proposed, considering a psycho-visual feature that image stretching and mean changing (in some limits) do not considerably influence human perception of images[32].

3)Structural comparison-based metric

UQI: universal image quality index(UQI)[151]is designed by modeling any image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion.

SSIM: structural similarity (SSIM)[152]is a modified version of the UQI. It operates in the spatial domain and performs three types of comparisons between the reference and distorted images: luminance comparison, contrast comparison, and structure comparison. The three components are combined according to the combination function $f(\cdot)$ as shown in equation (1.26). Here x is the original image having perfect quality, y is the reconstructed image for lossy compression

$$S(x, y) = f(l(x, y), c(x, y), s(x, y)) \quad (1.26)$$

Three functions, $l(x,y)$, $c(x,y)$, $s(x,y)$, Respectively represent luminance comparison, contrast comparison, and structure comparison. The three components are relatively independent.

MS-SSIM: a multi-scale version of SSIM (MS-SSIM)[153] uses five scales instead of the single-scale approach in SSIM. The contrast and structural comparisons are performed at all scales, while the luminance comparison takes place only at the final scale.

CW-SSIM: complex wavelets-SSIM (CW-SSIM)[154] is simultaneously insensitive to luminance change, contrast change, and spatial translation. The key idea is to make use of the fact that these image distortions lead to the consistent magnitude and/or phase changes of local wavelet coefficients.

4) Multi-strategy combined metrics

Some other metrics have two or more separate strategies, different from the metrics above, which adopt a single most relevant strategy used by HVS to assess image fidelity.

FSIM: feature SIMilarity (FSIM)[155] uses phase congruency (PC) and gradient magnitude(GM) as the primary and secondary features. PC and GM play complementary roles in characterizing the local image quality. It is proposed based on the fact that the human visual system (HVS) understands an image mainly according to its low-level features, such as edges and zero-crossings. The computation consists of two stages: first, the local similarity map is computed; then, the similarity map is pooled into a single similarity score.

$$S_L(x) = S_{PC}(x) \cdot S_G(x) \quad (1.27)$$

Then $S_{PC}(x)$ and $S_G(x)$ are combined to get the similarity $S_L(x)$ at each location x . However, different locations have different contributions to HVS perception of the image. The human visual cortex is sensitive to phase congruent structures, for example, edge locations. The PC value at a location can reflect how likely it is a perceptibly significant structure point. Then the max PC value $PC_m(x)$ between the original image and distorted image at location x is utilized to weigh the importance of $S_L(x)$

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (1.28)$$

where Ω means the whole image spatial domain. The FSIM index is designed for grayscale images or the luminance components of color images. It can be extended to the color version called $FSIM_c$ [155] by incorporating the chromatic information in a straightforward manner:

$$FSIM_c = \frac{\sum_{x \in \Omega} S_L(x) \cdot [S_c(x)]^\lambda PC_m(x)}{\sum_{x \in \Omega} PC_m(x)} \quad (1.29)$$

where $Sc(x)$ is chrominance similarity measure combined by I and Q components in YIQ color space, and $\lambda > 0$ is the parameter used to adjust the importance of the chromatic components.

VSI: A visual saliency-induced index (VSI)[156] is a simple but very effective full reference IQA method using visual saliency (VS). First, VS is used as a feature when computing the local quality map of the distorted image. Second, when pooling the quality score, VS is employed as a weighting function to reflect the importance of a local region.

DSS: DCT subbands similarity (DSS)[157] exploits essential characteristics of human visual perception by measuring change in structural information in subbands in the discrete cosine transform (DCT) domain and weighting the quality estimates for these subbands.

MDSI: mean deviation similarity index (MDSI)[158] is calculated by combined gradient similarity (GS), chromaticity similarity (CS), and deviation pooling (DP). Gradient magnitude is used to measure structural distortions, and chrominance features are used to measure color distortions, then the similarity maps are combined to build a gradient-chromaticity similarity map. Finally, the deviation pooling is used to compute the global variations, which is followed by power pooling.

According to MDSI, an image with perfect quality is assessed by a quality score of zero since there is no variation in its similarity map. Experiments prove that MDSI shows a very good compromise between prediction accuracy and model complexity.

5) Learning-based method

Benefitting from the development of neural networks, some researchers have proposed learning-based objective IQA metrics recently[159-161]. Generally, this method establishes a human visual system model by large-scale experiment data. Let us recall that the subjective IQA is based on a certain subjective standard and manual quality evaluation, which represents the natural human perception but at the expense of a considerable time. On the contrary, the objective IQA can evaluate the image quality quickly by the mathematic model to simulate the human visual perception. However, this fitting is often not perfect, so modification is required according to application conditions.

Learning-based objective IQA metrics can fit the HVS better than traditional objective IQA metrics. They can train the model in terms of application and the corresponding data but not calculate in a fixed equation. After the model is trained, it can automatically generate the value, and this processing is usually more efficient than the subjective IQA. This image quality evaluating method has been used in hyper spectral remote sensing images[159], and ground-based optical images of satellites[160]. The learning-based objective IQA shows its potential but not the main research direction in lossy compression since high computation efficiency is often required.

Through the review in this section, it can be concluded that visual quality is very important for the comprehensive evaluation and practical application of lossy compression, and the use of reliable objective visual quality indicators can sufficiently help algorithms to evaluate the visual quality of decompressed image, thereby to perform lossy compression with the quality that meets user requirements, at least no negative impact on human perception or further image processing and analysis. There are two factors that need to be considered in the selection of IQA metrics in this topic: the first is reliability, the consistency between the predicted visual quality and the visual quality of the distorted image will be compared in detail in the next chapter; the second is the calculation complexity, overly complex calculations will reduce the time efficiency of compression. Therefore, learning-based metrics are not further considered.

1.5 Analysis of methods of image quality controlling in lossy compression

Currently, the research on image quality in lossy compression focuses on two aspects. One is to provide better image quality in compression coding; the second is to study the metrics for evaluating image distortions and improve their consistency with the human visual system. However, few studies focus on the control of decompressed image distortions in lossy compression.

The distortion control does not mean providing a high compression ratio is not required. However, high compression ratios inevitably lead to more information being discarded, thereby reducing the quality of the decompressed image. In order to balance

both, Some trade-off should be found between a compressed image quality and a produced CR [162]. Besides, in lossy image compression, image quality has to be described using some adequate (application or service-oriented) quality metrics [32, 38, 150, 155, 158, 163]. The trade-off of lossy compression can be reached if:

- a) an adequate metric is available;
- b) a tool (algorithm) for quality variation is available;
- c) a method for providing the desired quality with appropriate accuracy is available, and it is able to work quickly enough.

In other words, one has to provide high computational efficiency, high reliability, and accuracy of an approach to providing the desired quality of compressed images[44].It is not an easy task because it depends on many factors, such as lossy compression coder adopted, image to be compressed, quality metric employed. Fortunately, the method of controlling the distortion in lossy compression also attracts a part of research interest. Some researchers have done some preliminary related research and achieved some results.

1.5.1 Visually lossless lossy compression

Visually lossless is a simple tradeoff based on visual perception. These methods lie in between lossy and lossless compression: they introduce compression distortions but ensure that those distortions are unlikely to be visible.

In [164], visually lossless compression was first introduced to compress medical images, where the vision model was embedded into the SPIHT coder. Visually lossless coding refers to the compression of images without evoking any perceptible degradation in image quality through the suppression of visible distortions. The implementation of the code consists of three stages, a forward wavelet transform, vision modeling, and SPIHT encoding.

Another visual lossless compression method was applied in remote sensing images within the framework of JPEG 2000 in [165], where the quantization step size needed for visually lossless display was a resolution function. This visually lossless compression was

achieved by applying multiple visibility thresholds corresponding to various display resolutions in each subband.

Recently, a CNN-based method was proposed to train an accurate visibility metric for visually lossless image compression[166]. The visually lossless threshold (VLT) was utilized to set the CCP that produces the smallest image file while ensuring visually lossless quality.

1.5.2 JND point

In addition to the visual lossless threshold, the researchers found that humans have other perception points for the visual quality of images. The visual Just-Noticeable-Difference (JND) metric is characterized by the detectable minimum amount of two visual stimuli. The JND metric can be used to save coding bitrates by exploiting the special characteristics of the human visual system[167].

In [167], a new angle based on JND was utilized to evaluate the JPEG coded image quality. The results show that the human perceived quality of coded images can be characterized by a piecewise constant function of the QF(the main CCP in JPEG) with discontinuities at JND locations. However, these locations are content-dependent and statistically distributed.

The fine-grained distortion levels were given for perceptual-based image compression in [168].A reliable subjective experiment with pair-wise comparison was utilized to rank the qualities of these distorted images.

However, among these JND points, the first one is the most significant, seen as the boundary between lossy and visually lossless compression. In [169], a high level of correlation was found existing between a simple image feature – mean gradient magnitude and the peak signal-to-noise ratio (PSNR) of the first just noticeable difference point for JPEG image compression. A method was proposed to estimate the JPEG quality factor, which represents the effective limit between perceptually lossy and lossless coding as the PSNR of the first just noticeable difference point. The subjective trial results are presented through stair-like quality functions (SQF) obtained through analysis and post-processing

of raw JND data. They show that human observers differentiate between 4 and 8 quality levels.

The JND-based method can be well based on the minor difference the human visual system can perceive to obtain a suitable compression ratio. But this approach relies on subjective experimentation, and there are certain difficulties in embedding it into a lossy compression algorithm because it also relies on image content.

1.5.3 Iterative method

Visually lossless compression and JND-based compression aim to provide a special image quality in lossy compression, around a certain visual threshold. However, the requirements of visual quality in applications are individual. Obviously, none of the above methods can provide arbitrary visual quality values for lossy compression. Some researchers have used iterative methods in lossy compression to achieve perfect precision visual quality.

First, the iterative method was used to achieve better accuracy for decompressed image quality. A novel scheme for lossy compression of an encrypted image with a flexible compression ratio was proposed in [84]. The better quality was achieved while reconstructing the principal content of the original image by iteratively updating the values of coefficients.

In [81], iterative image compression procedures are proposed to provide a desired visual quality with high accuracy. The PCC values attained at the end of the iterative procedure may heavily depend upon the coder used and the complexity of the image. However, the reasonable selection of the starting value and the variation interval for CCP help reducing the number of iterations.

In [139], a quality-controlled reconstruction of ECG signal by the formulation of 2D Discrete Cosine Transform (DCT) coefficient and iterative JPEG2000 encoding scheme was presented for Electrocardiogram (ECG) data compression. Iterative JPEG2000 encoding is utilized for achieving the overall predefined reconstruction error.

This method can get better quality accuracy in lossy compression by presuming multiple compression/decompression, but the number of iterations is uncertain, and sometimes it is very large, which reduces the time efficiency of compression.

1.5.4 Prediction method

Predicting the image quality before using a compression algorithm is an approach with better time efficiency than iterative one.

In [170], a classification-based image compression approach has been proposed, where images are assigned to one of two classes before their compression by a JPEG algorithm considering their quality after compression. This method allows predicting and automatically selecting the proper value of QF for each image before compressing, thereby saving storage space while maintaining sufficiently high image quality. Several features are selected to describe the image, and a majority of the features are extracted from images processed by Prewitt edge detection filter. Suppose the $SSIM > 0.91$ and $PSNR > 33$ dB, the quality of the decompressed image is considered to be satisfactory (high quality). Otherwise, it is classified as low quality.

In [171], a fast and rather accurate approach to the prediction of mean square error (MSE) or peak signal-to-noise ratio (PSNR) was developed for the ADCT coder. It analyzes statistics of DCT coefficients in a limited number of blocks of size 8x8 pixels. This prediction method is fast but has a limited accuracy.

Although the extraction of image features in the prediction method can estimate the quality of the decompressed image and avoid the operation of multiple compression/decompression, the prediction accuracy is limited. It is difficult to determine the CCP in lossy compression by purely predictive methods to provide desired visual quality.

1.6 Conclusions and formulation of research objectives

This chapter firstly introduces digital image technology, clarifies the role and application scope of image compression in digital image technology, and its relationship

with other image processing and analysis technologies. In the second section, a brief introduction to lossy compression technology and lossless compression, the basic implementation scheme and algorithm characteristics of the technology are given; it is shown that lossy compression has been widely used in digital image technology due to its superior compression ratio. In Section 1.3, an overview of currently existing lossy compression encoders is given. Certainly, the introduction of distortions also makes it necessary to evaluate the decompressed image quality in lossy compression. Subjective evaluation involves a lot of manual labor, which cannot be achieved in automatic calculation. In contrast, the proposal of an objective evaluation metric enables the image quality to be quickly calculated, which can be directly fed back to the lossy compression process. Therefore, the fourth section of this chapter reviews the commonly used full-parameter image quality evaluation indicators.

From the previous literature review in the fourth section, it can be seen that how much loss can be allowed is application-dependent. If the terminal is a human, the distortion is often required to be undetectable, that is, visually lossless. Further, it is required to provide image quality corresponding to different levels of visual perception. If the image requires further processing or analysis rather than direct viewing, the requirement for decompressed image quality will then be individualized, as it depends on the impact of image quality on the task, and it is desirable not to negatively affect the performance or accuracy of the task. Moreover, selecting the image quality metric is also important since the correlation between the values calculated by different metrics and the accuracy of the final task is not the same[172].

In order to obtain compression that is safer for the user, i.e., without loss of visual perception or adverse effects on subsequent image processing or analysis, the controlling of images distortions introduced by lossy compression is demanded in many applications. Then, in the fifth section, a review is given to analyze the methods of image quality controlling in lossy compression. Visual lossless and JND-based visual quality provides distortion control for the perceptual threshold of human vision. In short, as long as the visual quality is guaranteed to be higher than a particular value, there is no clear requirement for the precision control of the provided visual quality. However, in some

image processing or analysis algorithms, the requirements for image quality are more specific. At this time, it is necessary to provide the expected visual quality in lossy compression for these tasks and meet certain accuracy requirements. The prediction method can extract information and features through a certain number of blocks of original images to establish a particular model to predict the quality of the decompressed image. It can also be used to set appropriate parameters for providing the desired visual quality. However, the accuracy of this prediction method is not satisfactory at present. It is mainly used for rough classification. If it is directly applied to provide the desired visual quality for lossy compression, its accuracy needs to be improved. The iterative method can achieve ideal control accuracy at the cost of multiple compression/decompression. This number depends on the image content and the compression encoder and sometimes can be huge. This disadvantage limits its use in platforms with high real-time requirements or limited computing resources.

In view of the fact that none of the current methods can satisfactorily provide the desired visual quality in lossy compression, it is urgent to design an effective visual quality providing method; thus, the main objectives of the study are:

- Develop an effective method of predicting the visual quality of decompressed images for a given parameter, which will work with a broader range of lossy compressions, such as DCT-based and DWT-based coders.
- Develop a method of providing the desired visual quality in lossy compression, which will have low computational complexity, and adequate control accuracy.
- Develop an intelligent method of the auto setting parameter according to the desired visual quality, which will allow enhancing the primary method's robustness by referring to some of the image characteristics.
- Develop a set of practical recommendations on the remote sensing images lossy compression, based on the influence of distortion on the classification accuracy.

CHAPTER 2

RESEARCH AND DEVELOPMENT OF PRIMARY METHOD OF PROVIDING THE DESIRED VISUAL QUALITY IN LOSSY COMPRESSION

This section studies and analyzes existing image quality metrics, and the degree of consistency of quality measures with subjective assessments of the created databases of test images is assessed. The performances of the existing popular lossy compression coders are also compared and analyzed, according to decompressed image visual quality under the same CR. Further, the method of predicting the visual quality in lossy compression is improved for a given parameter. In addition, the primary two-step compression method of providing the desired visual quality is developed for grayscale images, where the HVS-based metrics are employed to evaluate the decompressed image quality.

2.1 Selection of appropriate visual quality metrics

In order to provide the desired visual quality in lossy compression, the decompressed image quality needs to be evaluated using appropriate visual evaluation metrics. Consequently, the first task in our study is selecting of visual quality metric to be employed.

The literature review in chapter 1 shows that the objective FR metric is preferred to lossy compression, and many evaluation metrics have already been proposed. These objective evaluation metrics aim to predict the visual quality of images through various computational models and are continuously optimized to reflect the fidelity of images better and maintain higher consistency with the evaluation results of the human visual system.

A commonly used method is to calculate the correlation between the value given by the objective evaluation metrics and subjective mean opinion scores (MOS) on a large dataset. The Spearman rank-order correlation coefficient (SROCC) is a commonly used measure of the method's prediction monotonicity[143]. There are other rank-order-based

measures, such as Kendall's rank-order correlation coefficient (KROCC), which are found to be highly consistent with the SROCC measure.

To figure out the appropriate metrics, we have tested test 48 existing popular FR metrics, and some of them have been discussed in chapter 1. The SROCC value of a metric on a database with N images is calculated as:

$$SROCC(O, S) = 1 - \left[\frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \right] \quad (2.1)$$

where d_i denotes the difference between the i -th image's ranks in the objective (O) and subjective (S) scores.

The testing was done on the image database TID2013[145], which contains 3000 test images obtained from 25 reference images, 24 types of distortions for each reference image, and 5 levels for each type of distortion. Each image is associated with a MOS in the range from 0 to 9. Three types of distortions are considered in our testing. They are introduced by lossy compression with the JPEG coder, with the JPEG2000 coder, and by compression of noisy images by the AGU or ADCT coders. These three distortions represent the types of distortions typically introduced by lossy compression. The results are shown in Table 2.1.

Table 2.1–The results of 48 FR metrics on TID2013 database in terms of SROCC

FR metric	SROCC	FR metric	SROCC	FR metric	SROCC
PSNR	0,9143	PSNRHMA	0,9667	ADD_SSIM	0,9561
MSE	-0,9143	FSIM	0,9606	MCSD	-0,9690
WSNR	0,9399	FSIMc	0,9616	MDSI	-0,9663
UQI	0,8245	IWSSIM	0,9405	UNIQUE	0,9025
SSIM	0,8933	ADM	0,9319	MSUNIQUE	0,9061
MSSIM	0,9345	GSM	0,9588	PSIM	0,9699
IFC	0,8747	IGM	0,9638	CVSSI	-0,9702
VIF	0,9477	SRSIM	0,9675	PSNRHMA_mod	0,9667
VIFP	0,9257	SFF	0,9553	DSI	-0,9297
MSVD	-0,4059	GMSD	-0,9706	CSSIM	0,9560
QILV	0,9019	ESSIM	0,9592	CSSIM4	0,9342
VSNR	0,9146	WASH	0,3886	SSIM4	0,9497
PSNRHVS	0,9645	VSI	0,9699	HaarPSI	0,9683
PSNRHVSM	0,9606	IQM2	0,9592	RVSIM	0,9480
CWSSIM	0,8624	DSS	0,9529	NQM	0,9042
PSNRHA	0,9658	ADD_GSIM	0,9962	RFSIM	0,9433

Larger absolute SROCC values indicate a closer relation with the human subjective evaluation. From Table 2.1, it can be seen that some metrics have a very high correlation coefficient in terms of SROCC, such as PSNR-HVS, PSNR-HVS-M, PSNR-HMA, FSIM, FSIMc, MDSI, etc. They have scores higher than 0.96, which means these metrics can describe the quality of decompressed images very well. Besides, the analysis also shows that the HVS-based metrics and the multi-strategy metrics have about the same performance.

In view of the experiment results, we have a motivation to employ the following metrics to evaluate the decompressed image quality in this study:

- 1) PSNR-HVS;
- 2) PSNR-HVS-M;
- 3) PSNR-HMA;
- 4) FSIM;
- 5) FSIMc;
- 6) MDSI.

In addition, the classic metric PSNR is also employed in some cases to be a reference for comparison.

2.2 Comparison of various lossy compression coders

Concerning the selection of the lossy compression coders, we have discussed it in chapter 1, section 1.4. Based on the result of the literature review, conclusions can be drawn that:

- 1) The DCT-based AGU series coders can provide a better compression ratio than JPEG by changing the partition method and reduce the blocking effect.
- 2) Compared with JPEG2000, the DWT-based SPIHT coder reduces the computational complexity.
- 3) The emerging BPG coder can provide visual quality comparable to the coder based on deep learning technology with a light algorithm.

Given the importance of visual quality evaluation in lossy compression encoder selection, we conduct some image-compression comparison experiments. Considering the decompressed image visual quality relates to the image content, particularly the complexity, three images are chosen as the typical example (shown in Fig.2.1).

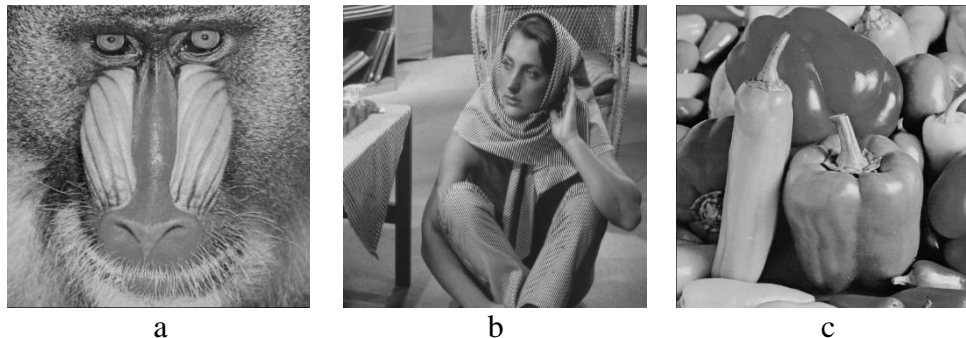


Figure 2.1 – Three typical images: a) Baboon, b) Barbara, c) Peppers

Among the three ones, the image Baboon has the highest complexity due to its high image texture. The image Barbara is moderately complex since it has highly textured regions as well as homogeneous areas. The image Peppers represents simple images for its large homogeneous areas and simple image structure. We conducted lossy compression on these images with the AGU, BPG, and JPEG coders, respectively. According to the analysis in section1, we have chosen three visual quality metrics, PSNR, PSNR-HVS-M, and FSIM to evaluate the decompressed images.

Table 2.2– Comparison of compression results of different coders

CR	Image	AGU			BPG			JPEG		
		PSNR	PSNR-HVS-M	FSIM	PSNR	PSNR-HVS-M	FSIM	PSNR	PSNR-HVS-M	FSIM
5	Baboon	33,305	39,010	0,990	34,245	41,020	0,992	29,140	47,609	0,994
	Barbara	42,848	49,340	0,998	43,618	51,560	0,998	37,920	53,465	0,998
	Peppers	41,416	47,027	0,997	42,850	49,774	0,998	37,710	56,523	0,998
10	Baboon	28,380	31,163	0,969	29,007	32,814	0,9734	25,403	34,616	0,974
	Barbara	37,710	40,964	0,992	38,688	42,919	0,9937	31,947	41,264	0,989
	Peppers	37,307	41,232	0,992	38,743	43,772	0,995	35,468	44,852	0,995
20	Baboon	25,233	26,058	0,936	25,322	26,813	0,937	22,824	24,700	0,918
	Barbara	33,318	34,132	0,977	34,162	35,665	0,981	26,900	30,941	0,954
	Peppers	34,828	37,105	0,985	36,081	39,453	0,989	32,957	35,825	0,979

The comparison of compression results corresponding to CR equal to 5, 10, and 20, respectively, is shown in Table 2.2. A general trend is that visual quality degrades as CR increases, regardless of the compression technique used and the complexity of the image

being compressed. It can be seen from Table that when the CR=5, the results obtained by the PSNR metric based on pixel error have a large deviation from the quality prediction results of the HVS-based PSNR-HVS-M metric and the multi-strategy-based FSIM metric. When the CR=10, this deviation becomes smaller, and when the CR=20, the results of the three are basically consistent. However, in any case, the results of PSNR-HVS-M and FSIM are consistent. Let us give an example of the image Baboon.

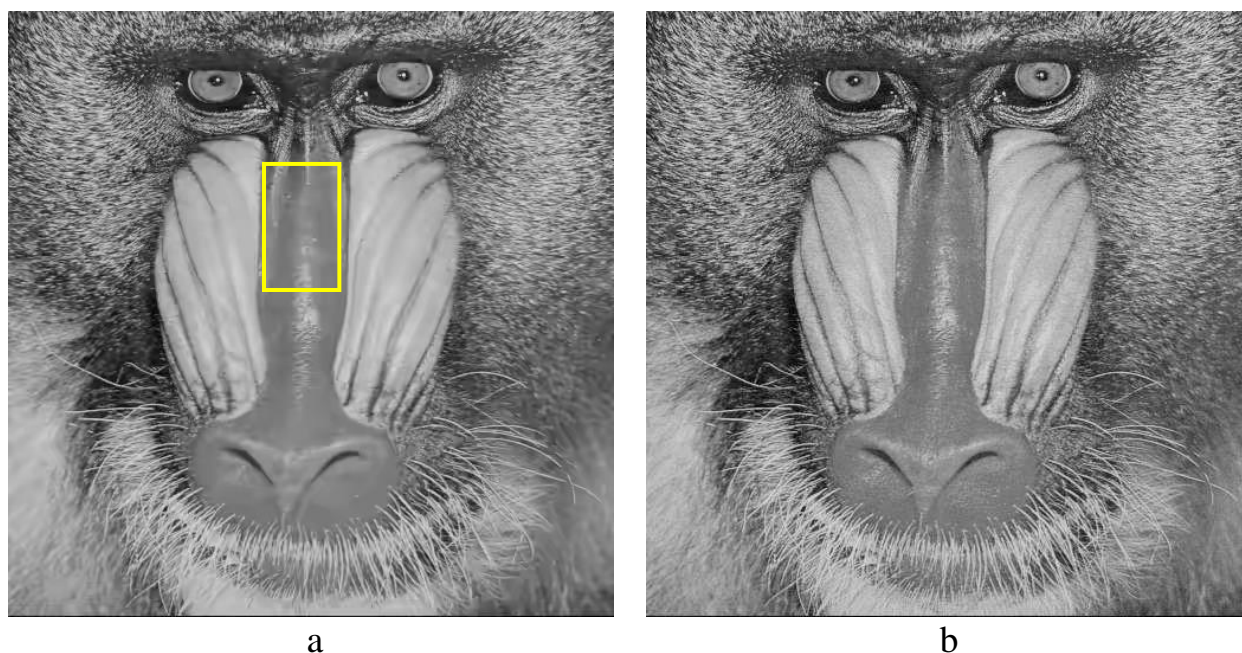


Figure 2.2 –Decompressed image Baboon, a) CR=10, BPG, b) CR=5, JPEG

Compared to the original image (shown in Fig 2.1.a), Fig 2.2.a introduces more distortion than Fig 2.2.b, accompanied by higher CR. The distortion can be noticed easily, e.g., the yellow rectangle area in Fig 2.2.a, while Fig 2.2.a is almost indistinguishable from the original. However, the evaluation score obtained by the PSNR metric is basically the same (29.14 dB and 29.007 dB). On the contrary, the evaluation scores obtained by PSNR-HVS-M and FSIM are largely consistent with human observations. Higher scores correspond to high-quality images. For PSNR-HVS-M, the scores are 32.814 and 47.608 dB, respectively, and the same is true for FSIM. This given evidence that the HVS-based metric and multi-strategy-based metrics are more reliable than the classic PSNR metric. It is necessary to employ them in distortion controlling for lossy compression.

Through the data in Table 2.2, we can also find that according to the results of PSNR-HVS-M and FSIM evaluation, a) in the case of small CR (5~10), the decompressed image quality of the JPEG coder is the best, and the distortions are invisible; b) With the

compression ratio increases, the advantages of the AGU and BPG coders gradually become significantly compared to the JPEG coder; c) when the compression ratio reaches 20, modern coders clearly outperform JPEG, the BPG coder obtains the best performance, followed by the AGU coder.

Based on these comparisons and observations, the following conclusions can be drawn:

- 1) the choice of visual quality metric is important in work that provides the desired visual quality value for lossy compression;
- 2) modern coders, such as AGU, BPG, etc., can get better performance than the classic compression coder JPEG at larger CR;
- 3) the image complexity also has a great impact on the quality of the decompressed image.

Choice of compression technique is determined by many factors, including priority of requirements as rate-distortion characteristics, simplicity of providing a desired quality or CR, computational efficiency, availability of portable device realizations or the corresponding platforms, etc.[40]. In this study, we investigated the modern lossy compression coder, AGU, ADCTC, SPIHT, and BPG to compress the general images, medical images, and remote sensing images. Among them, both the gray-scale images and RGB or three-channel images are considered.

2.3 Prediction of visual quality in lossy compression

The prediction method is one way to provide a desired quality of image compressed in a lossy manner since the CCP setting can be taken based on this prediction. It employs pre-established dependences of a given metric on CCP based on some approximation, and no iteration is involved in the entire process[101, 173]. However, the accuracy of quality providing is worth improving.

Recently, a group of methods has been introduced[173, 174], which obtain and “analyze” statistics in a limited number of blocks in the DCT domain. However, these methods are mainly intended for the prediction and providing of such quality metrics as

MSE or PSNR that are known to be inadequate for the visual quality of distorted images. Therefore, the prediction method needs more reliable evaluation metrics to test. In addition, it needs to be tested on other compression encoders to expand its application breadth.

2.3.1 Prediction of visual quality in lossy compression for medical images

Concerning lossy compression of medical images, the so-called visually lossless compression is often required. This means that the introduced distortions should be invisible to enable image diagnostic value.

In visually lossless compression, there are several important requirements to be satisfied. First, to be reliable, compression should be based on using adequate visual quality metrics and invisibility thresholds [163]. Second, a desired value of a used visual quality metric has to be provided quickly enough and with appropriate accuracy. Since, in this paper, we consider grayscale images, it is possible to use PSNR-HVS-M [163], which is one of the most reliable visual quality metrics and which has a priori known threshold of distortion invisibility approximately equal to 42 dB [163]. Besides, as a particular case, we consider the coder AGU based on DCT[108] that performs better than JPEG and slightly better than JPEG2000.

1) Basic dependences for image compression by AGU

In order to evaluate the distortion introduced by lossy compression more comprehensively, three evaluation metrics are employed. One is standard PSNR, and two other ones are PSNR-HVS and PSNR-HVS-M[163], which are defined as equation (1.23) (1.24) (1.25), respectively. Based on their computational model, there is a certainty of behavior that PSNR, PSNR-HVS, and PSNR-HVS-M have equal values if distortions are like additive white Gaussian noise and the masking effect is absent.

The dependences obtained for 512x512 pixel test image MRT_{prepared} are presented in Fig. 2.3.a. Analysis of the plots allows concluding the following. For $QS < 7$, all metrics are larger than 40 dB (PSNR-HVS-M exceeds 50 dB). This means that introduced distortions are invisible. In [174], the following approximation has been introduced for small QS:

$$MSE_{HVS-M} = 0,02896 \times QS^{1.976} \quad (2.2)$$

where MSE_{HVS-M} is calculated as a mean value of local estimates of in blocks, the number of such local estimates is about $I_m J_m / 64$, I_m and J_m denote an image size. For the traditional approach, all possible block positions have to be taken into account and have to be calculated after image compression and decompression by a given coder.

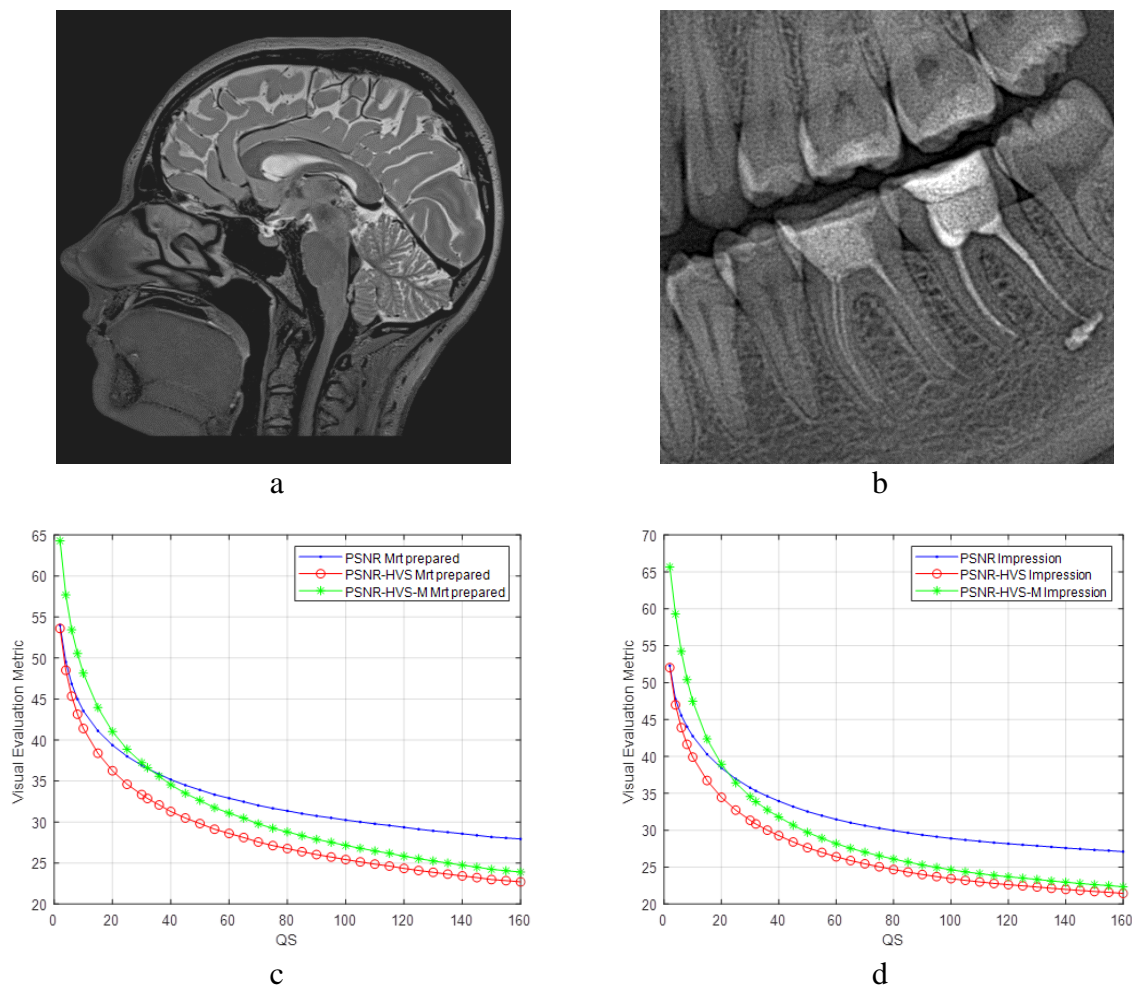


Figure 2.3 - Medical test images and dependences of quality metrics on QS, a) image Mrt_prepared, b) image dental, c) dependence curve of the image Mrt_prepared, d) dependence curve of the image dental

The interval of larger QS (from 7 to approximately 30 corresponds to MSE_{HVS-M} from 1 to about 30, Fig. 2.3) relates to the most important practical cases. When QS increases, all metrics decrease (Fig. 2.3.c). PSNR-HVS becomes sufficiently smaller than PSNR for the same QS.

Fig. 2.4, taken from [174] with the fitted approximation curve [175], shows the scatter-plot for which each point corresponds to determine one test image compressed with a given QS. As it is seen, values for the same QS can be quite different, especially for $QS > 30$, where values can differ by several times.

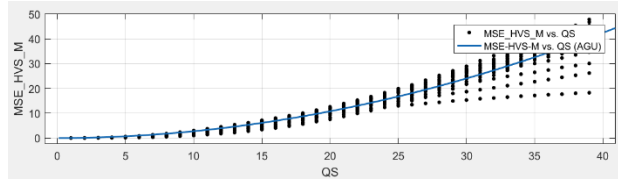


Figure 2.4 - Scatter-plot of MSE_{HVS-M} on QS for lossy compression of grayscale images for the coder AGU

This means that predictions like (2.2) can lead to errors for particular images that can be too large, and a more accurate prediction is needed. An approach is to use image statistics in a limited number of image blocks[28].

2) Prediction method based on preliminary analysis

In this study, MSE_{HVS-M} is calculated as an estimate before compression using a limited number of blocks of size 8×8 pixel. It is based on the assumption that there is a correlation between MSE_{HVS-M} for AGU coder and MSE_{HVS-M} determined in a limited number of 8×8 pixel blocks. The choice of 8×8 pixel blocks aims to carry out fast processing since it has fast hardware and software realizations[176].

First, N 8×8 pixel blocks are chosen. Then, 2D DCT is calculated for each block by getting $D(k,l,n)$ as follows.

$$D_{dq}(k,l,n) = QS * ([D(k,l,n)/QS]) \quad (2.3)$$

where $[\cdot]$ denotes rounding-off to the nearest integer. $k=0, \dots, 7$ and $l=0, \dots, 7$ are spatial frequency indices, and $n=1, \dots, N$ is a block index.

Second, $MSE_{HVS-M}(n)$ is calculated using the Table of spatial frequency weights and masking rule (see [13] for more details). Finally, the average $MSE_{HVS-M\Sigma}$ is calculated for all blocks, and MSE_{HVS-M} for AGU is predicted.

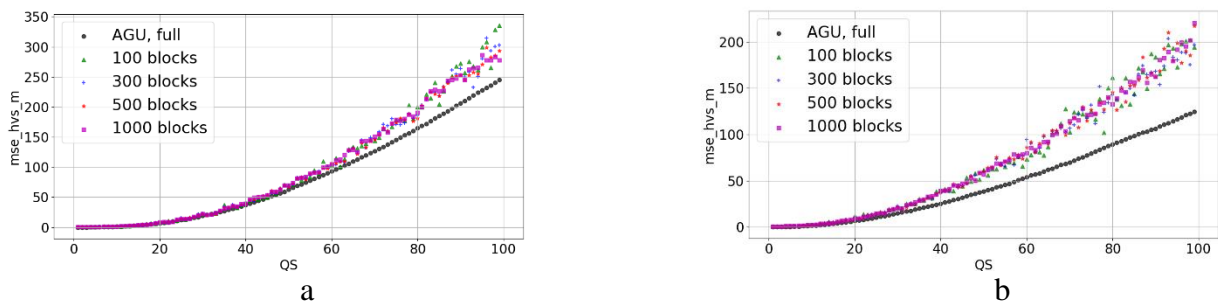


Figure 2.5– Scatter-plots of estimates of $MSE_{HVS-M\Sigma}$ for different number of analyzed blocks and MSE_{HVS-M_AGU} for images compressed by AGU for Baboon (a) and Barbara (b)

The scatterplots are presented for two more typical test images, Baboon (Fig.2.1.a) and Barbara (Fig.2.1.b), where the former is highly textural, and the latter is the middle complexity image (Fig.2.5).

For each QS and each N, we have calculated the mean and standard deviation of the ratio $S = \text{MSE}_{\text{HVS-M}\Sigma} / \text{MSE}_{\text{HVS-M}_{\text{AGU}}}$. The obtained data are presented in Fig. 2.6.

There is an interval of QS from 1 to approximately 7 where the mean of S is quite large (considerably larger than unity) with an obvious tendency to decrease quickly. Standard deviation values are large in this interval too. This effect cannot be seen in Fig.2.5 since although the values of $\text{MSE}_{\text{HVS-M}\Sigma}$ are larger than $\text{MSE}_{\text{HVS-M}_{\text{AGU}}}$, they are still very small. If $\text{QS} > 7$, then the ratio has stable values close to 1.34. This ratio does not depend upon the number of analyzed blocks. This means that the predicted $\text{MSE}_{\text{HVS-M}_{\text{AGU}}}$ can be calculated easily as

$$\text{MSE}_{\text{HVS-M}_{\text{pred}}} = \text{MSE}_{\text{HVS-M}\Sigma} / 1,34 \quad (2.4)$$

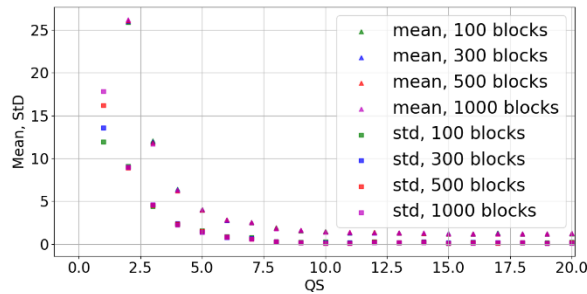


Figure 2.6 - Mean and standard deviations of the ratio $\text{MSE}_{\text{HVS-M}\Sigma} / \text{MSE}_{\text{HVS-M}_{\text{AGU}}}$ for all test images compressed by AGU

Based on the above analysis, an algorithm can be proposed as follows.

- 1) calculate initial QS as $\text{QS}_{\text{init}} = (\text{MSE}_{\text{HVS-M}_{\text{des}}} / 0.02896)^{1/2}$;
- 2) if $\text{QS}_{\text{init}} \leq 7$, use this QS_{init} as the final quantization step QS_{fin} for compression;
- 3) if $\text{QS}_{\text{init}} > 7$, then calculate $\text{MSE}_{\text{HVS-M}\Sigma}$ using a limited number of blocks and recalculate it to $\text{MSE}_{\text{HVS-M}_{\text{AGU}}}$;
- 4) compare this $\text{MSE}_{\text{HVS-M}_{\text{AGU}}}$ to $\text{MSE}_{\text{HVS-M}_{\text{des}}}$; if accuracy is satisfactory (e.g., $\text{MSE}_{\text{HVS-M}_{\text{AGU}}}$ differs from $\text{MSE}_{\text{HVS-M}_{\text{des}}}$ by less than $\epsilon\%$), then use QS_{init} as QS_{fin} ; otherwise, determine QS_{fin} as

$$QS_{\text{fin}} = QS_{\text{init}} \left(\frac{MSE_{HVS-M_{\text{des}}}}{MSE_{HVS-M_{AGU}}} \right)^{1/2} \quad (2.5)$$

where it is supposed that $MSE_{HVS-M_{AGU}}$ is approximately proportional to QS^2 .

3) Verification results

We have applied the proposed algorithm to nine test images for two values of $MSE_{HVS-M_{\text{des}}}$, namely equal to 4, which corresponds to the invisibility of distortions, and equal to 10, which relates to visible but not annoying distortions. The obtained data are presented in Tables 2.3 and 2.4, respectively. As it follows from the analysis of data in Table 2.3, $MSE_{HVS-M_{\Sigma}}$ and $MSE_{HVS-M_{AGU}}$ derived for QS_{init} are mostly considerably smaller than $MSE_{HVS-M_{\text{des}}}$. This means that the approximation (2.1) is not quite accurate for small QS . Thus, correction like (2.4) is needed. After correction, the provided MSE_{prov} is around $MSE_{HVS-M_{\text{des}}}$, some values of MSE_{prov} are larger than $MSE_{HVS-M_{\text{des}}}$ (this happens for complex structure images), and some are smaller (this takes place for simple structure images). Note that CR values vary in very wide limits from 4...5 for complex structure images to almost 19 for simple structure ones.

Table 2.3 – Data for $MSE_{HVS-M_{\text{des}}}=4$

Image	QS_{init}	<i>MSE_{HVS-MΣ}</i>	<i>MSE_{HVS-M_{AGU}}</i>	QS_{fin}	<i>MSE_{prov}</i>	CR
Aerial	12	2,014	1,503	20	6,195	5,75
Airfield	12	2,431	1,814	18	6,129	4,26
Baboon	12	2,041	1,523	19	5,769	4,50
Barbara	12	3,653	2,726	15	3,505	8,70
Diego	12	1,840	1,373	20	6,790	4,50
Frisco	12	4,314	3,220	13	3,482	18,5
Goldhill	12	2,744	2,047	17	5,648	8,60
Lenna	12	3,648	2,722	15	4,254	12,5
Mrt _{prepared}	12	2,381	1,777	18	4,130	14,9

Table 2.4 – Data for $MSE_{HVS-M_{\text{des}}}=10$

Image	QS_{init}	<i>MSE_{HVS-MΣ}</i>	<i>MSE_{HVS-M_{AGU}}</i>	QS_{fin}	<i>MSE_{prov}</i>	CR
Aerial	19	6,136	4,579	28	14,33	7,72
Airfield	19	7,984	5,958	25	13,79	5,61
Baboon	19	6,626	4,945	27	14,27	5,95
Barbara	19	7,937	5,923	25	10,38	13,2
Diego	19	5,785	4,317	29	18,50	6,34
Frisco	19	8,292	6,188	24	9,132	30,8
Goldhill	19	9,066	6,765	23	11,24	12,1
Lenna	19	8,949	6,678	23	8,909	19,2
Mrt _{prepared}	19	6,126	4,572	28	10,82	21,2

Consider now the data in Table 2.4. Again, $MSE_{HVS-M\Sigma}$ and MSE_{HVS-M_AGU} for QS_{init} are considerably smaller than $MSE_{HVS-M\ des}$. After correction, they become to be around $MSE_{HVS-M\ des}$, although some of them occur to be too large as for the test image Diego. Since more distortions are introduced, CR values have increased compared to data in Table 2.3.

Fig. 2.7 presents the test image $MRT_{prepared}$ used in our analysis with the marked fragments that should be paid attention to in the analysis. The enlarged green frame fragment for compression with providing PSNR=38 dB is presented in Fig. 2.7.b, where the introduced losses are invisible. The same fragment for compression with providing PSNR=36 dB is presented in Fig. 2.7.c, where the introduced losses become noticeable.

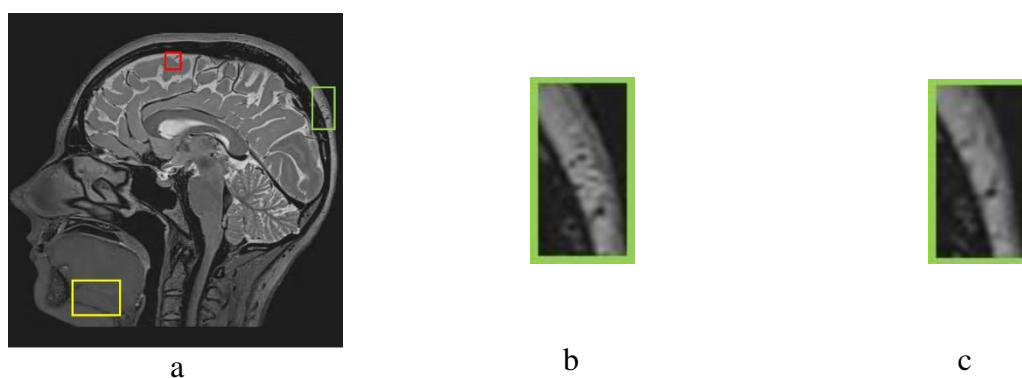


Figure 2.7 - The test image $MRT_{prepared}$, a) original image, b) compression with providing PSNR=38 dB, c) compression with providing PSNR=36 dB

2.3.2 Prediction of visual quality in lossy compression by SPIHT

Given the results achieved by AGU encoders for a range of visual metric prediction methods, we hope to generalize it to other encoders, such as DWT-based encoders.

The SPIHT coder is a typical representative of lossy compression techniques based on DWT, which was developed from the early EZW (embedded zero-tree wavelet) image compression technology. Compared with DWT-based JPEG2000, SPIHT has lower computational complexity, which makes it widely used in lossy compression. It is easy for SPIHT to provide a desired CR. In addition to CR characterizing the encoder and a compressed image, an end-user also needs to consider whether the compressed image quality meets the quality requirements, i.e., is a desired quality provided. The evaluation or

prediction of compressed image quality is essential for numerous applications, and many quality metrics can be used for this purpose[137, 177, 178].

There are two favorable obstacles that allow predicting visual quality for the SPIHT coder in an indirect but quite fast and accurate way. One obstacle is that the performance characteristics of SPIHT are quite similar to those ones [179] of the DCT-based coder AGU. Another obstacle is that several simple and fast procedures for predicting the performance of AGU have been proposed recently[40, 177]. Thus, our idea consists in exploiting experience in the prediction of AGU coder characteristics for the prediction of SPIHT performance parameters.

1) Comparison of the Dependences for SPIHT and AGU

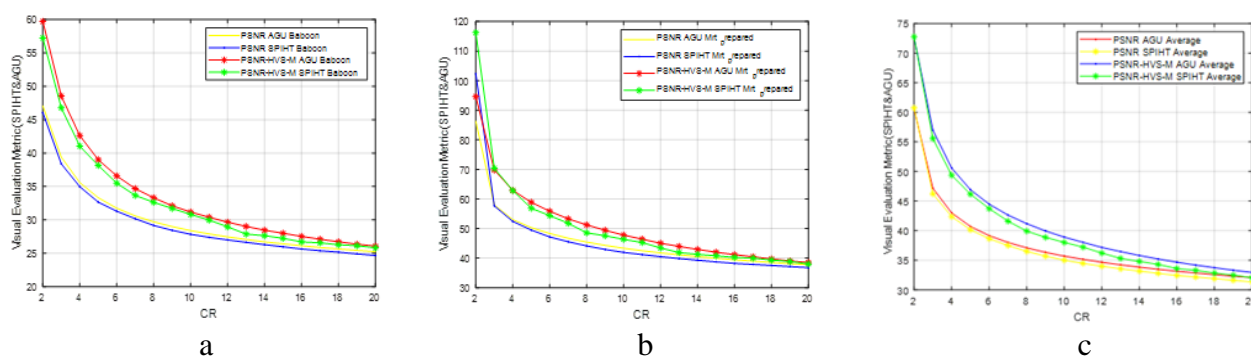


Figure.2.8 - Test images dependence curve a) high complexity image Baboon; b) low complexity image $MRT_{prepared}$; c) average dependences for AGU and SPIHT

Let us analyze rate-distortion curves obtained for two test images, the images Baboon (see Fig.2.2.a) and $Mrt_{prepared}$ (see Fig.2.3.a). They are presented in Figures 2.8.a and 2.8.b as dependences of PSNR and PSNR-HVS-M on CR. It is possible to see that there is a strict connection between the dependences for the considered coders. To get a better understanding, the average dependence curves of visual quality metrics on CR under two encoders are given in Fig. 2.8.c.

If the visual quality values of multiple images are averaged, the compressed visual quality dependence curves of SPIHT and AGU codes are still similar, and the deviation is basically fixed for each CR value. After calculation, the average deviation of PSNR (SPIHT with respect to AGU) is 0,6232 dB, and the average deviation of PSNR-HVS-M (SPIHT with respect to AGU) is 0,9337 dB. Then, the visual quality metrics value for SPIHT can be calculated as

$$PSNR_{SPIHT} = PSNR_{AGU} - 0,6232 \text{ dB} \quad (2.6)$$

$$PSNR\text{-HVS}\text{-}M_{SPIHT} = PSNR\text{-HVS}\text{-}M_{AGU} - 0,9337 \text{ dB} \quad (2.7)$$

2) Prediction method for SPIHT

By comparing the compression dependence curves of AGU and SPIHT, it can be expected that the prediction of SPIHT compression can be quickly realized by means of the prediction result of the decompressed image quality of AGU.

However, there is one key problem to be solved. The control parameter of SPIHT is bpp, but the PCC of AGU is QS. In addition, the predicted connection between AGU and SPIHT is based on the dependence of decompressed image quality on CR. Consequently, to predict the compressed image quality of SPIHT for a given bpp, one has to determine the corresponding QS for AGU.

For a given BPP, the first stage task is to convert it to CR. In SPIHT, CR and bpp are strictly dependent where on 8-bit grayscale images. The conversion can be achieved by the following equation[180].

$$CR \approx 8/\text{bpp} \quad (2.8)$$

The second stage is the conversion from CR to QS (used CCP in AGU). It is not an easy work because it depends on the image at hand. As known, there is direct dependence between distortions due to DCT coefficient quantization and losses in compressed data. We will use the parameter P_{0q} , which denotes the mean probability that quantized DCT coefficients in 8x8 blocks are equal to zero [40, 177].

$$P_{0q} = (\sum_{n=1}^{N_{bl}} N_n) / 64N_{bl} \quad (2.9)$$

where N_n denotes the number of DCT coefficients to be zeroed after quantization (this happens if the DCT coefficient magnitude is less than $QS/2$) for an n-th block, N_{bl} denotes the number of the considered 8x8 pixel blocks randomly chosen in a considered image. The larger P_{0q} (more coefficients are assigned zero values), the greater CR. The function that approximates this relationship can be expressed as[177]:

$$CR = 0,9462 \exp(2,895P_{0q}) + 1,045 \times 10^{-13} \exp(35,52P_{0q}) \quad (2.10)$$

This means that P_{0q} can be predicted for a given CR or bpp. In [40], a method for predicting the PSNR for AGU is proposed. This method calculates the DCT coefficients in

8x8 pixel image blocks and quantizes them as $D_q(k,l,n)$ with (2.2). As one knows, standard PSNR is estimated as (1.22) (1.23). Besides, a quick and accurate MSE prediction approach has been proposed for AGU based on a given QS. Let us calculate the differences as

$$\Delta D_q(n, k, l) = QS \times D_q(n, k, l) - D(n, k, l), k = 0, \dots, 7; l = 0, \dots, 7 \quad (2.10)$$

and then the MSE for an n-th block can be predicted by the equation (2.11).

$$MSE_n = \frac{1}{64} \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(n, k, l))^2 \quad (2.11)$$

The MSE estimate for the entire image can be calculated as follows:

$$MSE = \frac{1}{N} \sum_{n=1}^N MSE_n = \frac{1}{64N} \sum_{n=1}^N \sum_{k=0}^7 \sum_{l=0}^7 (\Delta D_q(n, k, l))^2 \quad (2.12)$$

After that, the metric PSNR can be estimated according to (1.23). The coefficients $D_q(k,l,n)$ are collected for all considered blocks, and, thus, one can easily predict visual quality metrics for any QS.

Concerning the visual quality metric PSNR-HVS-M, our paper[40] shows how MSE_{HVS-M} can be quickly and quite accurately predicted for AGU (see also Section 2.3.1).

P_{0q} reflects the percentage of DCT coefficient values returned to zero after quantization. Known CR allows determining P_{0q} using (2.9). Of course, it is not easy to find the inverse function, but P_{0q} can be derived using piecewise linear interpolation. Then, one needs an algorithm for obtaining the QS corresponding to the determined P_{0q} and, respectively, CR for the considered image. One very simple version of this algorithm is shown in Fig. 2.9.

First, set an initial value of QS, e.g., equal to 0, and then gradually increase it until the percentage of DCT coefficients with absolute values smaller than $QS/2$ is smaller than P_{0q} . When this happens, record QS value. The fourth stage in the prediction process is to calculate the PSNR and PSNR-HVS-M values corresponding to the recorded the QS. After this, complete the entire prediction process of the visual quality of the SPIHT compression coder via AGU using expressions (2.6) and (2.7).

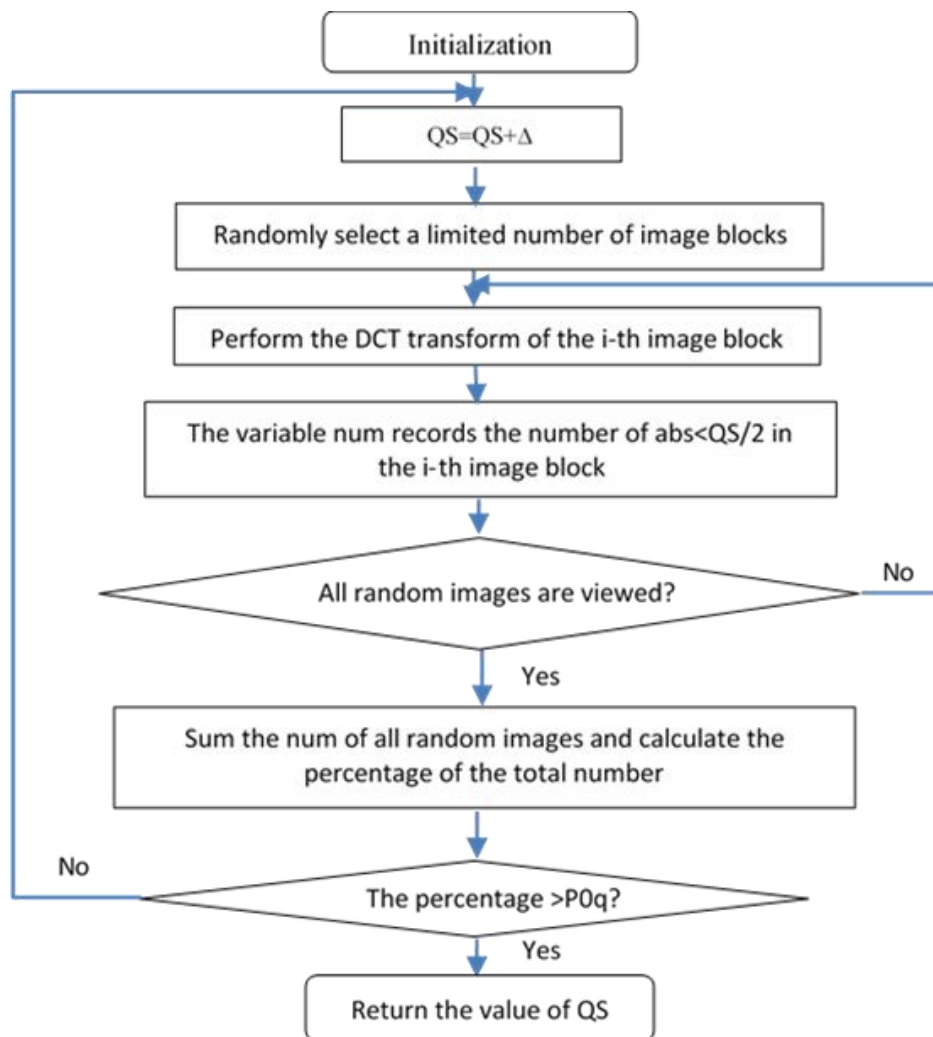


Figure 2.9- Flow chart for calculating QS

3) The prediction results and analysis

We have considered two values of bpp in our experiments, namely 0.5 and 1. To analyze prediction accuracy, two visual quality metrics, PSNR and PSNR-HVS-M, have been employed. In the experiment, we selected 300, 500 and 1000 as the number of random image blocks. The comparison has shown that the data accuracy for 500, and 1000 blocks has not improved compared to the case of 300 blocks, but the time consumption has increased because of the increase in the number of calculations. Therefore, in this paper, 300 is selected as the number of random image blocks for the experiment. In the experiment, we selected nine commonly used images (including highly complex texture images, moderately complex texture images, and simple texture images). The PSNR prediction experimental data are presented in Table 2.5, corresponding to the visual quality prediction data with bpp of 0,5 and 1, respectively.

Table1.5–Statistics of PSNR data in prediction for SPIHT

Test image	bpp=0.5				bpp=1			
	PSNR _{pre}	PSNR _{real}	Error	Time (s)	PSNR _{pre}	PSNR _{real}	Error	Time (s)
Goldhill	33,6655	33,1294	-0,5361	0,0977	36,2791	36,6047	0,3256	0,0828
Baboon	26,1117	25,6281	-0,4836	0,1964	30,7143	29,8422	-0,8721	0,1227
Lenna	37,0936	37,2353	0,1417	0,0859	40,778	40,4585	-0,3195	0,0629
Barbara	35,2759	32,0839	-3,192	0,1097	39,0924	37,4122	-1,6802	0,0689
Aerial	28,6229	28,7333	0,1104	0,1832	32,7749	33,1914	0,4165	0,1079
Airfield	29,0245	27,4969	-1,5276	0,1558	31,9097	30,3014	-1,6083	0,1056
Frisco	42,8915	42,6874	-0,2041	0,0598	48,3571	47,7694	-0,5877	0,0638
Diego	26,5196	26,6317	0,1121	0,1912	29,4761	29,4905	0,0144	0,1250
Mrt _{prepared}	39,4486	38,2106	-1,238	0,0869	47,7793	44,1214	-3,6579	0,0578
SSD			1,0274				1,2120	

In this Table, PSNR_{pre} represents the PSNR value of the image processed by the SPIHT encoder calculated by the program, and PSNR_{real} is the PSNR value of the image actually compressed/decompressed by SPIHT. The error refers to the deviation between the predicted value and the actual value, and time refers to the time spent in the prediction process. In order to statistically analyze the accuracy of the prediction, the SSD (sample standard deviation) calculation was performed for each group of errors as

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2} \quad (2.13)$$

Table2.6–Statistics of PSNR-HVS-M data in prediction for SPIHT

Test image	bpp=0.5				bpp=1			
	PSNR-HVS-M _{pre}	PSNR-HVS-M _{real}	Error	Time (s)	PSNR-HVS-M _{pre}	PSNR-HVS-M _{real}	Error	Time (s)
Goldhill	36,2594	33,9034	-2,3560	0,2408	40,8164	40,1590	-0,6574	0,2105
Baboon	27,8017	26,6977	-1,1040	0,3336	35,0034	32,6220	-2,3814	0,2583
Lenna	39,0606	39,5323	0,4717	0,2125	44,1417	44,9048	0,7632	0,1926
Barbara	36,4551	33,6409	-2,8142	0,2305	42,6237	40,5339	-2,0898	0,2024
Aerial	30,2541	29,3158	-0,9383	0,3006	37,7715	36,6634	-1,1081	0,2786
Airfield	31,3978	29,1119	-2,2859	0,2706	36,7416	33,5935	-3,1481	0,2538
Frisco	42,9987	42,5198	-0,4789	0,2005	46,7146	48,9363	2,2217	0,1926
Diego	28,2534	27,4636	-0,7898	0,3229	33,7246	33,3702	-0,3544	0,2790
Mrt _{prepared}	42,7600	40,2362	-2,5238	0,1905	44,8965	48,5444	3,6479	0,1584
SSD			1,0531				2,0946	

From the analysis of data in Table 2.5, it can be seen that the prediction accuracy of this method is good, the maximum deviation is about 3,66 dB, and the overall distribution

of errors is reasonable. The sample standard deviations based on nine images are 1,0274 dB and 1,212 dB, respectively. Thus, accuracy is better than for the two-step method[181].

The data on PSNR-HVS-M are given in Table 2.6. Here $PSNR-HVS-M_{pre}$ is the PSNR-HVS-M value of the image processed by the SPIHT encoder calculated by the program, and $PSNR-HVS-M_{real}$ is the PSNR-HVS-M value of the image actually compressed/decompressed by SPIHT.

The deviation of PSNR-HVS-M when bpp is 0,5 is slightly larger than that one for bpp is 1. The maximum deviation is 3,65 dB, and the sample standard deviations are 1,0531 dB and 2,0946 dB, respectively. At the same time, because the calculation of PSNR-HVS-M involves DCT transformation and DCT inverse transformation, the time consumed increases, about two times the PSNR prediction. From the statistical data of PSNR and PNSR-HVS-M, this prediction method can predict SPIHT compressed image quality very well on the basis of AGU, and the accuracy has been improved. Table 2.7 presents data for analysis of time efficiency. In order to more objectively understand the time efficiency of the SPIHT visual quality prediction method, the average prediction time of 9 test images is used as a measurement factor in Table 2.7.

Table 2.7–Average time consumption(s) comparison (PSNR and PSNR-HVS-M)

bpp value	Time_{pre}	Time_{Two-step}	Time_{SPIHT}
0,5	0,2558	0,8937	0,3914
1	0,2251	0,9220	0,3005

In Table 2.7, two sets of data were measured for bpp of 0,5 and 1, and the prediction and calculation of PSNR and PSNR-HVS-M two visual quality metrics systems were conducted. $Time_{pre}$ refers to the time required to obtain the visual quality value through the prediction method; $Time_{two-step}$ represents the time required in the first step of compression/decompression through the two-step method and obtains the visual quality value; $Time_{SPIHT}$ is the time required to SPIHT compressed images.

As can be seen from the data in Table 2.7, the newly proposed prediction method has greatly improved the time efficiency because it avoids the compression/decompression step. The average time required is 1/3 of the two-step method and is less than the time required for the actual compression process of SPIHT.

2.4 The primary two-step compression method for providing the desired visual quality

Given the low time efficiency of iterative methods and the unsatisfactory accuracy of prediction methods, an intelligent and efficient method to provide the desired visual quality for lossy compression is demanded. It should have low computational complexity and take into account some aforementioned requirements to lossy compression. Besides, it is also required to have acceptable control accuracy and ubiquity on existing popular lossy encoders, as well as compatibility with various advanced visual quality evaluation metrics. To address this problem, an original two-step compression method that takes into account both time efficiency and accuracy has been proposed in[42, 181].

The primary two-step method[42] is implemented as follows: Through the initial (first step) compression with CCP determined on the basis of a priori data, a more reasonable visual quality metric value is obtained, and the CCP value is then corrected to perform the final compression that, on the average, ensures that visual quality satisfies the user's requirement better. However, the drawbacks of this approach are not studied and explained. The case of visual quality metrics is not considered at all.

In this section, we apply this primary two-step compression method to AGU and BPG encoders, considering various visual quality evaluation metrics when evaluating decompressed image quality. The performance of the basic method in the application is deeply analyzed.

2.4.1 The primary two-step approach to providing a desired visual quality in AGU lossy compression

As a case study, we analyze the compression method for AGU[108]based on DCT, for which the varied CCP is the quantization step (QS). The metrics PSNR, PSNR-HVS, and PSNR-HVS-M [38, 150]are employed to evaluate the decompressed image quality. The two latter ones are applicable when the end-user of the image-based system is a

human being. The experiments are conducted on the gray-scale images. Its accuracy and performance are tested from different aspects.

1) The fundamental of primary two-step method in lossy compression

The two-step method is proposed based on the average distortion curve for trend prediction. In fact, the average distortion curve provides preconditions for calculating the initial (rough) value of CCP. The method of obtaining the average distortion curve is explained below.

First, some images are chosen as a basic image set, which should be typical for different image application fields to make the average distortion curve trend more general.

Second, when obtaining the average distortion curve, it is necessary to first set QS to a series of values for each image and obtain the corresponding value of the visual quality metric. Then, one averages data for each QS point and, thus, gets the average rate/distortion curve. In this way, the average distortion curves have been obtained for the aforementioned three visual quality metrics. The results are shown in Fig.2.10.

It can be found that no matter for PSNR, PSNR-HVS, or PSNR-HVS-M, the changing trend of the distortion curves is very similar. Based on this characteristic, the average rate/distortion curve (obtained in advance offline) is utilized to obtain the initial value of PCC and perform the first compression. The metric value obtained by the first compression will be close to the desired value, but the error is uncertain (different pictures have different errors; of course, this can also be observed in Fig. 2.10). Then, the second compression is performed.

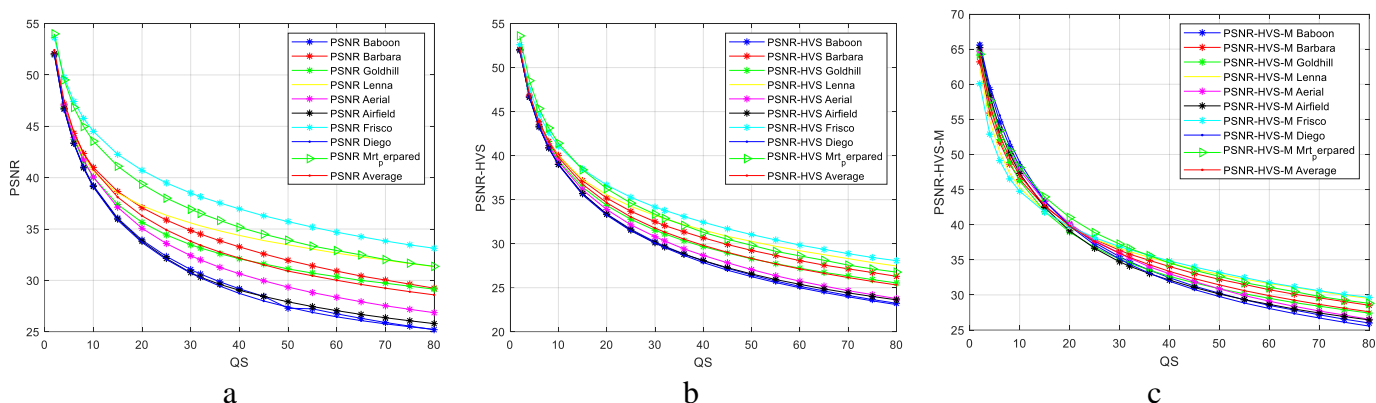


Figure. 2.10 -The average rate/distortion curves of AGU coder for gray-scale images, a) PSNR, b) PSNR-HVS, c) PSNR-HVS-M

Third, the CCP needs to be corrected before the second compression step. This correction is related to the error obtained by the first compression. The following formula can be used in the parameter correction.

$$QS_{des} = QS_{init} + \frac{PSNR_{des} - PSNR_{init}}{M'} \quad (2.14)$$

where QS_{init} is the initial value of QS (rough value from average distortion curve), $PSNR_{init}$ is the PSNR value obtained after the first step compression, $PSNR_{des}$ is the desired value of PSNR, M' is the average distortion curve derivative corresponding to QS_{init} . After linear estimation, a QS_{des} value closer to the desired value can be obtained for each image. In the second step, compression is performed with the QS corrected for the image as a parameter.

The calculation methods for the other two visual quality metrics are the same and are executed with the help of the following formulas.

$$QS_{des} = QS_{init} + \frac{PSNR-HVS_{des} - PSNR-HVS_{init}}{M'} \quad (2.15)$$

$$QS_{des} = QS_{init} + \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{init}}{M'} \quad (2.16)$$

By correcting QS once, a large increase in accuracy is expected to be achieved in the second compression step.

2) Experiment implemented for AGU coder

To better understand how this method works, we set desired values of our metrics equal to 30, 35, and 40 dB in our experiments. They have been carried out for PSNR, PSNR-HVS, and PSNR-HVS-M. Detailed data for one of the sets of tests are shown in Table 2.8.

Table 2.8–Statistics and parameters of providing $PSNR-HVS_{des}=30$ dB

Test Image	QS_{init}	$PSNR-HVS_{init}$	ΔQS	QS_{rec}	$PSNR-HVS_{prov}$
Goldhill	38,9283	29,8194	-1,0614	37,8669	30,0071
Baboon	38,9283	28,2555	-9,7143	29,214	30,4306
Barbara	38,9283	30,8412	4,6843	43,6126	30,1307
Lenna	38,9283	31,7084	9,5133	48,4416	30,4254
Aerial	38,9283	28,8464	-6,4239	32,5044	30,1949
Airfield	38,9283	28,2629	-9,6731	29,2552	30,3089
Frisco	38,9283	32,6046	14,5038	53,4321	30,605
Diego	38,9283	28,0752	-10,718	28,21	30,5263
Mrt _{prepared}	38,9283	31,5009	8,3578	47,2861	30,1604
Variance		2,9739			0,0401

The variances after the first compression and the second compression are calculated, respectively. From the results of this set of data, after the second step of compression, the variance is greatly reduced, and the error is well controlled. For PSNR and two HVS-based metrics for three desired levels of visual quality, we conducted nine sets of experiments, and the statistical information of the experimental data is shown in Table 2.9.

Table 2.9–Statistics obtained for nine test images

Quality Metric	M_{des}	VAR_{fis}	VAR_{sec}	$MAX\Delta_{final}$
PSNR	40	4,4876	0,4714	1,2125
PSNR	35	7,1179	1,6761	3,6067
PSNR	30	9,1183	9,0517	5,2102
PSNR-HVS	40	0,7906	0,0408	0,3583
PSNR-HVS	35	1,6575	0,0369	0,6017
PSNR-HVS	30	2,9739	0,0401	0,605
PSNR-HVS-M	40	0,3675	0,0108	0,257
PSNR-HVS-M	35	1,0119	0,1041	0,720
PSNR-HVS-M	30	2,0312	0,0280	0,465

Here M_{des} is the desired value of the considered visual quality metric, VAR_{fis} is the variance of visual quality metric for nine test images obtained after the first-stage compression, VAR_{sec} is the variance of visual quality metric obtained after the second (correcting) step of compression, $MAX\Delta_{final}$ is Maximum error between M_{des} and the provided values M_{pro} .

3) Analysis of the results

It can be seen from Table 2.8 that after the two-step approach, the variances tend to reduce, and the data become more stable than the ones in the first step; the error between the desired visual quality ($PSNR-HVS_{des}$) and the provided visual quality ($PSNR-HVS_{prov}$) does not exceed 0,605 dB, which is a sufficiently good accuracy. Compared with the analysis of PSNR results in the paper [182], for HVS-metrics, the provided accuracy does not depend on the complexity of a compressed image.

In Table 2.9, the results show that the two-step approach has generally improved the accuracy of providing the value of M_{des} for the metrics PSNR-HVS and PSNR-HVS-M. After two steps, the variance has reduced to about 1/10 to 1/70 of the one-step variance, indicating that the QS correction has played a positive role. The maximum variance is 0,1,

indicating that data fluctuations are small and the results are relatively stable. Meanwhile, the maximum error is 0,605 dB (for PSNR-HVS) and 0,720 dB (for PSNR-HVS-M).

In order to analyze the performance of the visual quality metrics values provided by this method, a comparison experiment has been carried out. We got three images for the test image Goldhill for three values of QS, namely QS=30, 35, and 40. The compressed images are shown in Fig.2.11, where metrics' values are given as well.



a. original image



b. QS=30, PSNR=33,4639 dB, PSNR-HVS=31,5263 dB, PSNR-HVS-M =35,0879 dB, CR =16,7483



c. QS=35, PSNR=32,7341 dB, PSNR-HVS=30,5316 dB, PSNR-HVS-M =33,7804 dB, CR =20,2303



d. QS=40, PSNR =32,1059 dB, PSNR-HVS=29,6638 dB, PSNR-HVS-M =32,6397 dB, CR =23,9576

Figure. 2.11– Comparison of images for, a) original image, b) QS=30, c) QS=35,d) QS=40

A comparison of the two images below (for $QS=30$ and 35) shows that the difference values for the corresponding metrics are about 1 dB (the difference is the smallest for PSNR and the largest for PSNR-HVS-M). For this case, slight differences in compressed images can be noticed, in particular, for fragments marked by green rectangles.

Comparing the images for $QS=30$ and $QS=40$ (Fig.2a and Fig.2c), larger differences in metrics' values are observed (they are about 1,36 dB for PSNR, about 1,86 dB for PSNR-HVS, and about 2,45 dB for PSNR-HVS-M). More obvious differences (shown by blue marks) can be noticed. Then, it can be concluded that if PSNR (or PSNR-HVS or PSNR-HVS-M) changes by more than 0,5...1 dB, changes in visual quality can be noticed.

Thus, these differences can be used as an accuracy threshold. When the error of providing a desired quality is less or approximately equal to 0.5dB or less, the difference is more difficult to observe.

Recalling the results in Table 2.8 and 2.9, we can conclude that such errors are acceptable. It can also be considered that the proposed method is able to provide image quality close to a desired value with almost no error (difference) for visual quality metrics PSNR-HVS and PSNR-HVS-M. Meanwhile, the results for $PSNR_{des}=30$ dB show that there are some problems with the two-step approach in this case that should be studied more thoroughly.

The use of the proposed method for DCT-based lossy compression is shown to be able to provide a desired visual quality of compressed images, enabling faster and better compression.

2.4.2 A two-step approach to providing a desired visual quality in BPG lossy compression

In [42], it is proved that the two-step method is able to provide the desired visual quality for the AGU coder By referring to the prior knowledge of the average rate-distortion curve. However, different compression codes are implemented based on different principles. Consequently, it is necessary to verify whether the two-step method is suitable for other encoders.

BPG encoder is a new compression technique that outperforms JPEG in terms of compression quality and size of compressed images [112]. Because of its superior performance, the encoder has quickly gained popularity, especially in portable devices and online encoding applications. In view of this, it is worthwhile to offer an opportunity of providing a desired visual quality of images compressed by this encoder with high accuracy. In this section, we propose to use a two-step method in BPG compression to improve the accuracy of visual quality providing. In BPG, the quantizer parameter Q is used as the main control compression parameter [113], and it is also used as a CCP in the two-step method.

To test the applicability of the two-step method in the new BPG encoder, we analyze the accuracy of experimental data and study the feasibility of providing a desired visual quality in the BPG encoder.

1) Two-step method on BPG coder

The two-step compression method is implemented based on the average rate/distortion curve obtained off-line, which reflects the dependence of visual quality on CCP. In the BPG encoder, the quantization parameter Q is used as CCP, here $1 \leq Q \leq 51$. A larger Q results in a higher CR, corresponding to lower visual quality.

In order to obtain the average rate-distortion curve, it is necessary to collect a certain amount of off-line image set data and average the visual quality metric values corresponding to certain CCP values of the image set so as to obtain the dependence of the average value on the CCP. Twelve images have been selected by us as the sample image set, including general images and texture ones. Fig. 2.12.a and Fig. 2.12.b show the dependences of PSNR and PSNR-HVS-M on Q , respectively.

The curve in Fig.2.12.a shows that the PSNR values for $PSNR > 40$ dB approximately coincide for almost all images (except Frisco), and in the $30\text{dB} < PSNR \leq 40\text{dB}$ interval, the image curves are nearly parallel, in particular, for the interval $30\text{dB} < PSNR < 35\text{dB}$. The average distortion curve and each image curve obtained on the basis of all data are also approximately parallel. In general, the degree of dispersion between the curves is smaller than for AGU[42]. Nevertheless, PSNR values for the same Q can differ considerably, by up to 7 dB. This causes problems in providing a desired quality.

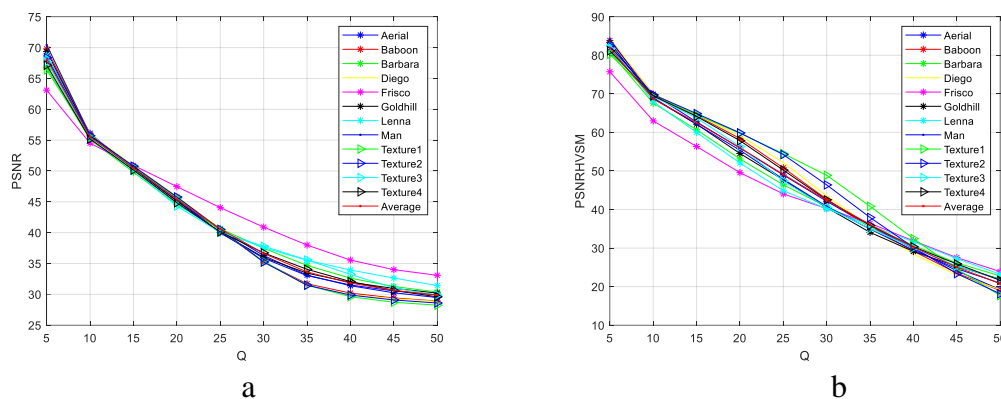


Figure. 2.12 – Visual quality metrics dependence on Q in BPG coder, a) PSNR, b)PSNR-HVS-M

The dependences of PSNR-HVS-M on Q show the same trend, but the curves go in a more compact manner than the dependences of PSNR. Compared with the AGU coder in [42], it can be found that the visual quality linearity in BPG coder is better than the one in AGU, and the CCP is the control parameter Q the intention of which is to provide a desired quality. Meanwhile, PSNR-HVS-M values for the same Q can differ by up to 10 dB. Recall here that, e.g., PSNR-HVS-M=35dB relates to obvious distortions, while for PSNR-HVS-M=42 dB, it is practically guaranteed that distortions are invisible[163]. Thus, the use of a given Q can lead to the different visual quality of compressed images.

2) BPG-based experiments

The average distortion curve of the BPG coder reflects the dependence of visual quality on the control parameter Q. Those curves have been obtained from the experimental statistical data of the image sets shown in Fig.2.13.



Figure 2.13- gray-scale mage set for BPG coder

This image set contains images that are standard in image processing, remote sensing images, and selected typical texture images. It can be seen that the types of images in the set are diversified, and the complexity of the images varies. The experimental data of this image set has also become the original data of the two-step method design.

The implementation process of the two-step method on the BPG encoder is similar to that of the AGU. First of all, a group statistical data for the dependence of metric on Q is obtained from the image set, and then the average rate-distortion curve can be drawn as Fig 2.12. Then, the first step compression is conducted with the initial Q. The calculation basis is that the average value of the visual quality corresponding to the estimated value is sufficiently close to the desired value. This estimated value is calculated by equation (2.17).

$$Q_{init} = Q_{est} + \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{ave}}{M'} \quad (2.17)$$

where Q_{est} is the right margin of the interval of the average rate-distortion curve.

After the first step compression, the visual metric is calculated for the decompressed image. As a case, let us take the image Goldhill and metric PSNR-HVS-M as the example, which is 34,071dB. This value is quite close to the desired value, but the accuracy can still be improved by the second step. In the second step, Q is corrected by $PSNR-HVS-M_{init}$. It is calculated using Equation 2.18.

$$Q_{des} = Q_{init} + \frac{PSNR-HVS-M_{des} - PSNR-HVS-M_{init}}{M'} \quad (2.18)$$

where Q_{des} is the corrected Q value used at the second step of compression.

After the second step, the provided PSNR-HVS-M is 35,394dB and is closer to the desired value. Differently from the AGU coder, the PCC Q in BPG is only allowed to be set as an integer. Does this feature affect the two-step method for BPG? More experiments and analyses need to be carried out.

3) Analysis of compression accuracy

Three typical metric values, 40dB, 35dB, and 30dB, are chosen for visual quality metrics PSNR and PSNR-HVS-M to analyze the effect of the two-step method on BPG. The statistical results of the experiment are shown in Table 2.10.

Table 2.10 –Statistics of BPG coder for gray-scale test images

Quality metric	M_{des}	VAR_{fis}	VAR_{sec}	$MAX_{\Delta fin}$
PSNR	40	1,227	0,108	1,032
PSNR	35	3,754	0,175	1,325
PSNR	30	1,856	0,781	2,87
PSNR-HVS-M	40	6,243	0,343	1,05
PSNR-HVS-M	35	3,087	0,916	2,545
PSNR-HVS-M	30	1,463	0,274	1,072

It can be seen from the statistical data that although the control parameter of BPG supports only integer settings, the variance after compression in the second step (after parameter correction) is about an order of magnitude lower than that in the first step, and the error after two-step compression is acceptable for practice. The error for PSNR does not exceed 2,87dB, and the error for PSNR-HVS-M does not exceed 2,55dB. It can be seen that the two-step compression method is suitable for BPG compressors and can provide sufficient accuracy improvement. In order to understand the specific data details, we use the desired value of PSNR-HVS-M to be 35dB for further analysis.

The Q_{init} shown in Table 2.11 is the parameter used for the first step of compression. It comes from the average rate/distortion curve and equal to 35 for all images (recall the calculation in Section 2.3). PSNR-HVS-M $_{init}$ after the first step of compression is different for different images. The interval of its changing is [34,071 ~ 40,838], the variance is 3,087, and the maximum error is 5,838dB. It can be seen that although PSNR-HVS-M $_{init}$ is around 35dB, there is a large range of its variation. At the same time, this value is used for feedback and calculation of the Q correction for each particular image (using Equation (2.18)). Using Q_{des} as the parameter for the second step of compression, the PSNR-HVS-M value range that can be provided is narrowed. The range is [32,455 ~ 36,097], the variance is 0,916, and the maximum error is 2,545dB. Generally speaking, regardless of the range of change or variance, the maximum error is, in general, improved after the second step of compression. In particular, it can be seen that in many cases, the corrected parameter Q in the second step is the same as that in the first step. In this case, only one step is required to achieve the desired visual quality. The second step compression can be omitted to improve time efficiency.

Table 2.11 –Statistics and parameters of providing PSNR-HVS- $M_{des}=35$ dB for BPG coder in gray-scale image compression

Test image	Q_{init}	PSNR-HVS- M_{init}	Q_{desr}	PSNR-HVS- M_{pro}	CR
Aerial	35,000	35,683	35,000	35,683	5,997
Baboon	35,000	35,374	35,000	35,374	6,929
Barbara	35,000	35,665	35,000	35,665	16,755
Diego	35,000	35,447	35,000	35,447	7,827
Frisco	35,000	36,414	36,000	35,495	29,486
Goldhill	35,000	34,071	34,000	35,394	16,831
Lenna	35,000	35,871	35,000	35,871	31,028
Man	35,000	34,926	34,000	36,097	10,955
Texture 1	35,000	40,838	40,000	32,455	7,427
Texture 2	35,000	37,937	37,000	34,762	8,456
Texture 3	35,000	35,042	35,000	35,042	25,348
Texture 4	35,000	35,869	35,000	35,869	17,187
Variance		3,087		0,916	

There are two more interesting aspects that follow from the analysis of data in Table 2.11: 1) The largest error occurs in Texture1, not in the Frisco image. Recall that in Fig.2.13, the largest deviation from the average distortion curve was for the test image Frisco. The selected data here is one group of data in 6 groups, and the largest deviations in the other groups of data appear for the test images of Texture9, Diego, Man, and Frisco, respectively. This shows that the occurrence of the error is not related to the original curve, and a good improvement effect can be achieved through the second step of correction; 2) The obtained compression ratio for different images varies greatly. Under the same visual quality, CRs vary from 6 to 31. Therefore, the two-step method provides BPG-compression with the desired visual quality that guarantees pre-requisites for subsequent image processing; thus, the advantages of BPG can be maximized.

2.5 Conclusions

Iterative methods can provide satisfactory desired image quality for lossy compression, but at the cost of multiple compression/decompression runs. Time efficiency cannot meet application requirements on some occasions.

Compared with iterative methods, prediction methods can provide decompressed image quality predictions for lossy compression faster, which facilitates better parameter setting. However, the current prediction methods still need to be improved in accuracy. The newly proposed two-step compression method can achieve a balance between accuracy and time efficiency and only requires at most two compressions to achieve appropriate accuracy. Experiments show that this simple two-step compression method can work well on DCT-based AGU encoders and HEVC-based BPG encoders, providing better accuracy than prediction methods.

For BPG, even a single step can meet the accuracy requirements for some images. By controlling the visual quality, the encoder can be set with appropriate parameters with maximizing the BPG compression ratio.

CHAPTER 3

IMPROVEMENTS OF TWO-STEP COMPRESSION METHOD FOR GRAY-SCALE IMAGE

The primary two-step compression method[42, 43] often provides the desired visual quality for grayscale image compression for AGU and BPG encoders. In this chapter, this method is further studied. Aiming at the problem of unsatisfactory accuracy in the low expected value of PSNR on the AGU encoder, a conditional constraint is imposed on the CCP correction scheme. Furthermore, with this constraint, the use of the two-step compression method can be generalized to the DCT-based ADCTC encoder and the DWT-based SPIHT encoder. In addition to this constraint, this chapter also optimizes the curve correction model in the basic two-step compression method, which improves the robustness of the scheme for the SPIHT and AGU encoders. Finally, the visual quality accuracy is further improved through a hybrid lifting scheme combining constraints and optimized curve models.

3.1 Modification of correction scheme of CCP in the second step

The primary two-step compression method can correct the control parameters by averaging the distortion curve so as to calculate the specific parameters for an image to be compressed and finally provide the compression of the desired visual quality value. However, in special cases, the results obtained by this correction are unsatisfactory, the error cannot be reduced, and even CCP calculated at the second step occur to be negative which is out of sense. Let us call this situation "over-correction", and in this case, some constraints need to be applied to prevent over-correction and improve visual quality accuracy in the second step of compression. In this section, we modify the preliminary two-step method by judging the parameter correction bias and proposing a constraint. Through experiments in AGU, the constraint scheme works well, effectively avoiding overcorrected results. With this constraint scheme, we also generalize the two-step approach to the DCT-based ADCTC encoder and the DWT-based SPIHT coder.

3.1.1 Analysis of two-step approach for compression of texture images with desired quality

Through the two-step compression method, the process of providing the desired visual quality in lossy compression has simplified and become faster than the previous method[81, 139]. In some practically critical situations, its accuracy is acceptable [181]. However, through in-depth research, it has been shown that if one needs to provide the desired quality characterized by the Peak Signal-to-Noise Ratio (PSNR) of about 30dB... 35dB or smaller, the accuracy radically decreases and becomes unacceptable[42]. This happens more often if a compressed image has a complex structure. Meanwhile, complex structure (highly textural) images are ubiquitous in remote sensing imaging [183], industrial [184], and intelligent medical-assisted diagnostic applications [77, 88, 111]. Texture features are of great significance in image classification and automatic recognition [183].

To address the above problem in the AGU coder, we study the compression performance of complex texture images mainly employing PSNR as a quality metric, analyze the results, and give the limitations of the two-step method in practical applications. A correction scheme is proposed to improve the accuracy of quality providing for the two-step method after correction.

1) Basic peculiarities of the two-step approach of AGU coder

The term image complexity is widely used in image processing, although it has not been strictly specified yet. So, let us try to explain it verbally with several examples. It is a known fact in lossless compression[185] that an attained CR varies in certain limits depending upon an image lossless compression is applied to. There are images with “unpredicted” or “hardly predictable” structures that are compressed with CR close to unity. These are images with many “locally active areas” as edges, details, and/or textures. Similarly, there are noisy images for which efficient denoising is impossible [186]. Again, these are images with a high percentage of pixels that belong to locally active areas. Coming back to lossy compression, we can state that for complex structure images, either PSNR is smaller if the same CR is provided (or the same PCC is used) or CR is smaller if

the same PSNR is provided. To prove this, Fig. 3.1.a presents two examples of the rate-distortion curves – dependences of PSNR on QS for the coder AGU for which QS serves as CCP. These dependences (Fig. 3.1.b) have been obtained for two test remote sensing images - Frisco and Diego – that are presented in Fig. 2.14 and which are good examples of simple and complex structure images, respectively. As one can see, PSNR values for a given QS can differ by several dB. Considering this, the proposition in [42, 181] was to apply some “good” initial QS at the first stage. Because of this, the primary method [42, 181] is based on using the average rate-distortion curve for trend prediction and QS setting. In fact, the average distortion curve provides appropriate preconditions for calculating the initial (rough) value of any CCP in general and QS for AGU in the considered particular case.

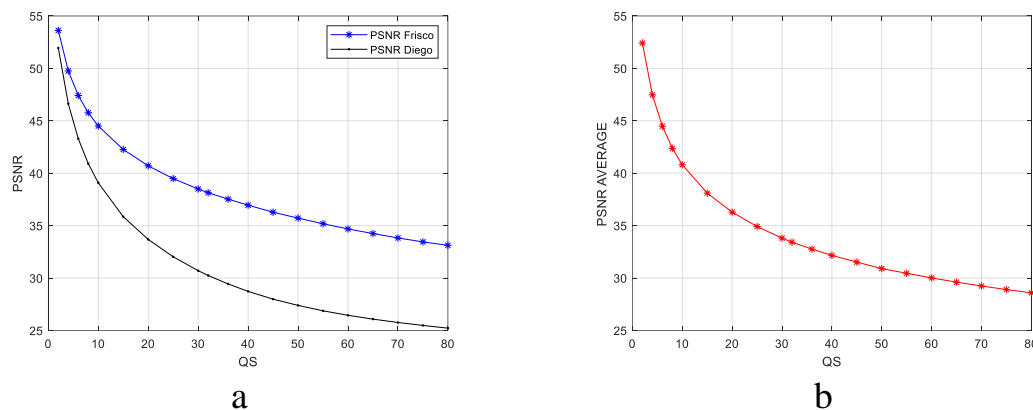


Figure 3.1 – Dependence of PSNR on QS for a) sample images and b) average curve for nine images compressed by the AGU coder

In order to understand this, some comparative experiments are necessary. Let us employ three visual quality metrics, namely PSNR, PSNR-HVS, and PSNR-HVS-M. The latter two metrics consider the human visual system (HVS), and they have been shown in the paper[42] to work well with the two-step method.

The comparison test results for the Airfield image are shown in Fig. 3.2. The similarity of images in Fig.3.2.a and Fig.3.2.b are very high, and some slight differences can be found (the parts marked in yellow), while one can more easily find some differences between images in Fig. 3.2.a and Fig. 3.2.c (these parts are marked in red). In this way, analysis of these images and the corresponding metrics' values allows drawing the following conclusions: when the difference of PSNR values (Δ PSNR) is less than 1dB,

the image difference is not obvious, but when ΔPSNR is greater than or equal to 1.5dB, the difference in compressed image quality can be easily observed.

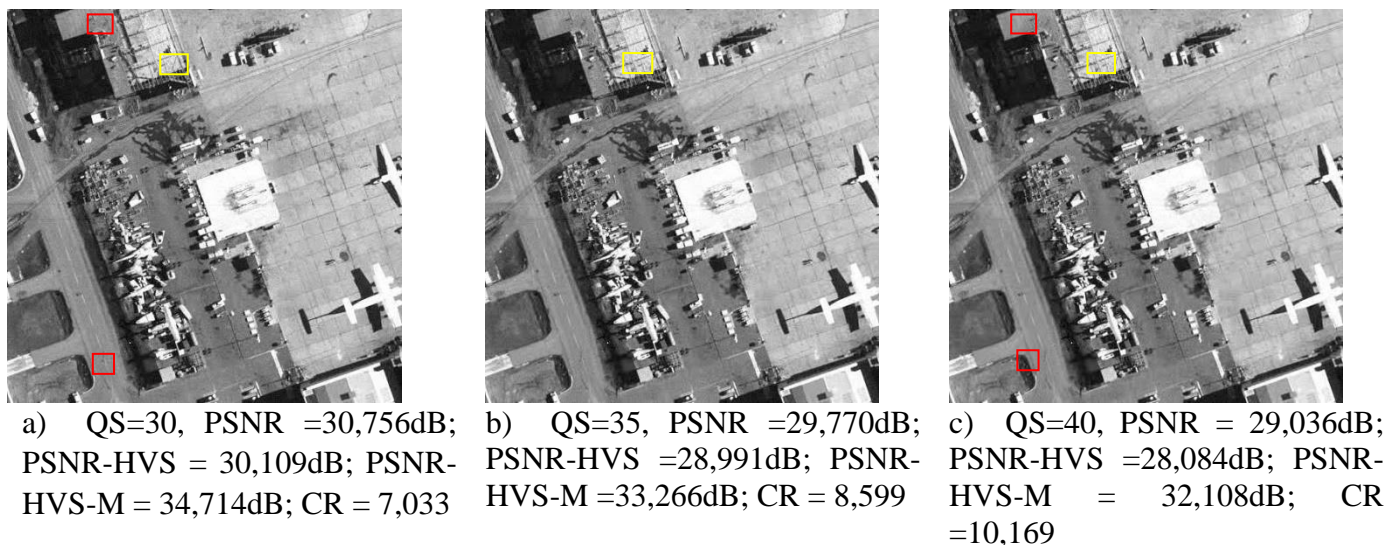


Figure 3.2 - Comparison of decompressed images Airfield for QS=30 (a), QS=35 (b), and QS=40 (c)

Particular images might have rate-distortion curves (see Fig. 3.1.a) that differ from the average curve (Fig. 3.1.b). This leads to the fact that for the images Frisco and Diego, the errors of providing $\text{PSNR}_{\text{des}}=35$ dB are 1,246 dB and 2,794 dB, respectively. This is less than after the first stage, but these errors are not appropriate. Such errors show that different accuracy of providing PSNR takes place, and this is reflected in the appearance of compressed images.

2) Analysis of quality providing for the two-step lossy compression approach

In [181], the results for the primary two-step compression method were analyzed for the set of test images (used in obtaining the average distortion curve). Three sets of data have been obtained for PSNR for three values of PSNR_{des} , namely, 34 dB, 37 dB, and 40 dB. It was found from the results that the proposed two-step procedure worked well enough if it was desired to provide PSNR larger than 37 dB, but the error of providing the desired PSNR occurred to be too large for PSNR_{des} smaller than 35.....37 dB[181].

To better understand how the two-step method works and what are the arising problems, let us consider the cases of PSNR_{des} equal to 35 and 30 dB in more detail. Table 3.1 presents data obtained for $\text{PSNR}_{\text{des}}=35$ dB. The lower line presents variance values of $\text{PSNR}_{\text{init}}$ and $\text{PSNR}_{\text{prov}}$. It is seen that, due to the second step, variance is significantly

reduced compared to one-step compression. It works well for simple and middle complexity images (consider data for such test images as Lenna, Barbara, Goldhill). However, there are four test images for which the error of providing $PSNR_{des}$ is greater than 1 dB. This means that people can clearly see the difference when looking at the corresponding images (recall the comparison in Fig.3.2).

Table 3.1 –Statistics and parameters of providing $PSNR_{des}=35dB$ in AGU coder for gray-scale images

Test image	QS_{init}	$PSNR_{init}$	ΔQS	QS_{rec}	$PSNR_{prov}$
Goldhill	24,60	34,503	-2,2476	22,3587	35,0117
Baboon	24,60	32,4512	-11,5268	13,0796	37,1222
Barbara	24,60	35,9636	4,3578	28,9642	35,0604
Lenna	24,60	36,3968	6,3169	30,9233	35,4641
Aerial	24,60	33,6942	-5,9054	18,701	35,5482
Airfield	24,60	32,2547	-12,415	12,191	38,6067
Frisco	24,60	39,574	20,6781	45,2845	36,2463
Diego	24,60	32,1536	-12,868	11,7384	37,794
Mrt _{prepared}	24,60	38,0092	13,609	38,2153	35,4429
Variance		7,1179			1,6761

It should be noted that the QS_{init} is obtained differently from the one in the paper [42], where Equation 3.1 is adopted to get a better estimate value instead of the rough value that occurs from the average distortion curve.

$$QS_{init} = QS_{est} + \frac{PSNR_{det} - PSNR_{ave}}{M'} \quad (3.1)$$

Here QS_{est} is the value estimated from the average distortion data, $PSNR_{ave}$ is the average PSNR value corresponding to the estimated QS_{est} .

Consider now the data for $PSNR_{des}=30$ dB. Recall here that for the cases of PSNR about 30 dB and less, the distortions introduced by lossy compression can be clearly observed, and the image quality is quite poor. The obtained data are presented in Table 3.2. Analysis shows the following. First, variance values for both $PSNR_{init}$ and $PSNR_{prov}$ have sufficiently increased compared to data in Table 3.1 and become practically equal. Second, the number of images with an error larger than 1 dB has risen to 5 out of 9. This shows that the two-step method cannot provide appropriate accuracy for producing the desired PSNR when $PSNR_{des}$ is about 30 dB. Again, the largest errors are observed for highly textural images such as Baboon, Airfield, and Diego.

Table 3.2 – Statistics and parameters of providing $PSNR_{des}=30dB$

Test image	QS_{init}	$PSNR_{init}$	ΔQS	QS_{rec}	$PSNR_{prov}$
Goldhill	60,15	30,34	-4,24	55,90	30,639
Baboon	60,15	26,72	-40,2	19,87	33,973
Barbara	60,15	30,88	10,87	71,01	29,947
Lenna	60,15	32,62	32,35	92,50	30,717
Aerial	60,15	28,34	-20,4	39,68	30,687
Airfield	60,15	27,05	-36,3	23,80	32,521
Frisco	60,15	34,68	57,62	117,7	31,155
Diego	60,15	26,44	-43,83	16,31	35,210
Mrt _{prepared}	60,15	32,89	35,66	95,81	24,510
Variance		9,1183			9,0517

3) Complex texture image test experiment

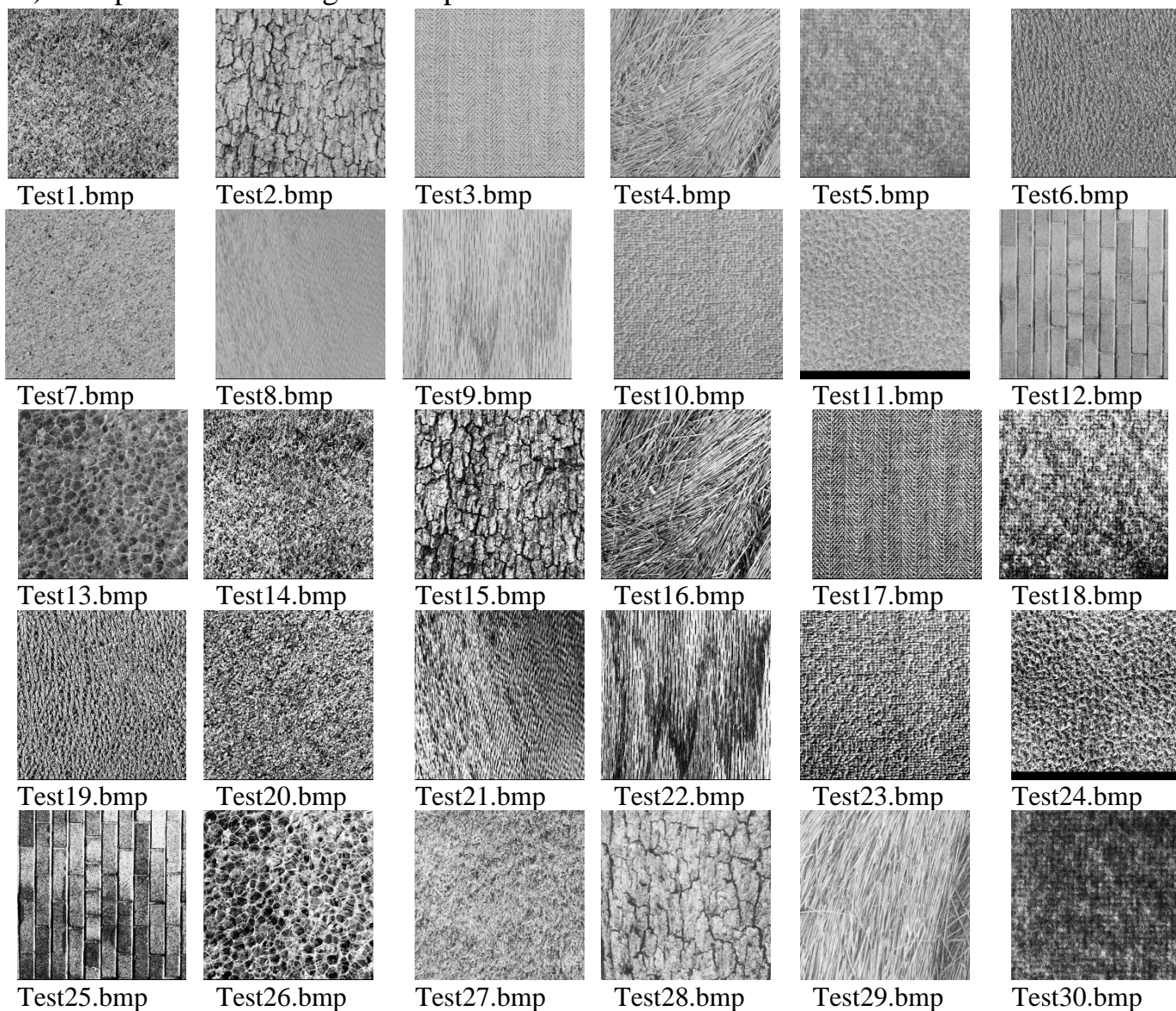


Figure 3.3 – Texture test image set for the AGU coder

By comparing the data, it can be seen that large errors of the designed two-step method often appear in images with complex structure (e.g., Baboon, Airfield, Diego). To analyze these effects more in detail, 30 texture images [187] have been chosen for testing. The image set and their titles are shown in Fig. 3.3. There are two reasons for choosing these images. First, these images are not in the image library from which the average rate-distortion curve was obtained earlier. The testing of these images can more objectively reflect the real application of the two-step method. Secondly, these images contain texture structures with different degrees of complexity. The test results can, to some extent, reflect the impact of the complexity of the texture structure of the image on the accuracy that can be achieved by the two-step compression.

Test statistics are shown in Table 3.3 for the images shown in Fig 3.3. Alongside the data for the PSNR metric is paid the main attention in this section, visual quality metrics are considered.

Table 3.3 –Statistics for texture images

Visual quality Metric	M_{des}	VAR_{fis}	VAR_{sec}	MAXΔ_{final}
PSNR	40	0,2852	0,280	1,5763
PSNR	35	1,5048	32,3906	29,3603
PSNR	30	3,692	N/A	
PSNR-HVS	40	0,0314	0,0011	0,0892
PSNR-HVS	35	0,1862	0,0097	0,3517
PSNR-HVS	30	0,6208	0,1528	1,3796
PSNR-HVS-M	40	5,6079	0,194	1,3653
PSNR-HVS-M	35	4,6928	0,4413	1,5423
PSNR-HVS-M	30	1,8633	0,2485	1,0542

From the statistical results, it can be seen that for the metrics PSNR-HVS and PSNR-HVS-M, the two-step method works well enough, and the variance after the two steps is considerable, by approximately one order, smaller than the variance of the first step. Thus, the error is also controlled well, which is better reflected for PSNR-HVS.

For the PSNR metric, more problems are revealed. When the desired value is 40 dB, the two-step method still works well. When the desired value is 35 dB and, especially 30 dB, abnormal situations occur. The experimental data when $PSNR_{des} = 30$ dB are shown in Table 3.4.

Table 3.4 – Statistics and parameters of providing PSNR_{des}=30 dB for texture images

Test image	QS _{init}	PSNR _{init}	Δ QS	QS _{rec}	PSNR _{prov}
Test1	60,148	25,705	-52,881	7,267	41,734
Test2	60,148	26,705	-40,571	19,577	33,955
Test3	60,148	26,48	-43,339	16,809	34,965
Test4	60,148	26,737	-40,175	19,973	33,948
Test5	60,148	26,454	-43,659	16,489	35,042
Test6	60,148	25,443	-56,107	4,041	46,525
Test7	60,148	26,836	-38,956	21,192	33,384
Test8	60,148	32,835	34,905	95,053	29,663
Test9	60,148	30,938	11,549	71,697	30,064
Test10	60,148	28,176	-22,458	37,691	30,757
Test11	60,148	28,088	-23,541	36,607	30,581
Test12	60,148	27,53	-30,411	29,737	31,182
Test13	60,148	27,297	-33,280	26,868	32,071
Test14	60,148	25,099	-60,342	-0,194	
Test15	60,148	25,211	-58,963	1,185	56,455
Test16	60,148	25,285	-58,052	2,096	51,6
Test17	60,148	24,68	-65,501	-5,353	
Test18	60,148	24,243	-70,882	-10,733	
Test19	60,148	24,398	-68,973	-8,825	
Test20	60,148	24,458	-68,234	-8,086	
Test21	60,148	25,231	-58,717	1,431	54,525
Test22	60,148	26,108	-47,919	12,229	37,446
Test23	60,148	25,561	-54,654	5,494	44,061
Test24	60,148	24,709	-65,144	-4,996	
Test25	60,148	24,548	-67,126	-6,978	
Test26	60,148	25,464	-55,848	4,300	46,07
Test27	60,148	26,863	-38,623	21,525	33,677
Test28	60,148	27,616	-29,352	30,796	31,571
Test29	60,148	29,249	-9,246	50,902	30,223
Test30	60,148	25,938	-50,012	10,136	38,92

As one can see, PSNR_{init} varies in wide limits starting from approximately 24dB and completing by approximately 33dB. This means that even if an image subject to lossy compression is fully textural (recall that the test images are taken from the database of texture images [187]), there is a probability that some of them can be compressed well, and Δ QS can be positive. However, for most textural images, the situation is the opposite. For the most complex structure (problematic) images, the recommended value of the

recommended QS occurs to be negative (marked by red), and this causes the procedure to stop running because QS should be positive by definition. Another observation to note is that for data (marked by blue), the operation of the two-step procedure is also improper since the errors of providing $PSNR_{des}$ are generally very large.

Table 3.5–Statistics and parameters of providing $PSNR_{des}=30$ dB with constraint

Test image	QS_{init}	$PSNR_{init}$	ΔQS	QS_{rec}	$PSNR_{prov}$
Test1	60,148	25,705	-52,881	30,074	30,486
Test2	60,148	26,705	-40,571	30,074	30,926
Test3	60,148	26,48	-43,339	30,074	30,717
Test4	60,148	26,737	-40,175	30,074	31,065
Test5	60,148	26,454	-43,659	30,074	30,594
Test6	60,148	25,443	-56,107	30,074	30,146
Test7	60,148	26,836	-38,956	30,074	30,931
Test8	60,148	32,835	34,905	95,053	29,663
Test9	60,148	30,938	11,549	71,697	30,064
Test10	60,148	28,176	-22,458	37,691	30,757
Test11	60,148	28,088	-23,541	36,607	30,581
Test12	60,148	27,53	-30,411	29,737	31,182
Test13	60,148	27,297	-33,280	30,074	31,320
Test14	60,148	25,099	-60,342	30,074	30,153
Test15	60,148	25,211	-58,963	30,074	30,136
Test16	60,148	25,285	-58,052	30,074	30,228
Test17	60,148	24,68	-65,501	30,074	29,876
Test18	60,148	24,243	-70,882	30,074	29,685
Test19	60,148	24,398	-68,973	30,074	29,744
Test20	60,148	24,458	-68,234	30,074	29,779
Test21	60,148	25,231	-58,717	30,074	30,158
Test22	60,148	26,108	-47,919	30,074	30,560
Test23	60,148	25,561	-54,654	30,074	30,446
Test24	60,148	24,709	-65,144	30,074	29,956
Test25	60,148	24,548	-67,126	30,074	29,856
Test26	60,148	25,464	-55,848	30,074	30,387
Test27	60,148	26,863	-38,623	30,074	31,374
Test28	60,148	27,616	-29,352	30,7958	31,571
Test29	60,148	29,249	-9,246	50,9079	30,223
Test30	60,148	25,938	-50,012	30,074	30,476
Var	3,823		0,281		

This means that some actions have to be undertaken to make the two-step procedure operation better. It is possible to observe from data in Table 3.4 that when $|\Delta QS|$ is greater

than $0,5QS_{init}$, the correction starts losing sense. In our view, there are the following reasons behind this. Recall that the basis of the two-step method is that, within a certain area, the average rate-distortion curve and the rate-distortion curves of particular images are approximately parallel, and linear approximations can be used to describe the curves locally. It can be seen from the Fig 3.1 that when $QS < 30$, this effect gradually disappears. This is why the data marked in blue in Table 3.4 have large errors. If the error of the M_{init} at the first step continues to increase, the corrected QS will be negative. Besides, the correction is calculated based on the corresponding derivative M' at M_{init} . When the error of the M_{init} value is large, it is not consistent with the actual situation to carry out correction using M' corresponding to M_{init} .

This shows that the primary two-step method has a certain range of applications. If it exceeds this range, its use needs to be restricted or the method has to be modified. We propose to analyze ΔQS . If $|\Delta QS| > 0,5QS_{init}$, then set

$$\Delta QS = 0,5QS_{init}. \quad (3.2)$$

The data for the modified algorithm that employs (2.14), (3.1) and (3.2) are shown in Table 3.5.

The experimental data show that when the application of the standard two-step method is restricted, the modified method of obtaining QS has achieved good results. The variance has been reduced by one order, and the maximum error does not exceed 1,58 dB.

3.1.2 A two-step procedure for image lossy compression by ADCTC with a desired quality

ADCTC is another DCT-based lossy compression coder, and we study how to provide desired visual quality in this coder for two reasons: first, it usually outperforms JPEG, JPEG2000, and many other compression techniques in rate-distortion sense by providing a sufficiently larger CR (compared to other coders) for many images in a wide range of quality variation [81, 119]. Meanwhile, ADCTC performs slower than many other compression techniques since it uses an optimized partition scheme. Then, we come to the urgent necessity to ensure that this compression technique is able to provide a

desired quality quickly enough to minimize power consumption at the software or hardware level to make the compression process environmentally friendly and green.

1) Two-step compression method for ADCTC

According to our previous work dealing with the two-step method, the first step is to obtain the average rate-distortion curve and its derivative values (or their approximations) for any QS.

The resulting distortion curves are shown in Fig.3.4.a and Fig. 3.4.b as a test, which is obtained from the image library. Among them, there are four standard images (Lenna, Barbara, Baboon, and Goldhill), four remote sensing images, and one medical image, which represent a variety of images that can be subject to lossy compression. In particular, there are highly textural images Baboon and Diego, whilst the test images Frisco and $MRT_{prepared}$ are quite simple. These images are shown in Fig.2.14, Fig.2.3.a, and Fig.3.2, respectively.

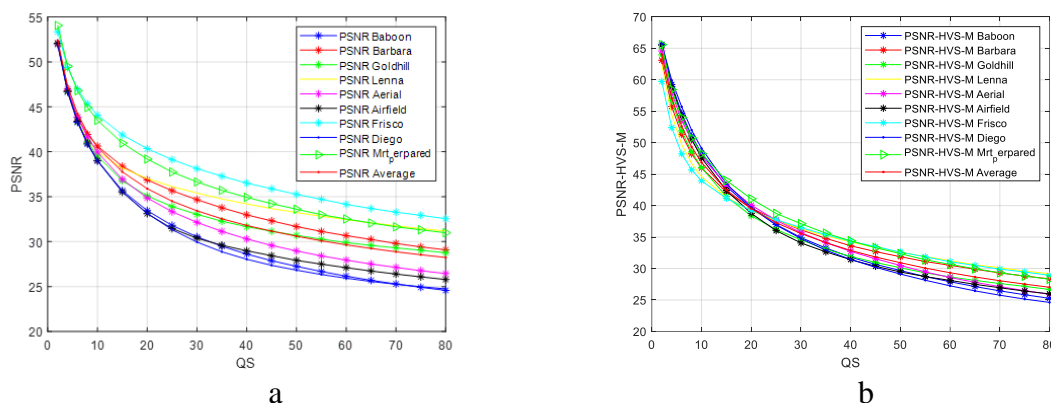


Figure 3.4 - Particular and average rate-distortion curves for ADCTC, a) PSNR, b) PSNR-HVS-M

In the first step of compression, we propose to set the CCP according to the measured average rate/distortion curve, which is the quantization step (QS) in ADCTC coding. Then the equation (3.1) is utilized to calculate the initial QS. After the initial value of QS is obtained, the image is compressed at the first step and then decompressed to obtain the visual quality value corresponding to a considered image. In the second step of the compression procedure, a QS corrected value for this image needs to be calculated with the equation (2.14) using the metric value obtained in the first step to improve accuracy.

The corrected (desired) QS value is used at the second step of compression. The whole processing is similar to the AGU coder.

2) Modification and analysis of the results

The first nine test images (the images in the library used to obtain the average distortion curve) are employed to test the applicability of the two-step method to ADCTC. In addition, to test its practical versatility, some texture images outside the image library will be used for verification [183].

Let us first examine the performance of the two-step method on these nine test images. Taking the visual quality value 40 dB of metric PSNR-HVS-M as an example, the statistical data are shown in Table 3.6.

Table 3.6 - Statistics and parameters of providing PSNR-HVS-M_{des}=40 dB for ADCTC coder

Test image	QS _{init}	PSNR-HVS-M _{init}	ΔQS	QS _{rec}	PSNR-HVS-M _{prov}
Goldhill	19,16	38,802	-2,02	17,13	40,001
Baboon	19,16	40,080	0,135	19,29	39,993
Barbara	19,16	40,000	0,000	19,15	40,000
Lenna	19,16	39,490	-0,86	18,29	39,895
Aerial	19,16	40,547	0,923	20,08	40,026
Airfield	19,16	39,272	-1,23	17,93	40,116
Frisco	19,16	39,516	-0,82	18,34	39,804
Diego	19,16	40,290	0,490	19,65	39,919
Mrt _{prepared}	19,16	41,516	2,56	21,71	40,187
Variance		0,6401			0,0131

The analysis of data in Table 3.6 shows that all values of the visual quality metric after the two-step compression have improved compared to the corresponding values of visual quality obtained after the first step of compression. We have calculated variance values for the first step and for two steps. The variance value for two steps also proves that the two-step compression is considerably more accurate.

It has been shown in [44] for AGU coder that the two-step method performs well for providing desired values of the metric PSNR-HVS-M, but it runs into problems if the desired PSNR is less than 35 dB. The experiment for such PSNR values has also been

conducted for ADCTC. Similarly to [44], we used the following modification. If the absolute value of ΔQS is larger than $QS_{init}/2$ and ΔQS is negative, then set

$$QS_{rec} = QS_{init}/2, \quad (3.3)$$

and continue the two-step compression. Due to this modification, the accuracy of providing the desired quality is guaranteed. The experimental data proving this are shown in Table 3.7.

Table 3.7 Statistics and parameters of providing $PSNR_{des}=30$ dB for ADCTC coder

Test image	QS_{init}	$PSNR_{init}$	ΔQS	QS_{rec}	$PSNR_{prov}$
Goldhill	56,05	30,17	1,988	58,033	30,038
Baboon	56,05	26,56	-38,21	17,835	31,018
Barbara	56,05	31,04	11,65	67,691	29,981
Lenna	56,05	32,73	30,33	86,372	30,883
Aerial	56,05	28,29	-18,91	37,134	30,764
Airfield	56,05	27,41	-28,78	27,259	30,775
Frisco	56,05	34,59	50,99	107,03	31,099
Diego	56,05	26,25	-41,62	14,418	30,449
Mrt_prepared	56,05	32,90	32,218	88,263	30,423
Variance		9,1895			0,1643

The reasons for improving the method can be explained by the following analysis. For the Baboon image in Table 3.7, the ΔQS calculated after the first compression/decompression step is equal to -38.21, which is greater than half of the initial value and is negative, and result in the highly erroneous $PSNR_{prov}$ (equals to 34,314dB). The same situation takes place for another highly textural image Diego (-41,62), for which the $PSNR_{prov}$ equals to 35,887dB. Let us come back to the average rate/distortion curve of PSNR (Fig. 3.4.a). It can be found that the curves of Baboon and Diego are the farthest (bottom) from the average rate/distortion curve, and the farthest above is the curve for the Frisco image, which also has highly erroneous $PSNR_{prov}=38,495$ dB. Due to this, ΔQS values for these three curves are the largest, where the values for Baboon and Diego are negative (marked by red), and the values of Frisco are positive (marked by green).

In Fig. 3.4.a, it can be seen that when QS is large, all the distortion curves change slowly, and the derivatives do not change much, while when QS is small, the derivatives of the distortion curve change greatly. Therefore, when ΔQS is negative and its absolute

value is greater than $QS_{init} / 2$ (negative direction), it is no longer accurate to predict the QS by the derivative corresponding to the QS_{init} , and it is necessary to adopt improved solutions, such as for the highly textural test images as Baboon and Diego. If ΔQS is greater than $QS_{init}/2$ (positive direction), it will not affect the accuracy, such as for image Frisco. However, after modification of CCP through the equation (3.3), we finally provide a better $PSNR_{prov}$ (see Table 3.7), thereby reducing the error. Six sets of statistical data for the metrics PSNR and PSNR-HVS-M are presented in Table 3.8.

Table 3.8 Statistics for nine test images for ADCTC coder

Visual quality metric	M_{des}	VAR_{fis}	VAR_{sec}	MAX_{Δfin}
PSNR	40	4,064	0,268	1,231
PSNR	35	7,776	0,788	2,368
PSNR	30	9,189	0,164	1,099
PSNR-HVS-M	40	0,640	0,013	0,187
PSNR-HVS-M	35	1,168	0,012	0,336
PSNR-HVS-M	30	2,217	0,058	0,843

From the analysis of these results, it can be seen that for the metric PSNR-HVS-M, the two-step method works well enough, and the variance after the two steps is considerable, by approximately one order, smaller than the variance of the first step. Meanwhile, the error is also well controlled –its variance does not exceed $0,85 \text{ dB}^2$.

By adopting the modified two-step method for this set of data with $PSNR_{des}=30 \text{ dB}$, the variance after two steps is also about one order smaller than the variance after the first step of compression. The error is also controlled well, and it does not exceed $2,4 \text{ dB}$.

Regarding the universality of this method, images outside the image library need to be used for verification. For this purpose, twenty highly textured images with different types and properties of textures have been selected as a verification test image set [187].

Through these 20 images, the applicability of the two-step method to high-texture images can be tested, and the error range of the two-step method in practical applications can be estimated. In the experiment, three representative visual quality values of 30, 35, and 40 dB were used as the desired quality, and six sets of data of PSNR and PSNR-HVS-M visual quality metrics were obtained. The results are shown in Table 3.9.

In this set of statistical data, for two sets of data, we used the modified two-step procedure. These are the cases for $PSNR_{des}$ equal to 30 dB and 35 dB. In both sets of data, there were cases when ΔQS was negative, and its absolute value was greater than $QS_{init} / 2$.

Table 3.9 Statistics for texture test images for ADCTC codeer

Visual quality metric	M_{des}	VAR_{fis}	VAR_{sec}	$MAX_{\Delta fin}$
PSNR	40	0,1677	0,1653	1,314
PSNR	35	1,8509	0,808	2,729
PSNR	30	5,092	0,1121	1,424
PSNR-HVS-M	40	7,3716	0,2819	1,502
PSNR-HVS-M	35	6,3311	0,4281	1,370
PSNR-HVS-M	30	2,9301	0,738	1,463

It can be concluded from the analysis of statistical data in Table 3.9 that, for ADCTC, the two-step method is universally applicable. For three considered values of PSNR-HVS-M, the procedure works well. The variance of residual errors after two-step compression is reduced by about an order of magnitude. For the three considered values of PSNR, the effectiveness of the two-step method is proved as well. It is worth noting that when the desired PSNR is equal to 40 dB, the variance after the two-step compression practically does not change compared to the variance in the first step. If the desired PSNR is equal to 35 dB or 30 dB, considerable improvement is observed, especially for $PSNR_{des}=30$ dB.

To facilitate detailed analysis, data for this group of test images are given in Table 3.10. As it can be seen from the analysis of data in Table 3.10, accuracy has been sufficiently improved – the variance of errors after the two-step compression is about 50 times smaller than after the first step. For almost all test images (except the test image #3 and the test image #13), $PSNR_{init}$ is smaller than $PSNR_{des}$. Most ΔQS are negative, and their absolute values are larger than $QS_{init}/2$ (there are even cases when absolute values of ΔQS are larger than QS_{init} and, then, QS_{rec} (according to (2.14)) occurs to be negative, which is impractical. In such cases, expression (3.2) is applied that ensures $PSNR_{prov}$ very close to the desired value of PSNR (30 dB). The maximal error is equal to 1.424 dB, and this is appropriate in practice.

Table 3.10 Statistics and parameters of providing PSNR_{des}=30 dB for texture images for ADCTC coder

Test Image	QS _{init}	PSNR _{init}	Δ QS	QS _{rec}	PSNR _{prov}
Test1	56,045	26,328	-40,800	28,023	30,455
Test2	56,045	25,144	-53,951	28,023	30,0986
Test3	56,045	30,863	9,590	65,635	30,092
Test4	56,045	27,247	-30,585	28,023	30,624
Test5	56,045	24,959	-56,003	28,023	30,2574
Test6	56,045	24,189	-64,557	28,023	30,2024
Test7	56,045	24,928	-56,347	28,023	29,9809
Tes8	56,045	24,539	-60,671	28,023	30,3694
Test9	56,045	26,728	-36,347	28,023	31,424
Test10	56,045	25,562	-49,311	28,023	30,354
Test11	56,045	26,481	-39,101	28,023	30,794
Test12	56,045	26,158	-42,689	28,023	30,152
Test13	56,045	32,706	30,066	86,111	30,186
Test14	56,045	27,803	-24,411	31,634	30,756
Test15	56,045	25,049	-55,002	28,023	30,427
Test16	56,045	24,473	-61,401	28,023	30,119
Test17	56,045	24,329	-63,006	28,023	30,163
Test18	56,045	25,372	-51,412	28,023	30,502
Test19	56,045	25,425	-50,825	28,023	30,565
Test20	56,045	29,140	-9,559	46,486	30,187
Variance		5,092			0,1121

3.1.3 Two-step providing of desired quality in lossy image compression by SPIHT

In the previous research work, the Two-step compression method was tested and verified on the encoders based on DCT, and the results have shown that this method can produce the compression efficiency and accuracy. However, the applicability of this method to encoders based on DWT such as JPEG2000 and SPIHT is uncertain and has not been studied yet. In this section, the related experiments are carried out for the coder SPIHT, which is based on the wavelet transform. Our goal is to analyze whether or not the two-step principle is applicable for SPIHT. Another goal is to understand are there any peculiarities of providing a desired peak signal-to-noise ratio (PSNR) and other quality metrics in this case. Finally, a general feasibility analysis of the two-step compression method is performed.

1) Two-step approach for SPIHT

In order to understand the dependence of image quality and compression ratio on control parameters, the Goldhill image was used as an example for testing. The experimental results are shown in Fig. 3.5.

For the SPIHT encoder, BPP is used as the control parameter. As can be seen in Fig. 3.5, as BPP increases, the compression ratio decreases, and the image quality improves. CR and BPP are strictly dependent where, for 8-bit representation of grayscale images, $CR \approx 8/BPP$ and, vice versa, $BPP \approx 8/CR$. Thus, knowing a desired CR, one calculates BPP, and SPIHT provides BPP slightly less than $8/CR$. Dependences of quality metrics on BPP are monotonously increasing (or non-decreasing) but they are not as smooth as for the DCT-based encoders due to specific properties of the SPIHT encoder.

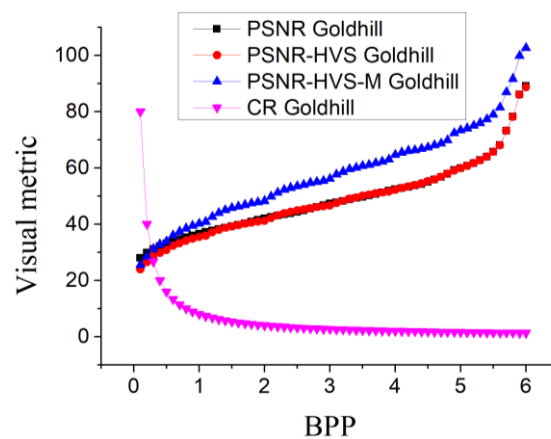


Figure 3.5 - Dependence of image quality and compression ratio on BPP for SPIHT of image Goldhill

To provide a desired image quality through the two-step method, the average rate/distortion curve needs to be obtained first. Here, nine images are used as the image library so as to obtain the average distortion curve of a small sample of images. The graphs for three visual quality metrics are shown in Fig. 3.6. Among the image library, there are four standard images (Lenna, Barbara, Baboon, and Goldhill), four remote sensing images, and one medical image, which represent a variety of images that can be subject to lossy compression. In particular, there are highly textural images Baboon and Diego whilst the test images Frisco and $MRT_{preparted}$ are quite simple.

In Fig. 3.6.a, the rate/distortion curve of the PSNR on BPP for each image is relatively smooth. For the same BPP value, metric values for different images differ greatly (up to 23 dB). The simple image Frisco's curve is at the top, and it has the highest

PSNR value under any BPP. The curves for Baboon and Diego images that are complex structure images are at the bottom, which means that the distortions are the largest under the same CR (BPP). However, the overall change trend is consistent. The average distortion curve obtained from these nine images are approximately parallel to other curves locally.

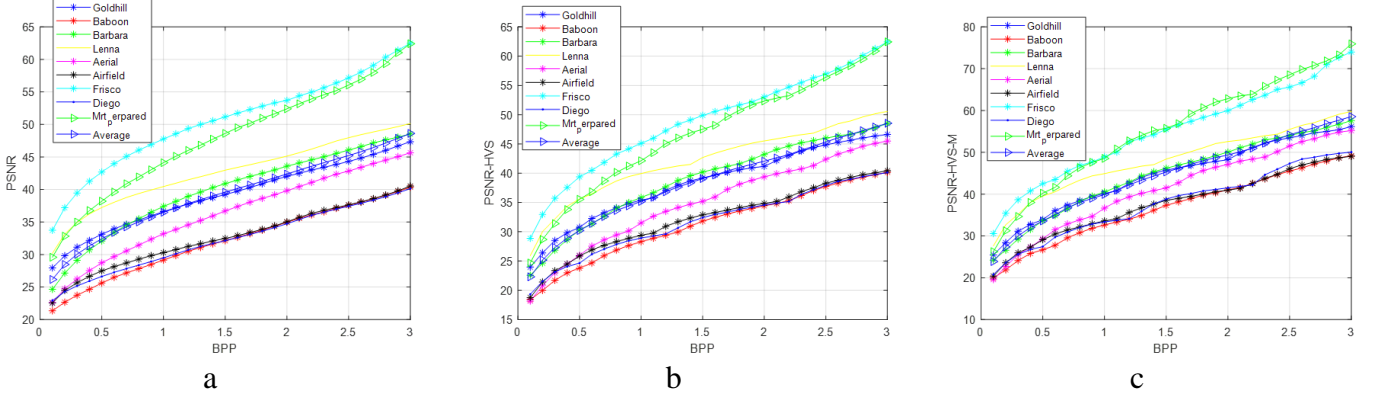


Figure 3.6 – Particular and average distortion curves for SPIHT for the metrics: a) PSNR, b) PSNR-HVS, c) PSNR-HVS-M

In Fig. 3.6.b and 3.6.c, one can observe that, for the same BPP, values of PSNR-HVS and PSNR-HVS-M are also sufficiently different. However, for PSNR-HVS-M, this difference is smaller than for PSNR, Complex structure images (such as Baboon) have the worst visual quality, and correspondingly, simple structure images (such as Frisco) have the best visual quality. The average distortion curves are again locally approximately parallel to dependences for particular images.

2) SPIHT-based experiment implementation

From the previous analysis, it can be concluded that, in SPIHT, it is feasible to perform compression prediction through the average distortion curve. In the first step of compression, we propose to set the CCP according to the measured average rate/distortion curve, the argument of which in SPIHT is BPP. Similar to the methods in other coders[18, 42, 44, 181], we can calculate the initial BPP value by the following formula:

$$BPP_{init} = BPP_{est} + \frac{PSNR_{des} - PSNR_{ave}}{M'} \quad (3.4)$$

After the initial value of BPP is obtained, the image is compressed at the first step and then decompressed to obtain the visual quality value corresponding to a considered image.

In the second step of the compression procedure, a BPP corrected value for this image should be calculated using the $PSNR_{init}$ value obtained at the first step to improve accuracy. We propose to do this as follows:

$$BPP_{des} = BPP_{init} + \frac{PSNR_{des} - PSNR_{init}}{M'} \quad (3.5)$$

The corrected (desired) BPP value is used in the second step of compression. After the second step of compression, the provided value $PSNR_{prov}$ is closer to the desired value $PSNR_{des}$ than the initial value $PSNR_{init}$. However, can the two-step method also provide precision control for the compression of other images? Further experiments and statistical data are necessary.

3) Analysis of compression accuracy

After the average rate/distortion curves based on the three visual evaluation metrics are obtained in advance, any image can be compressed by the two-step compression method. Taking the desired value of PNSR-HVS-M as 40dB as an example, nine images were tested, as shown in Table 3.11. In this Table, we give the following data: BPP_{init} calculated according to (3.4), $PSNR_{init}$ calculated after compression with BPP_{init} and decompression, ΔBPP calculated as $\frac{PSNR_{des} - PSNR_{init}}{M'}$ in (3.5), BPP_{des} calculated according to (3.5).

Table3.11- Two-step statistics and paramters of providing $PSNR_{des}=40dB$ for SPIHT

Test image	BPP_{init}	$PSNR_{HVS-M_{init}}$	ΔBPP	BPP_{des}	$PSNR_{HVS-M_{prov}}$
Goldhill	1,007	40,215	-0,021	0,986	40,031
Baboon	1,007	32,688	0,721	1,729	39,186
Barbara	1,007	40,616	-0,061	0,947	39,912
Lenna	1,007	44,934	-0,487	0,521	39,762
Aerial	1,007	36,721	0,323	1,331	40,455
Airfield	1,007	33,631	0,628	1,636	39,168
Frisco	1,007	48,944	-0,882	0,504	42,565
Diego	1,007	33,388	0,652	1,660	39,962
Mrt _{prepared}	1,007	48,632	-0,851	0,504	40,316
Var		40,971		1,013	

Before analyzing these data in detail, let us recall the following. There are several reasons why the proposed two-step procedure can lead to residual errors in providing a

desired value of an employed metric. First of all, linear interpolation is used in (3.5), and it is valid only in a certain (not too large) neighborhood of BPP_{init} and under the condition that an approximated function behaves linearly (absolute values of the second and higher derivatives are close to zero). Clearly, this is not guaranteed in our case. Secondly, linear approximation (3.5) exploits M' obtained for the average curve as a derivative estimate for all particular dependences. Nevertheless, this is also the idealization.

One can see in Table 3.11 that $|\Delta BPP|$ obtained for Frisco and $MRT_{prepared}$ images after the first compression step is quite large - it exceeds half of the BPP_{init} value. The values obtained after BPP calculation using (3.5) are 0,125 and 0,156, respectively, and the PSNR-HVS-M values obtained after such parameter compression are 32,085dB and 29,186 dB, respectively. So, the accuracy of visual quality provided has not been improved after the second step of compression. This is due to the aforementioned factors. Really, when BPP is smaller, the curve changes steeper, and the derivative changes more. At this time, the BPP is predicted by the derivative of the initial value point, which causes $|\Delta BPP|$ to be too large. This means that the expected improvement of providing visual quality accuracy cannot be ensured by Equation (3.5) in some cases. Similar situations happened for the DCT-based encoders such as AGU and ADCTC [18, 44], controlled by the quantization step. For SPIHT, we propose to use similar correction formula for BPP_{des} - if the absolute value of ΔBPP is larger than $BPP_{init} / 2$ and ΔBPP is negative, then set

$$BPP_{rec} = BPP_{init} / 2, \quad (3.6)$$

then continue the two-step compression. Due to this modification, the improvement in accuracy of providing a desired quality is guaranteed. The experimental data proving this are given in Table 3.11. After using the modified formula, the obtained visual quality values are 42,565 dB and 40,316 dB (closer to the desired visual quality value). A comparison of variance values shows that, after the second step, variance has decreased by about 40 times, i.e. the considerable benefit is provided. Meanwhile, a comparison to the data for AGU[44] shows that the accuracy of providing a desired PSNR-HVS-M for SPIHT is worse (for AGU, the variance of providing PSNR-HVS-M after second step is equal to 0,0108).

Let us test accuracy for the range of metric values that is the most important in practice. Three typical thresholds of 30 dB, 35 dB, and 40 dB for our visual quality metrics, PSNR, PSNR-HVS, and PSNR-HVS-M, respectively, have been obtained for nine test images of experimental data, which are summarized in Table 3.12.

Table 3.12-Two-step method statistics for test images for SPIHT

Quality metric	M_{des}	VAR_{fis}	VAR_{sec}	$MAX_{\Delta fin}$
PSNR	40	50.715	4.213	5.950
PSNR	35	39.577	9.598	7.673
PSNR	30	29.432	10.82	7.168
PSNR-HVS	40	45.320	2.175	3.517
PSNR-HVS	35	37.937	4.603	4.263
PSNR-HVS	30	31.349	3.599	4.369
PSNR-HVS-M	40	40.971	1.013	2.565
PSNR-HVS-M	35	32.867	2.922	3.598
PSNR-HVS-M	30	26.890	4.028	3.314

It can be seen from the data in this Table that the variance after the first compression step is large (up to $50,715\text{dB}^2$), and the variance after the second (correction) step is an order of magnitude lower than VAR_{fis} (except $PSNR_{des} = 30$ dB). The higher the desired visual quality, the smaller the error. In addition, through comparison, it can be concluded that the two-step method works better with employing the HVS-based visual quality metrics, and the maximum error does not exceed 5 dB. In the evaluation based on PSNR, the variance is improved, but the maximum error is still large, and more effective improvements are needed to reduce the error.

3.2 Modification in curve model

The problem of "over-correction" can be avoided by adding constraints in the parameter correction process after the preliminary step compression so as to ensure that the parameter correction has a positive effect on the reduction of errors. However, this simple constraint has a limited effect on improving the accuracy, and the effect is related to the image itself. In this section, we propose a simple pre-classification of images, using an adapted average distortion rate/distortion curve according to the image complexity. Through optimization of this curve model, better visual quality accuracy can be provided in AGU and SPIHT encoders.

3.2.1 Adaptive two-step procedure of providing desired visual quality of compressed image

Through extensive experiments on this method, the results prove that this method is not only suitable for AGU coder, but also for other DCT-based encoders, such as ADCT coder[18]. The test results on SPIHT show that this method is also suitable for DWT-based encoders[45]. However, with the in-depth study of the two-step method, we have also found the shortcomings of the method. The two-step method can provide satisfactory HVS-based visual quality metric accuracy, but when the expected PSNR value is low (less than 35 dB), the visual quality and expected value of some images after compression/decompression have a large deviation. This situation exists in the experiments of all three codes used in the tests. In response to this problem, this paper analyzes and proposes an improved method to increase the accuracy provided by the two-step method when the desired value of PSNR is low.

1) PSNR dependence analysis

In [42, 44, 181], we have studied the dependence of PSNR on QS, and there is a common trend in general, but for each certain QS value, PSNR for different images varies greatly. In addition to QS, what are other factors that determine PSNR for a particular image and a given QS? Let us try to answer this question.

Some researchers have proposed compression performance prediction for different types of images, such as facial images [188], remote sensing images [189], etc. Of course, different types of images have different characteristics in texture. However, through comparative experiments with different image data sets, it has been found that the complexity of an image plays an important role in determining PSNR and other quality metrics. Fig.3.7.a and Fig.3.7.b show two average rate-distortion curves, respectively, are PSNR, and PSNR-HVS-M.

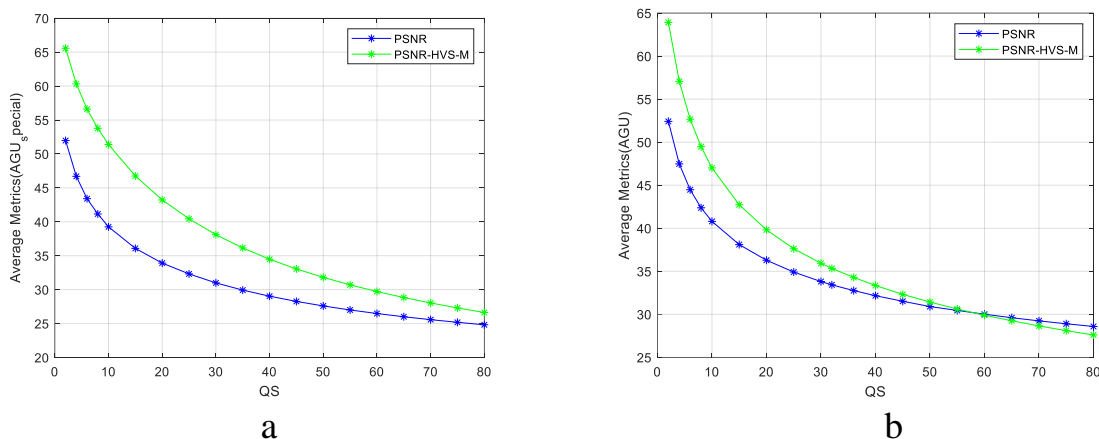


Figure 3.7 - Average metrics curve for AGU, a) texture images, b) nine general images

Figures 3.7 contain plots drawn from two sets of test images. The first set includes aforementioned nine test images (see [42] for more details), and the second set contains thirty texture images (see [44] for more details). It can be seen from the comparison of the plots in these two Figures that average PSNR-HVS-M values for the same QS differ by up to 2,5 dB. Meanwhile, the difference for PSNR is larger. For example, for QS=30, we have a mean PSNR of about 31 dB for the first set (texture images, Fig. 3.7.a) and about 34 dB for the second set (Fig. 3.7.b). Thus, the average rate-distortion curve of PSNR on QS can sufficiently depend on the choice of the image set. Recall here that M' in (2.14) is determined from the average dependence. Then, the difference in M' for a given image to be compressed and M' used in (2.14) can influence the accuracy of proving a desired quality.

Therefore, if the two-step procedure is applied to produce a desired PSNR-HVS-M, accuracy is high. At the same time, when it is desired to provide a given PSNR and this PSNR is low, accuracy can be unsatisfactory (see also data in [44]). It can be seen that even by increasing the number of test images to obtain some new average rate-distortion curve for PSNR, the improvement in accuracy cannot be reached because of the essential difference of dependences of PSNR on QS for different images.

An alternative way we propose to have more than one average rate-distortion curve suitable for images with different properties and to carry out some simple preliminary “classification” for each image to be compressed by referring it to two or more groups (classes). Such a classification should, in the simplest case, divide images into simple and

complex structures ones. Note that “classification” should be very fast so as not to lose the advantages of two-step compression.

2) Image complexity classification

Recall [42, 44, 181] that we are mostly interested in providing PSNR from about 30 dB to about 35 dB. Just for this interval, the accuracy of providing a desired quality according to PSNR can be inappropriate. Meanwhile, this is the main operation interval of lossy compression since $\text{PSNR} \geq 35$ dB corresponds to practically invisible distortions whilst $\text{PSNR} \leq 30$ dB relates to distortions that can be annoying. It follows from the analysis of the plots in Figures 3.7 that the aforementioned interval approximately corresponds to QS values from 15 to 50.

The simplest classification or discrimination procedures, if there are only two classes, are usually threshold ones, i.e., a considered parameter is compared to a threshold, and based on this comparison, a decision is undertaken. In such a case, obviously, one needs to find (or choose) a good input parameter and a proper threshold. For this purpose, let us, for the beginning, suppose that a simpler structure image, on the average, has a larger PSNR for a given QS (in other words, we suppose that an image has a simpler structure if it has a larger PSNR (and CR) for a given QS. Then, we need an input parameter that has a high correlation with PSNR. Strictly saying, it is not necessary to have a high Pearson correlation. It can be enough to have a high Spearman or Kendall correlation [190]. To carry out such an analysis, we have calculated the following parameters: CR, PSNR40, PSNR35, PSNR30, PSNRHVSM40, PSNRHVSM35, PSNRHVSM30. Here CR denotes the compression ratio for a lossless compression technique, namely, zip. An assumption is that since such a CR is determined by the information redundancy of an image, it can be connected with image quality when compression is lossy. In addition, we have analyzed six parameters that characterize compressed images for particular situations when the quantization step is set to 40, 35, and 30, respectively.

Considering the test images, nine general images and thirty-nine texture images have been adopted as the complete image set, and we have calculated all three correlation factors for all possible pairs of them. Fig.3.8.a and Fig.3.8.b, Fig.3.8.c are three correlation heat maps for Pearson, Spearman, and Kendall correlation. It can be seen that:

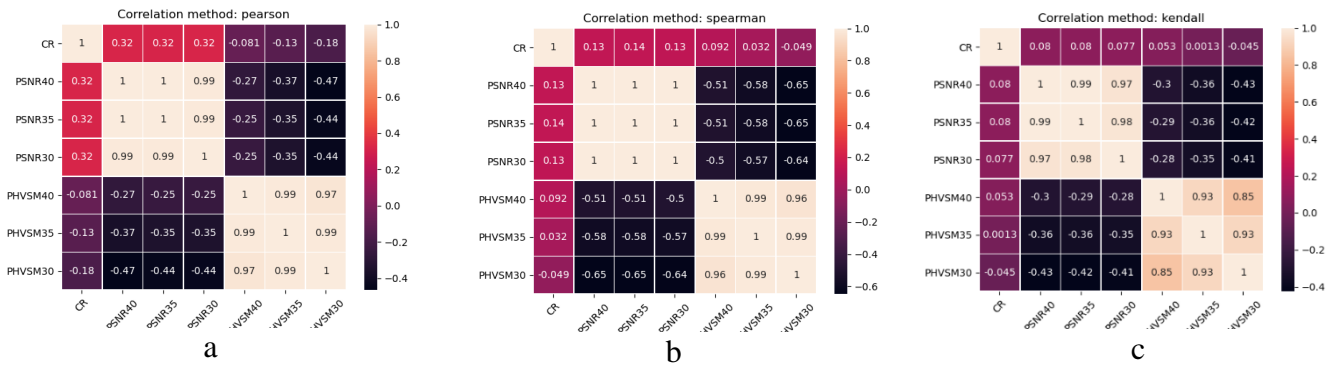


Figure 3.8 – Correlation coefficient heat map for seven parameters, a) Pearson; b) Spearman; c) Kendall

1) There is a certain correlation between CR and other parameters, but it is low. This means that CR provided by ZIP cannot be used for the characterization of image complexity (for the considered case).

2) The correlation coefficient between the same visual evaluation metrics for three values of QS is very high; it is above 0,97, which can be considered as highly correlated; this means in practice that if PSNR for QS=20 for one image is larger than for another image, it is very probable that the same holds for QS=30 and 40; in other words, the dependences PSNR on QS for different images intersect very rarely; this also means that if a given image belongs to a certain class for, e.g., QS=20, it belongs to the same class for other values of QS.

3) The same holds for PSNRHVSM40, PSNRHVSM35, PSNRHVSM30, although the correlation between PSNRHVSM40 and PSNRHVSM30 is slightly smaller than between PSNRHVSM40 and PSNRHVSM35.

4) There is also a certain correlation between PSNR and PSNR-HVS-M, but the correlation coefficients are all negative. It shows that the increase of one parameter may cause the decrease of another parameter.

5) The coefficient correlation between PSNR and CR is higher than the correlation between PSNR-HVS-M and CR.

Therefore, to improve the accuracy of the two-step method, a simple complexity classification of the image before pre-compression is a feasible way. Actually, PSNR itself has a certain correlation with CR, but this correlation is too low to rely on. On the contrary, PSNR for certain QS can be used to represent the image complexity to classify images.

For a specific QS value, a high PSNR indicates a low complexity of the image and vice versa.

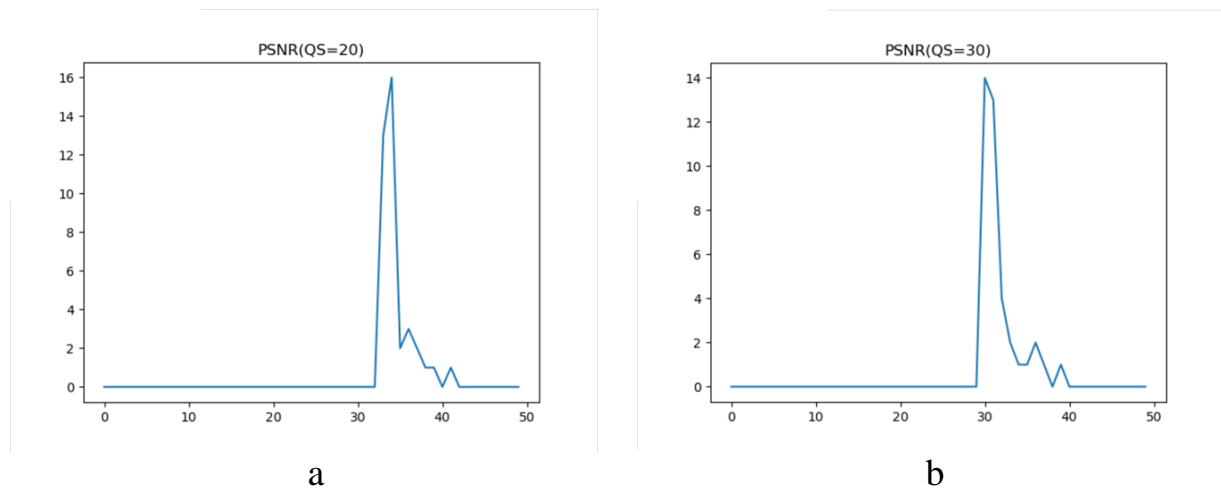


Figure 3.9 – PSNR histograms for AGU coder, a) QS=20, b) QS=30

In order to obtain the PSNR classification threshold and further analyze the previous results, experiments have been conducted on thirty-nine images. When QS is 20 and 30, respectively, the histograms of PSNR have been obtained and plotted in Fig.3.9.a and Fig.3.9.b. From the analysis of the plots in these two Figures, we can conclude that there is a common feature in the plots. A threshold is easy to find – it is approximately equal to 33 dB for data in Fig.3.9.a (for most textural images, PSNR is smaller than 33 dB). So, it can be considered that when QS=30, images with $\text{PSNR} > 33$ dB are treated as simple ones, while images with $\text{PSNR} \leq 33$ dB have higher complexity. According to this threshold, eight out of thirty-nine images are simple images, and thirty-one are complex images.

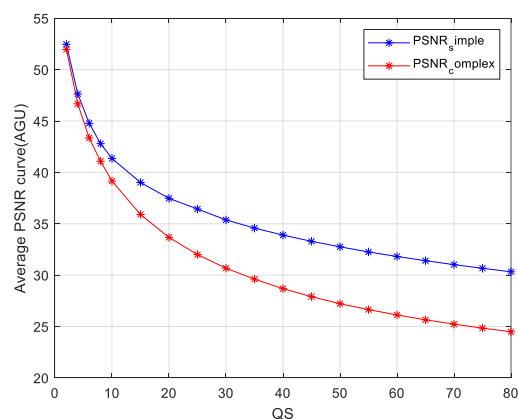


Figure 3.10 – PSNR average curves (simple & complex sets)

In Fig.3.10, two average rate-distortion curves are drawn for these two types of images. It can be seen from this Figure that the PSNR of simple images is higher than that

of complex images. When the desired quality PSNR is 30 dB, the initial QS value set according to the average rate-distortion curve in Fig. 3.7.b is about 60, and the corresponding derivative is also quite different. This leads to considerable errors in QS determination using (2.14) [42]. After the proposed subdivision of images into two groups, the error can be expected to reduce, and the accuracy can be better improved.

3) Proposed adaptive two-step method

According to the above analysis, to improve the accuracy of the two-step method, it is necessary to perform a PSNR30 calculation or prediction for an image to be compressed before determining the image complexity. In the compression prediction method for the AGU encoder [191], only a certain number of image blocks needs to be extracted for data processing to predict the PSNR30 value. Because no compression/decompression is involved, this operation does not considerably increase the time of overall two-step compression, and the proposed preliminary analysis can be performed quickly with satisfactory accuracy. The specific implementation method is as follows in Fig.3.11.

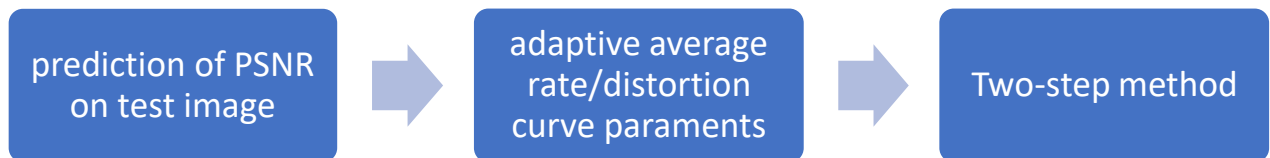


Figure 3.11 - Block diagram of adaptive two-step method for AGU

In this Section, thirty-nine images (nine general-purpose images and thirty texture images) have been divided into two groups according to the PSNR prediction results using PSNR30 and 33 dB as the threshold. Eight images belong to the simple set, and thirty-one belong to the complex set. Average rate-distortion curves for the two image sets have been obtained off-line, respectively (Fig. 3.10). The adaptive two-step method has been realized according to the flowchart in Fig.3.11.

1) The prediction program [191] is used to predict the PSNR30 of an image to be compressed and determine is it a simple image or a complex image. The experimental results prove that the classification effect can be achieved by selecting three hundred 8x8 image blocks, and the time efficiency can be guaranteed.

2) Based on the average rate-distortion curve corresponding to the determined image complexity, the initial QS value of the test image and the derivative of the average rate-distortion curve are adaptively set.

3) Calculate the initial QS value according to equation (3.1), and perform the first step of compression. After the PSNR feedback is obtained, the QS value is corrected according to equation (2.14), and the second step of compression is performed to provide a desired visual quality.

From the perspective of the overall implementation process of the adaptive two-step method, the improvement is mainly in two aspects: 1) First, when the average rate-distortion curves are obtained offline, the simple classification is first carried out according to the image complexity, the average rate-distortion curve of the simple image set and the average rate-distortion curve of complex image set are obtained. Because this process is done offline, it does not affect the time efficiency of the two-step method; 2) Secondly, the appropriate average rate-distortion curve is chosen.

4) Experimental results and analysis

The results of this method have to be verified. First, an experiment was performed on 39 images [18, 44] through the two average distortion curves in Fig.3.10. The experimental statistical results are shown in Table 3.12.

Table 3.12. PSNR statistical results of simple and complex images for image set

Image set	PSNR_{des}	VAR_{fis}	VAR_{sec}	MAX_{Δfinal}
Simple test image set	40	3,028	0,113	0,863
	35	3,1301	0,1413	1,062
	30	2,9887	0,2009	0,881
Complex test image set	40	0,0484	8,35E-04	0,133
	35	0,186	0,0017	0,159
	30	0,5904	0,0137	0,433

It can be seen from the statistical data in Table 3.12 that using the improved two-step method of predictive adaptation, the variance in the second step is one order of magnitude lower than that in the first step (almost two orders of magnitude in complex image sets). The QS in the second step is corrected more accurately, which can provide higher

accuracy. At the same time, the maximum deviation of the complex image set is obviously smaller than the maximum deviation of the simple image set. It can be concluded from these two indicators (variance and maximum deviation) that the adaptive two-step compression method can better ensure the compression quality accuracy of both simple and complex images. The maximal deviation of the provided PSNR with respect to the desired one is about 1 dB or less which is acceptable for practice.

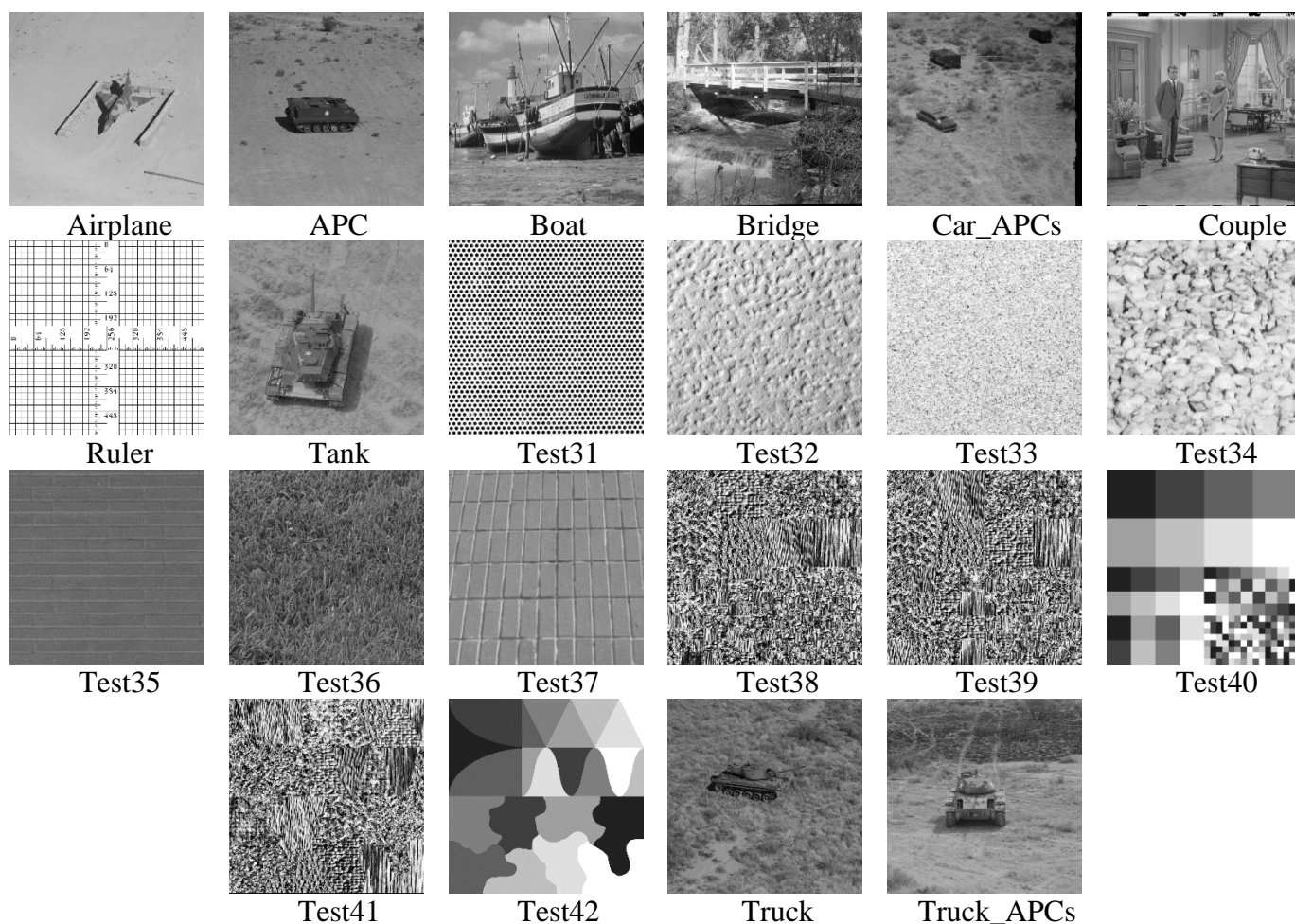


Figure 3.12 - test images set for adaptive two-step method for AGU coder

In order to better test the adaptive two-step method, the experiment first performed for thirty-nine image sets has also been done for the other 22 images shown in Fig.3.12. It contains general natural images, character images, drone aerial images, and texture images. Please note that “super-simple” images are also included here so that through experiments, we can learn whether this method is also applicable in extreme situations.

The experimental data statistics based on images in Fig.3.12 are shown in Table 3.13. According to the PSNR30 prediction result and the threshold of 33 dB in Fig.3.9, 22

images are divided into two groups, of which 11 images are simple images, and the other 11 images are complex images.

It can be seen from the statistical results that the adaptive two-step method normally works for this test image set. The variance of the second step is about an order of magnitude lower than the variance of the first step, and the data tends to converge. The effect in complex images is better than in simple images. The maximum error in complex images is 2,196 dB, while the maximum error in simple images is 4,063 dB (existing in extreme cases). Compared with the results in [42, 44, 181], the accuracy has been improved.

Table 3.13 - PSNR statistical results of simple and complex images for the AGU coder for 22 test images

Image set	PSNR _{des}	VAR _{fis}	VAR _{sec}	MAX Δ _{final}
Simple test image set	40	29,1022	1,4357	4,063
	35	18,076	1,7359	3,56
	30	9,1728	0,631	1,492
Complex test image set	40	0,0848	0,0006	0,053
	35	0,2344	0,0033	0,21
	30	1,3415	0,4693	2,196

Table 3.14 - Statistics and parameters of providing PSNR_{des}=35 dB for the AGU coder (simple images)

Test image	QS _{init}	PSNR _{init}	Δ QS	QS _{rec}	PSNR _{prov}
Airplane	32,37	36,545	9,751	42,120	35,557
APC	32,37	33,545	-9,184	23,185	34,908
Boat	32,37	33,389	-10,170	22,199	35,199
Car_APCs	32,37	33,572	-9,015	23,354	34,999
Couple	32,37	33,965	-6,534	25,835	35,202
Ruler	32,37	38,433	21,670	54,039	31,444
Truck	32,37	33,180	-11,493	20,876	35,355
test34	32,37	33,381	-10,222	22,147	35,014
test37	32,37	34,670	-2,082	30,287	34,825
test40	32,37	47,393	78,236	110,605	37,003
test42	32,37	38,301	20,838	53,207	35,384
Variance		18,076			1,7359

In order to understand the composition of the experimental data, the experimental data for $\text{PSNR}_{\text{des}}=35$ dB are presented in Table 3.14. It can be found that the data accuracy for nine images is satisfactory, and the data of two images have become the main factors for the large variance of the statistical data in the first compression study, namely Ruler and test40. For example, for the image test40, the PSNR value in the first compression step is as high as 47,393 dB. According to the image complexity classification rule in our study, this value indicates that the complexity of the image is extremely low and can be considered a super-simple image. Even so, after the second step of compression, the PSNR value is 37,003 dB, indicating that the selected average distortion curve parameter effectively corrects the QS value. The variance in the second step has also decreased from 18,076 to 1,7359. A similar situation also occurs when the PSNR_{des} is 40 dB.

In summary, the adaptive two-step compression method has improved its accuracy satisfactorily after introducing complexity prediction. The experimental results also show that the method is robust and can still work even in extreme conditions.

3.2.2 Adaptive two-step method for providing the desired visual quality for SPIHT

Our previous works have proved that this method works well for the DCT-based coder AGU and HEVC-based coder BPG[42-44], and its adaptive version reduces the errors for the metric PSNR [46]. However, the accuracy of providing a desired quality for the DWT-based coder SPIHT is sometimes insufficient [45]. In this section, our goal is to develop an adaptive two-step method for SPIHT to improve accuracy. The following tasks were solved. First, a prediction of visual quality for a particular parameter value is conducted. The prediction scheme is based on the information extraction from a certain number of image blocks to perform a visual quality calculation of the image compressed for a given CCP value. A threshold is adopted as the complexity grouping; images are divided into two groups: simple and complex images. Second, the results of the grouping determine the adaptive curve model adopted. Finally, a two-step compression method is applied according to this curve.

1) Analysis of two-step method on SPIHT

In our previous studies, the result of applying the two-step method to SPIHT has occurred considerably less optimistic than for the DCT-based coder AGU [42, 45]. In particular, the largest residual errors of providing a desired metric value took place for simple structure images and/or low desired quality. The reason is that these images' rate-distortion curves differ a lot from the average one, which led to the inappropriate initial CCP and erroneous estimate of derivative used in the calculation of the corrected CCP.

A part of the average rate-distortion curve is given in Fig 3.13. Two examples of particular rate/distortion curves are also presented for comparison and detailed analysis. The curve for the test image Goldhill is very similar to the average one; therefore, the accuracy of quality provided by the conventional two-step method for this test image is high, and the residual error is appropriate [45]. However, for the simple structure image Frisco (that contains large quasi-homogeneous regions), the rate/distortion curve differs a lot from the average curve. This results in the residual error that is the largest among the test images considered in [45]. Therefore, the accuracy is worth improving, especially for simple structure images and low desired PSNR. The two-step compression method [45] requires the simultaneous fast realization of both fast discrete wavelet and cosine transforms.

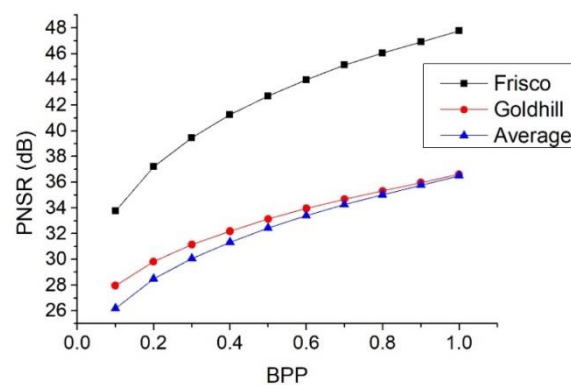


Figure 3.13 – comparison of average curve and example image curves of PSNR on BPP for SPIHT

2) Adaptive two-step method on SPIHT

Aiming at the problem that the difference in image complexity results in the mismatch of the average rate-distortion curve, an adaptive method was proposed for AGU [46]. Because of its effective improvement, and thanks to the image quality prediction

method proposed recently [41], it is possible to apply an adaptive two-step method to SPIHT.

The main idea consists of as follows. We assume that it is possible to easily, reliably, and quickly pre-classify an image to be compressed and refer it to two (or even more) classes (categories). Having average rate/distortion curves for all classes and assuming that particular rate/distortion curves for images of a given class are close to the corresponding average rate/distortion curve, it is possible to set the initial (first step) CCP better and to use a better estimate of derivative in the second step.

In this section, we have limited ourselves by considering two classes. The average rate-distortion curves were obtained from two basic image sets. In this paper, all basic images were divided into simple and complex groups. The grouping was based on the image quality prediction value for a fixed CCP, bit per pixel (BPP) in the SPIHT coder [41, 46].

The prediction approach for the SPIHT coder proposed in [41] provides an estimated value of PSNR for a given BPP using some calculation to replace the actual compression. The time consumption is about 2/3 of SPIHT compression, and the standard deviation of residual errors of providing a desired PSNR is about a few dB.

In our adaptive method, the predicted value (2.5) is utilized to pre-classify an image to be compressed into two groups, simple and complex structures. The BPP is given as 0.5, 300 8×8 random image blocks are chosen to calculate the prediction PSNR (2.5). The basic image set prediction results are shown in Figure 3.14.

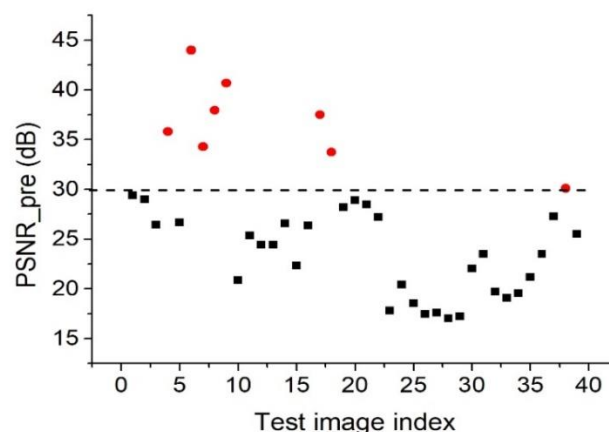


Figure 3.14 – Prediction of PSNR for SPIHT(BPP=0.5)

Let us set the PSNR=30 dB as the threshold (Horizontal dotted line in Fig. 3); if the predicted PSNR is larger than 30 dB, then the image is treated as the simple one (marked as a red dot). Otherwise, it belongs to the complex image set (marked as a black square).

As shown in Fig. 3.14, most images belong to the class of “complex structure images.” Only 7 out of 39 test images have been classified as simple structure ones. The average curves drawn from images of the two classes are shown in Figure 3.15. It can be seen from the analysis of these curves that the average values for simple images are considerably (by 10-15 dB) higher than for the complex ones due to different image complexity.

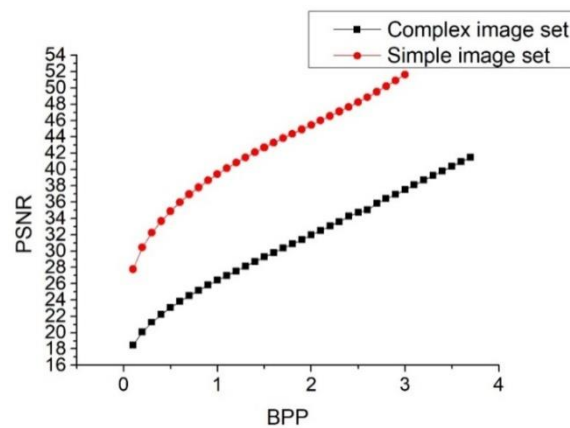


Figure 3.15 – Grouped dependence of PSNR on BPP for SPIHT

The proposed pre-classification approach is performed for each image to be compressed. The considered image is automatically classified as simple or complex images; the average curve is chosen adaptively; the two-step method is implemented according to the equations (3.4) and (3.5).

3) Validation experiment

The validation experiment has been conducted in three stages to verify its feasibility.

First, thirty-nine gray-scale images [187, 192] were chosen as the basic image set, including nine general-purpose images and thirty texture images. This image set was divided into two groups, and then serial experiments were conducted for each image and a wide range of BPP values; Finally, the average rate-distortion curves (see Fig 3.15) were obtained.

Second, twenty images [187, 192] have been chosen as the test image set to conduct the validation experiment with curve models from the basic image set. These test images

have also been split into two groups with the same prediction and classification strategy used in the basic image set.

Finally, an adaptive two-step method validation experiment has been implemented for two groups of images. Three typical values have been chosen for the metric PSNR, and the results are presented in the next section. For comparison, the results for the previous two-step method [45] for these test images have been obtained as well.

4) Result and discussion

The experiment results are presented to compare the previous two-step method and adaptive method for the SPIHT coder. The result of the previous method is shown in Table 3.15, and the result of the adaptive method is shown in Table 3.16, where the experiment was conducted for simple images and complex ones separately, but the data were combined in one table to facilitate comparison.

Table 3.15 Statistic result of basic method on test images for PSNR (dB)

PSNR_{des}	VAR_{fis}	VAR_{sec}	MAXΔ_{final}	MAE
40	37,6386	1,9183	5,9907	0,5049
35	41,831	5,4447	5,5956	1,5911
30	48,7648	13,5432	7,9147	2,5726

Table 3. 16 Statistic result of adaptive method on test images for PSNR (dB)

PSNR_{des}	VAR_{fis}	VAR_{sec}	MAXΔ_{final}	MAE
40	22,2946	2,4362	5,81	0,8618
35	24,4809	5,2552	4,7989	1,7228
30	32,9900	2,0219	5,378	0,7799

From these data comparisons, it is proved that the adaptive scheme improves the overall accuracy of the two-step method for SPIHT lossy compression. First, the variances in the first step compression have been reduced sufficiently with the better initial CCP; second, the residual errors of visual quality providing are smaller and more convergent due to the adaptive selection of the average rate-distortion curve, which is more significant at a low desired quality (30 dB).

For a detailed analysis of the adaptive method on remote sensing images [193], one example is shown in Fig 3.16. The example image was compressed for the desired quality equal to 40 dB, 35 dB, and 30 dB. The original image and decompressed image with small

CR are presented in Figure 3.16.a and 3.16.b, respectively. The image in Figure 3.16.b has an excellent quality, which is indistinguishable from the original image. The image in Figure 3.16.c has a good quality, and the distortion is not easy to notice; the image in Figure 3.16.d has relatively bad quality, but it allows to understand the content of the image in spite of distortions that are mainly concentrated in texture/detail areas; meanwhile, the high compression ratio (CR) is achieved. The example also shows that the largest residual error is observed for $\text{PSNR}_{\text{des}} = 30$ dB.

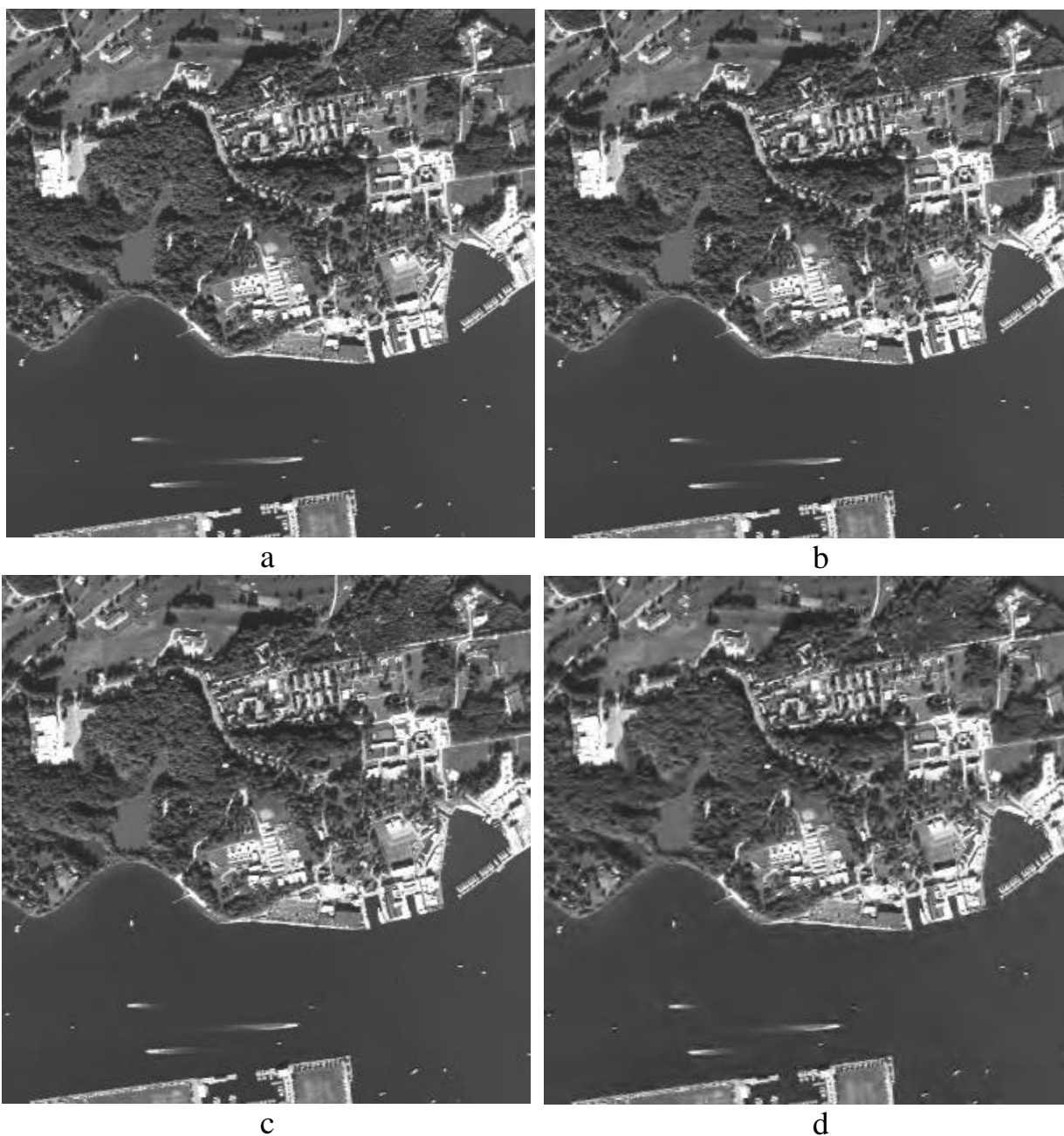


Figure 3.16 - Remote sensing image example, a) original images, b) $\text{PSNR}_{\text{pro}}=40,708$, $\text{CR}=4,086$, c) $\text{PSNR}_{\text{pro}}=33,944$, $\text{CR}=7,160$, d) $\text{PSNR}_{\text{pro}}=27,098$, $\text{CR}=19,112$

It is demonstrated that the proposed adaptive method for two-step compression is able to provide the desired quality for the SPIHT coder. The user can set a proper desired value according to requirements for a given application and achieve the highest CR.

3.2.3 Intelligent lossy compression method of providing a desired visual quality for images of different complexity

The research work in the previous section shows that for the two-step compression method in the case of unsatisfactory accuracy in the compression of some images in AGU, it is possible to carry out a simple classification of images before their lossy compression[46] and use different average curves for different groups of images. This partly helps, but there can be “strange” (super-simple) images. Besides, it is desired to carry out preliminary classification faster and more reliably.

In this section, we show that image classification can be done on the basis of entropy calculation and its comparison to a set of thresholds. This allows fast detection of “strange” images and division of other images into three groups (simple, middle, and high complexity). This results in better (more accurate) providing of the desired PSNR in the range where introduced distortions are visible.

1) Entropy-based fast classification of image complexity

Aiming to improve the accuracy of the two-step method of lossy compression, the complexity of images was introduced in [46]. Consequently, an alternative approach proposed was to divide images into simple and complex structures ones and have two average rate-distortion curves, respectively. The classification strategy in [46] was either to predict the PSNR value of an image set for some QS value(s) or to calculate CR for lossless compression (CR_{lossless}). It was supposed that these parameters being compared to thresholds are able to classify images. This approach partly helps to improve the accuracy, but the prediction of PSNR is still not fast as we desired. Another problem is that the extreme images are difficult to point out, which has a strange performance for the two-step method.

In this section, we propose a new approach to solve the problems mentioned above. Our idea is based on the assumption that the entropy is able to characterize the image complexity. In fact, entropy has been used to determine the complexity of images to detect defects on 3D printed surfaces [194]. The entropy data have been obtained for our image set (including 61 gray-scale images of the size of 512×512 pixels) to verify the theoretical feasibility. The result is shown in Fig 3.17.

For 61 images, entropy varies from 0,5 to 7,9, along with CR_{lossless} varying from 100.7 to 1. The correlation coefficients for these two indicators of image complexity are -0,74 for Pearson, -0,88 for Spearman, and -0,76 for Kendall, respectively. The larger CR_{lossless} corresponds to simpler images; meanwhile, the entropy value is smaller, and vice versa. It is easy to roughly divide the images set into three levels according to the entropy of an image: 1) complex image if entropy >7; 2) middle complexity image if $6 < \text{entropy} \leq 7$; 3) simple complexity image if entropy <6.

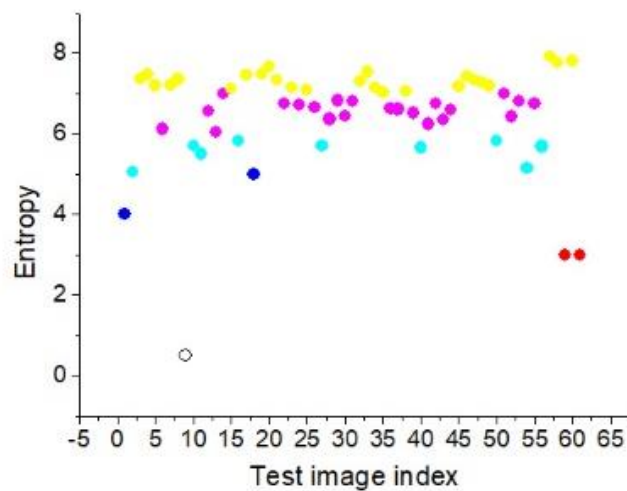


Figure 3.17 – Entropy values for 61 gray-scale images

Let us take three images as examples, corresponding to three levels of complexity. The first one is the remote sense image Frisco (see Fig.2.14), which has a large uniform area as the background, and it belongs to the simple image set according to entropy (5,8166). The second one is a texture image (numbered as #12 in Fig. 3.3), which has strong texture information, and it belongs to the middle complexity image set since its entropy equals 6,799. The last one is the image Baboon (see Fig. 2.14) with the entropy equal to 7,3579, which is a typical complex image for its rich texture information and

complex composition. The comparison of the three images proves that our classification strategy roughly corresponds to the complexity of the image.

Similarly to[46], the average rate-distortion curves have been obtained for three-level basic image sets, 4 images in the simple image set, 15 images in the middle complexity image set, and 17 images in the complex image set. These curves are shown in Figure 3.18; the curves of three example images are also present in the corresponding plots. Three average curves have the trend in common, PSNR decreases monotonically along with QS increasing, and the average curves are approximate “locally parallel” to other curves of the corresponding complexity images, but the curvature deviations of different images sets are different. Among them, the images in the simple image set have the largest diversity compared to the other two sets.

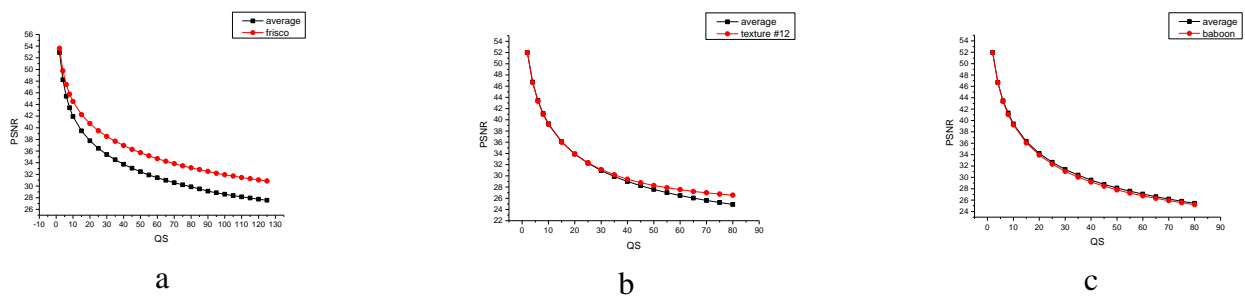


Figure 3.18 - Dependence of PSNR on QS, (a) simple images; (b) middle complexity images; (c) complex images

According to the two-step method, these curves are the most important data for calculating QS values for the first and second steps of compression for providing a desired visual quality. First, the estimated QS (QS_{est}) value is obtained from one point on the average rate-distortion curve (chosen according to the entropy of an image being compressed), where PSNR is the closest to the desired value $PSNR_{des}$. Subsequently, equation (3.1) is employed to calculate the initial QS (QS_{init}) used in the first step compression. Then decompress and calculate the visual quality of decompressed image with respect to the original image. This initial PSNR ($PSNR_{init}$) is usually quite close to $PSNR_{des}$, but it can be not close enough to meet the requirement of a user; therefore, the second step is needed to reduce the error. The QS is corrected by equation (2.14) using $PSNR_{init}$. The second step is conducted with the corrected QS (QS_{des}) and produces a compressed image file, which provides a PSNR ($PSNR_{pro}$) closer to $PSNR_{des}$ than $PSNR_{init}$.

2) Results and discussion

Test experiments were carried out on 61 images (including 36 basic images), and statistical results are shown in Table 3.17, where three PSNR values (40, 35, and 30dB) were selected as the desired quality. The result proves that the second step compression has effectively reduced the error; the variance of $PSNR_{pro}$ does not exceed $1,5 \text{ dB}^2$. Compared to the basic two-step method, the accuracy has been, in general, improved by our new approach with entropy classification. However, there is still some detailed information that needs to be noted. The errors in the simple image set are larger than in the other two sets, and the largest error in it occurs in the desired PSNR equal to 40dB, while in the other two sets, the largest error occurs in the lowest desired PSNR (as it was in our works [42, 44, 46]).

Table 3.17 -Statistical results of two-step compression method based on entropy

Image set	PSNR_{des}	VAR_{fis}	VAR_{sec}	MAX_{Afinal}
Simple test image set	40	26,1129	1,4352	4,4539
	35	20,153	1,4246	2,8279
	30	15,2624	1,1545	1,9385
Medium test image set	40	0,1583	0,0033	0,2394
	35	0,6986	0,486	0,8277
	30	2,4614	0,4111	1,907
Complex test image set	40	0,3173	0,0238	0,5398
	35	1,3866	0,1221	1,4162
	30	3,6813	0,3083	2,4022

The problems with the simple image set are due to three “strange” images presented in Fig 3.19. They are images Ruler (entropy=0,5), Texture #40 (entropy=2,9994), and Texture #42 (entropy=2,9833), their dependences of PSNR on QS are given in Fig 3.15. As one can see, these are artificially created images with a limited number of value levels (specific histograms of image values). Rate-distortion curves for them might not be monotonous or differ a lot from the average curve. This causes problems in using the expressions (3.1) and (2.14) in providing the desired quality. Then, a preliminary calculation of entropy can show in practice that one deals with a “strange” image (if entropy is smaller than 3, “be careful”).

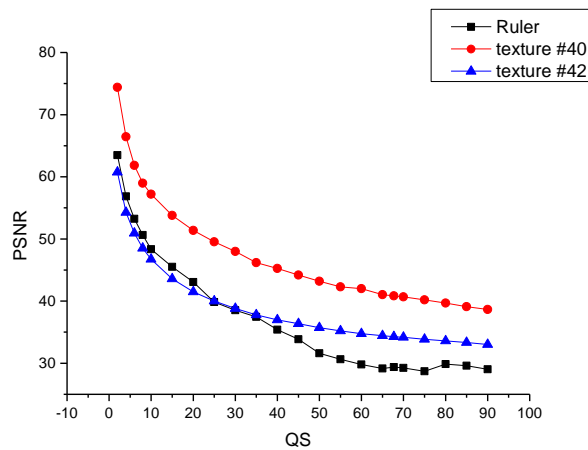


Figure 3.19 - Rate-distortion curves for three special (“strange”) images

3.3 Hybrid method of correction scheme of CCP in second step

After the exploration of methods such as adding correction constraints [44] and classifying according to the complexity of the image [46, 47], the two-step compression method has improved the visual quality control accuracy of the AGU encoder to a certain extent. In this subsection, we propose a hybrid correction method further to improve the control accuracy of the two-step compression method. The current work differs from the previous work in several ways. The main contributions of our modifications are the following:

- 1) previous two-step method has used a general image set to obtain the average rate-distortion curve and employed it to determine the initial PCC at the first step [42]. However, this curve cannot fit all kinds of images well; this leads to errors at the first step that is difficult to control. Here, we obtain the average rate-distortion curve from a particular category image set and use it in the corresponding category of images, for example, remote sensing images. This improvement reduces the error in the first step compression;

- 2) the previous two-step method as used the approximate derivative to correct the PCC in the second step; this approach works well if the error is small. However, the limitations on the possibility to apply linear interpolation arise if the error is relatively large; then, the correction at the second step occurs unable to provide the desired visual

quality sometimes [44]. It is needed to simulate the image rate-distortion curve in a better way, then predict a more precise PCC to meet the desired visual quality. We take the deviation tangent in average rate-distortion to simulate the objective image rate-distortion in the case of a large error that takes place in the first compression. This supplementary measure can produce a more accurate PCC for the second step compression and, eventually, improve the accuracy of providing the desired visual quality.

1) Two-step method on AGU for remote sensing images

In [42], typical nine grayscale optical test images have been chosen as the basic images set to obtain the average rate-distortion curve. This image set includes the images from the Kodak set, and remote sensing images, medical images. The purpose was to represent the universality of compressed images. Three metrics were employed, namely, PSNR, PSNR-HVS, and PSNR-HVS-M. Three typical values are used in the testing experiment in terms of three metrics.

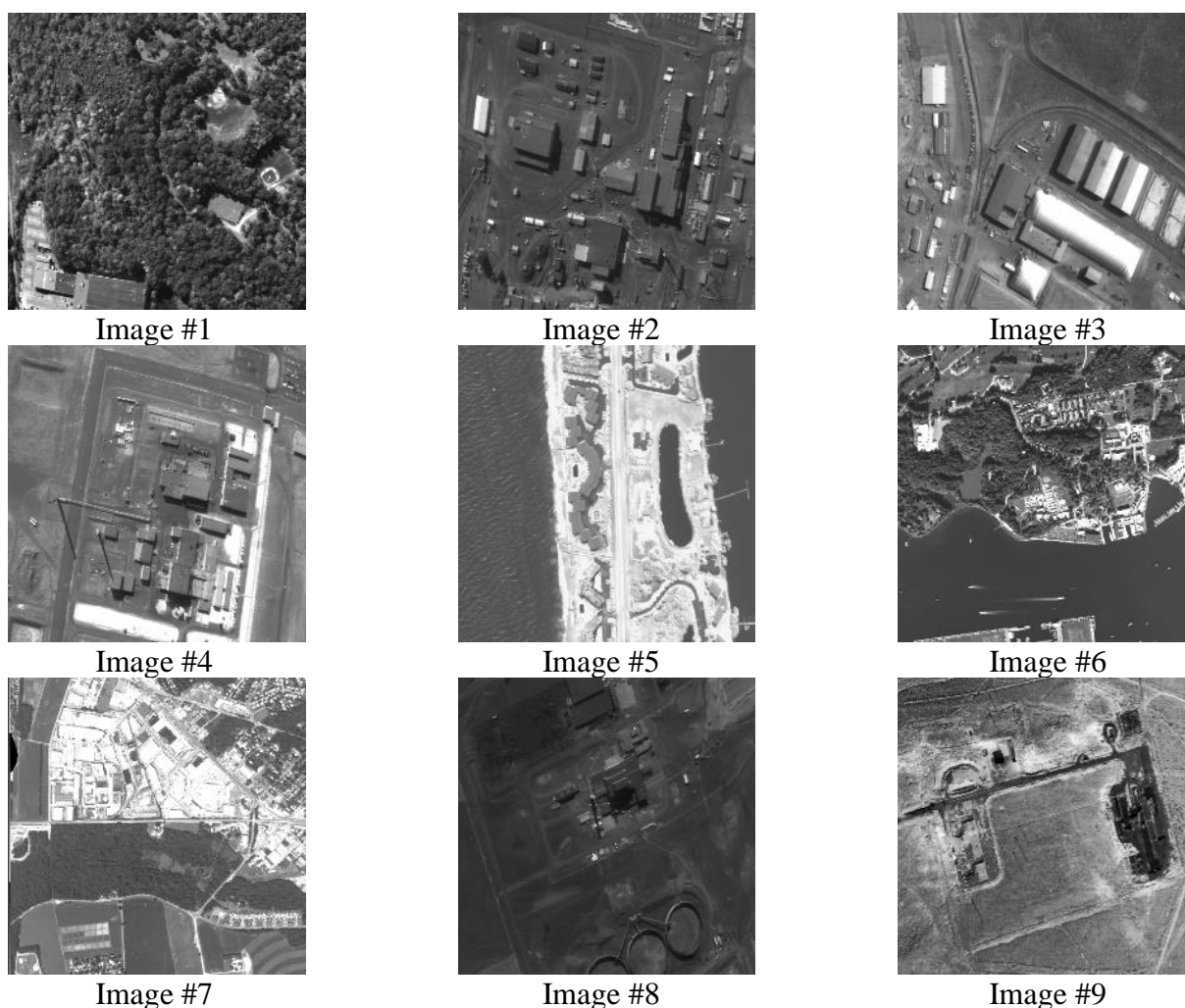


Figure 3.20 – gray-scale remote sensing image set for AGU coder

The application of the standard two-step method [42] is thus restricted. Due to modification (3.2), the variance has been reduced by one order, and the maximum error has not exceeded 1.58 dB. Meanwhile, errors less than 1 dB are desired.

This section studies remote sensing images, which are mostly high-textured images. They often contain several objects, and every object has a semantic meaning. Visual quality plays an important role in classification, target detection, and solving other typical remote sensing image processing tasks.

An experiment has been carried out for remote sensing set with the general average rate-distortion curve (the curve is obtained in [42]). The image set and image titles are shown in Fig. 3.20. In order to analyze the performance of the two-step method for RS images, the same three typical metric values are tested, respectively, 40dB, 35dB, and 30dB. Classical visual quality metric PSNR and HVS-based metric PSNR-HVS-M are employed to evaluate the visual quality of images after compression/decompression, and the statistical results are shown in Table 3.18.

Table 3.18 - Statistical results of conventional method for remote sensing images

Visual quality metric	M_{des}	VAR_{fri}	VAR_{sec}	MAX_{Δfinal}
PSNR	40	0,686	0,153	1,214
PSNR	35	2,572	0,273	1,406
PSNR	30	5,535	0,693	1,951
PSNR-HVS-M	40	0,870	0,025	0,317
PSNR-HVS-M	35	0,457	0,081	0,63
PSNR-HVS-M	30	1,015	0,071	0,485

In Table 3.18, the variance after compression in the second step is about 1/5 to 1/30 of the one in the first step. The error for PSNR does not exceed 1,95 dB, and the error for PSNR-HVS-M does not exceed 0,63dB. This result proves that the two-step method works well for remote sensing images, but the accuracy can be not sufficient enough, in particular, for PSNR_{des}=30 dB. Consequently, this method is worth modifying to improve its accuracy. Besides, it is desired to reduce the time consumption in a simple way to save computing resources.

2) Modified scheme of two-step method

In light of the shortcomings of the two-step method, we need some ways to meet the requirements of remote sensing image compression. Below we propose a modified scheme of the two-step method to improve the accuracy further and reduce the time consumption to a certain degree.

Our current work employs modifications in two aspects. First, for remote sensing images, we obtain the average rate-distortion curve using only remote sensing images. It is supposed that such a curve fits better the remote sensing images. Second, the two methods are used to correct the QS value in the second step, and the chosen strategy relates to the size of deviation in the first step.

a. Particular average rate-distortion

In this section, we choose one group of remote sensing images as the basic image set to obtain the particular average rate-distortion curve for the remote sensing data.

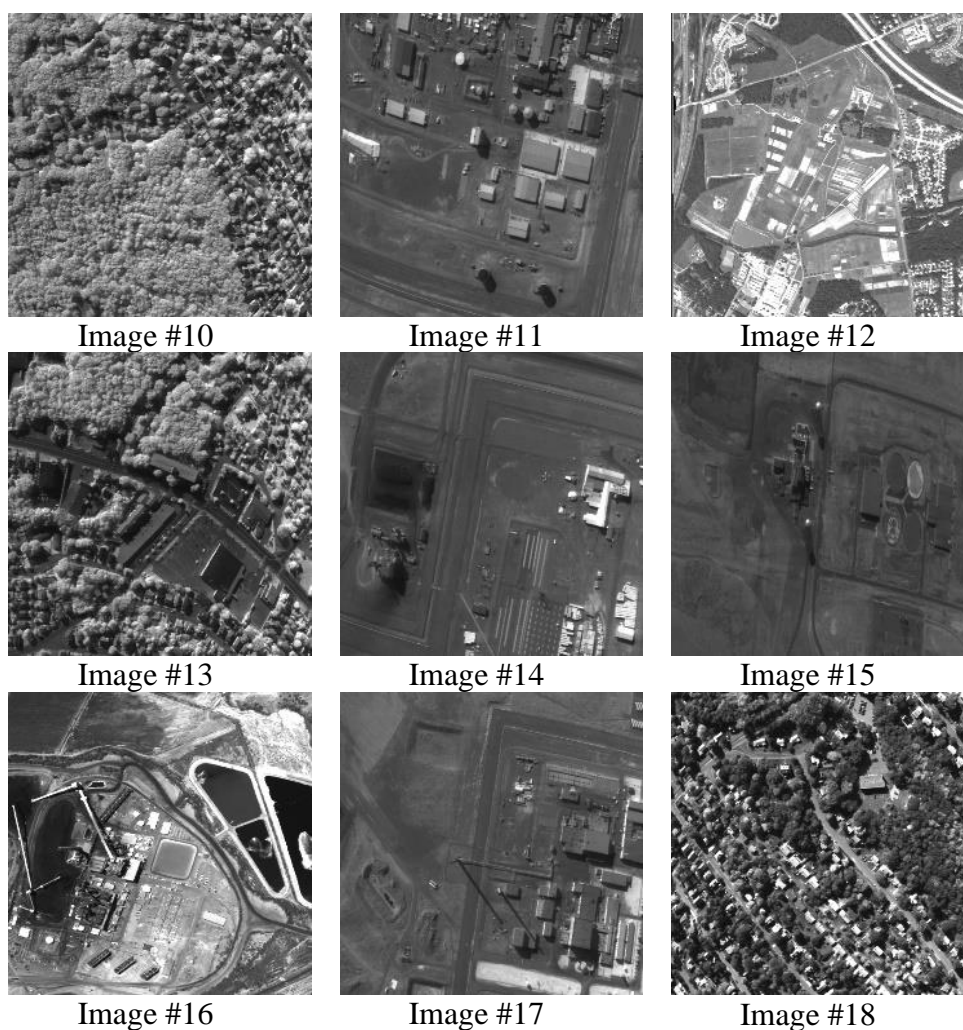


Figure 3.21 – gray-scale remote sensing basic image set for AGU coder

The basic image set and the image titles are shown in Fig.3.21. It includes nine gray-scale images of the size of 512×512 pixels, and the same interval of QS as in our previous work [42] is adopted in the original data while experiments are conducted to obtain the average rate-distortion curve. The remote sensing average rate-distortion curve is presented as $PSNR_R$ and $PSNR-HVS-M_R$ in Fig.3.22. For comparison, the previous general average rate-distortion curves from[42]are also presented as $PSNR_G$ and $PSNR-HVS-M_G$ in this plot.

It can be seen from Fig.3.22 that the remote sensing image average rate-distortion curves are similar to the earlier obtained (general) ones; in some areas, the curves almost coincide. But there are still some differences in other areas. For the metric PSNR, this range of difference is about from 45dB to 30dB; for the metric PSNR-HVS-M, it is about from 30dB to 23dB. This is the reason why the two-step method performs on PSNR-HVS-M better than on PSNR for the typical three values (40 dB, 35 dB, 30 dB) in our previous experiments. If we want to improve the accuracy of the two-step method, then the particular average rate-distortion curve needs to be adopted for remote sensing images. It can be expected that a better initial QS can be calculated according to the particular average rate-distortion curve.

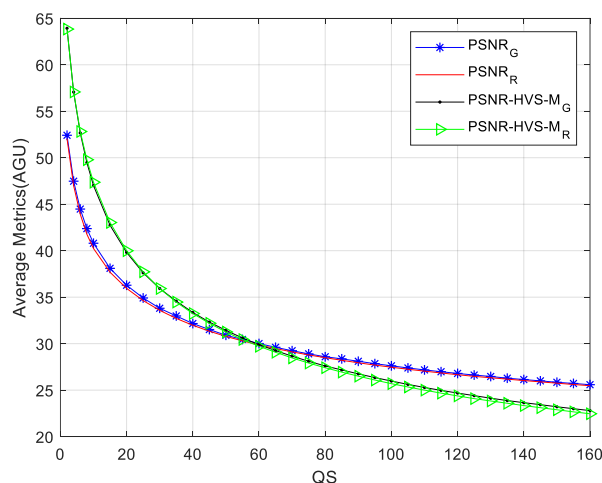


Figure 3.22 - Metrics dependence on QS for AGU

b. Hybrid correction method for QS

In the conventional two-step method[42], the correction of QS is conducted by the deviation corresponding to the initial QS. This approach is effective while the deviation is small in the first step. However, for large deviations, the positive effect of the second step

is not obvious. A modification was proposed to avoid overcorrection[18, 44, 45], it is a simple method, and it did work. But the accuracy is unstable since the corrected QS is not related to the metric value in the first step.

In most conditions, the error after the second step can be reduced to meet the desired visual quality. But it is uncertain when the error after the first step is large. Let us give some images as examples (images #8 and #9 shown in Fig. 3.20), which have abnormal results according to equations (3.1), (2.14), and (3.2). In Fig. 3.23, the desired value of PSNR equals 30dB, and the rate-distortion curves for the test image #8 and #9 are unknown before compression. We only have the average rate-distortion curve at hand.

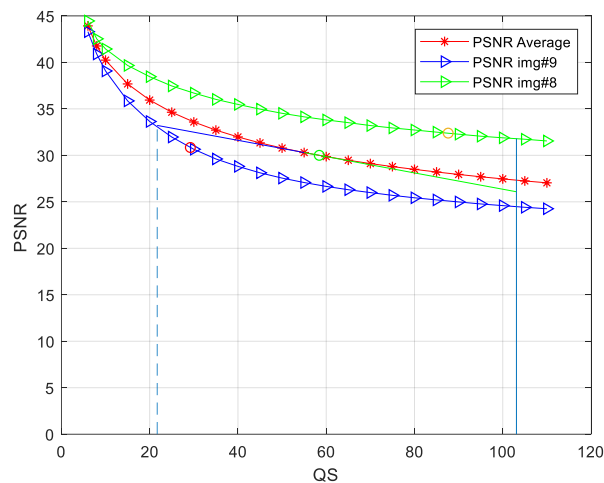


Figure 3.23 - Experiments for conventional two-step method

According to equation (3.1), for $\text{PSNR}_{\text{des}}=30\text{dB}$, the initial QS is calculated as the value of 58.424. It is drawn in green on the average rate-distortion curve (shown by red). First, let us analyze the experiment for image #9; the PSNR is 26,78 dB for $\text{QS}=58,424$, then the equation (2.14) is employed, and QS is corrected by the slash from the green point on the average rate-distortion curve, the value of QS_{des} is 21,694, and the corresponding PSNR is supposed to be 33,036 dB. However, the ΔQS is greater than $0,5*\text{QS}_{\text{init}}$. Therefore, equation (3.2) is adopted to calculate QS in the second step. Eventually, the QS is set as 29,212, and the PSNR equals 30.843 dB (it is drawn in red on the rate-distortion curve of image #9). Through data comparison, it is proved that the simple modification works for the case of image #9. However, this does not happen for image #8. In the second step, the QS is calculated as 103,1 in terms of equation (2.14), and the PSNR is supposed

as 31,795 dB then the modification is adopted to correct the QS for the $|\Delta QS| > 0.5 QS_{init}$. Consequently, the QS is set as 87,636. Then, the obtained PSNR is 32,391dB (it is drawn in red in the rate-distortion curve for image #8). This result shows that the modification still has limitations.

In this section, we propose a novel approach to correct the QS in the case if a large deviation takes place in the first step. It is shown in [42] that the difference can be easily observed visually when the error is approximately equal to 1dB. Therefore, it is expected to provide a desired quality with errors less than 1dB. If the deviation in the first step compression is larger than 1,5 dB, it is treated as a large error, and the modification is performed as the following.

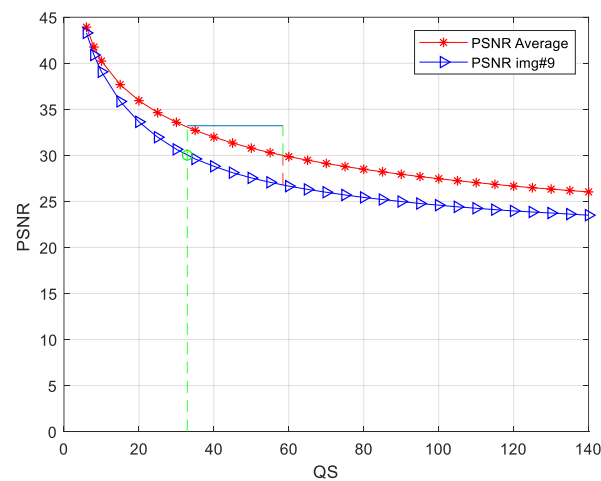


Figure 3.24 - Experiment for the new modified two-step method on the image #9

For the image #9, the PSNR after the first step compression is 3,22 dB less than $PSNR_{des}(30dB)$, which is drawn in the red dotted line in Fig.5. It can be learned from this difference that the curve of image #9 is below the average curve. The “higher” PSNR should be set on the average rate/distortion curve to obtain a better QS value. The value can be calculated as $PSNR_{des}$ plus 3,22dB, and it is 33,22 dB (drawn in the green dotted line corresponding to $x=58.424$). Corresponding to this PSNR value, the QS equals 32,899, and it is the QS value set in the second step compression. After the second step compression, the PSNR for image #9 is 30,017dB, which is better than the result in the conventional method (30,843dB).

On the contrary, for the image #8, the PSNR after the first step compression is 3,917 dB larger than $PSNR_{des}(30dB)$, then a lower PSNR value is supposed to be 26,083 dB. QS

corresponds to this PSNR value equals 137,003. After the second compression, the PSNR for image #8 is 30,786dB (drawn as a red point in the rate-distortion curve of image #8). This value is also better than the result in the conventional method (32,391dB).

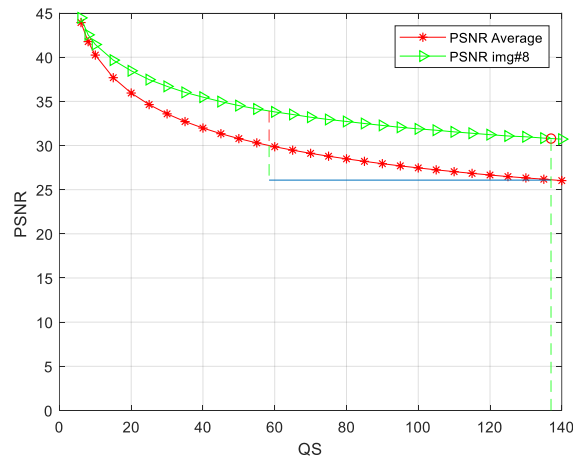


Figure 3.25 - Experiment for the new modified two-step method on image #8

3) Analysis of results

The proposed modification of the two-step method is proved to perform better than the conventional method on two images, namely, images #8 and #9. In this section, we analyze the overall performance of this method on nine remote sensing images. As in previous test experiments, the three typical values are set as the desired visual quality, 40dB, 35dB, and 30dB, respectively. The statistical results for remote sensing images are collected in Table 3.19.

Table 3.19 -Statistic result of hybrid correction method for remote sensing images

Visual quality metric	M_{des}	VAR_{fri}	VAR_{sec}	$MAX_{\Delta final}$
PSNR	40	0,543	0,124	0,972
PSNR	35	2,429	0,09	0,773
PSNR	30	5,248	0,072	0,786
PSNR-HVS-M	40	0,893	0,034	0,587
PSNR-HVS-M	35	0,449	0,016	0,296
PSNR-HVS-M	30	0,845	0,037	0,505

Compared to the conventional method, the variance in the first step is slightly improved; the variance after compression in the second step is about 1/5 to 1/70 of the one in the first step. Overall, these results are better than the ones for the conventional two-step method. The main improvement is reflected in the group of $PSNR_{des}=30dB$. The error for

PSNR does not exceed 0,972 dB, and the error for PSNR-HVS-M does not exceed 0,587dB. These results are much better than ones for the conventional method. Particularly, for the PSNR metric, the variance and maximum error are reduced effectively. The data for providing $PSNR_{des}=30$ dB are shown in Table 3.20 to understand the data in detail.

Table 3.20 Statistic and parameters of providing $PSNR_{des}=30$ dB in hybrid correction method

Test image	QS_{init}	$PSNR_{init}$	QS_{rec}	$PSNR_{pro}$
Img#1	58,424	27,982	40,990	29,926
Img#2	58,424	31,440	74,851	30,284
Img#3	58,424	31,096	70,927	30,261
Img#4	58,424	30,660	65,951	30,133
Img#5	58,424	29,987	58,276	29,999
Img#6	58,424	28,099	40,713	30,439
Img#7	58,424	27,654	38,154	30,362
Img#8	58,424	33,917	137,003	30,786
Img#9	58,424	26,780	32,899	30,017
Variance		5,248		0,072

In Table 3.20, five images have an initial error larger than 1,5dB; then, the modification scheme is adopted to improve the accuracy. All of the errors are reduced to less than 1 dB. Whether the $PSNR_{init}$ is more or less than $PSNR_{des}$, and how large the deviation is, the procedure can correct the QS in the right way.

Based on the overall analysis, it can be stated that our modification of the two-step method has two aspects of improvement:

1) All of the errors are less than 1dB (recall that human beings are able to detect difference [42] if it exceeds 1 dB). In our remote sensing image sample, 100% of images can achieve this accuracy whilst this probability is 85,2% for the conventional two-step method; the main improvement of accuracy relates to the PSNR metric, in particular, its lower value area, where the error is difficult to be controlled with the conventional two-step method;

2) Besides, this modification also reduces the time consumption in a simple way since more images achieve accuracy just after the first step compression. This is due to

getting a remote sensing average rate-distortion curve for which initial QS can be better set in the first step of compression.

3.4 Conclusions

This chapter improves the compression of grayscale images by the initial two-step compression method.

First, a constraint is placed on the range of parameter correction. This scheme can avoid the occurrence of parameter "over-correction," such as negative parameters, and ensure the positive effect of parameter correction in the second-step compression on visual quality error control, thus extending the use of the two-step method to the DCT-based ADCT coder and the DWT-based SPIHT coder.

Secondly, we propose to carry out image classification according to the image complexity and provide multiple average rate/distortion rate curves. The decompressed visual quality prediction classifies the to-be-compressed image and thus selects a suitable curve model. At the same time, entropy is also used as an evaluation index for image complexity, which can be used to pre-classify images quickly.

Finally, we propose a hybrid scheme in the compression of grayscale remote sensing images, which provides a specific average rate/distortion curve according to the image category, and adopts different correction schemes according to the size of the correction deviation value in the parameter correction, thereby further improving the precision.

CHAPTER 4

TWO-STEP COMPRESSION METHOD FOR COLOR AND THREE-CHANNEL IMAGES AND ITS USE IN REMOTE SENSING

The maximum compression ratio can be obtained by providing the desired visual quality in lossy compression of grayscale images without negatively affecting subsequent image processing. The same applies to color images and other multi-channel images. In this chapter, research is conducted on providing a desired visual quality in color and multi-channel images. The two-step compression method is extended to the color and multi-channel applications in the BPG coder. The quality of the decompressed images is evaluated by PSNR-HMA, FSIM, and MDSI visual quality assessment metrics. Finally, the effect of visual quality on the classification accuracy of remote sensing images is studied. Setting a reasonable visual quality threshold according to the accuracy requirements can effectively improve the efficiency of lossy compression.

4.1 Two-step method for color and three-channel images

A two-step method [29] is a fast and efficient choice, and it has proved to work well for several gray-scale image compression coders [22, 29]. Having a preliminary stage that employs simulations for a set of test images with getting averaged dependences, this approach provides fast processing being applied to a given image [30-32]. In view of its success in compressing gray-scale images, it is worth promoting this method use for color and three-channel images.

4.1.1 An automatic optimization method for BPG compression based on visual perception

Color and multichannel images contain richer information than the corresponding gray-scale ones. Color images are perceived by humans as reflections of surrounding real world. However, it is more complex to deal with color images and multichannel than with grayscale ones [195]. Even though the two-step method works well for BPG on grayscale

images, we still need to take into account certain peculiarities of color image perception by humans and quality assessment.

In this section, we propose and analyze an automatic optimization of a desired quality providing for the state-of-the-art BPG compression applied to color images.

1) Image sets

The basis for the two-step compression method to provide a desired visual quality in lossy compression is based on the average rate-distortion curve of this type of image, which is calculated by compressing/decompressing specific image data and calculating the visual quality corresponding to different control parameters. Then, the average value of the quality metric data is obtained. Therefore, in order to provide the desired visual quality for color image compression by BPG coding, firstly, it is necessary to select a suitable basic image set to obtain the average rate/distortion curve, and secondly, it is necessary to select the corresponding test image set to test the compression method.

In our experiment, 12 images (see Fig.4.1) from Kodak image dataset have been chosen to obtain the average rate-distortion curve off-line. This set simulates a variety of data that can be met in practice and contains various types of images including simple structure, textural, colorful, etc. All of these images have the size of 512x512 pixels.

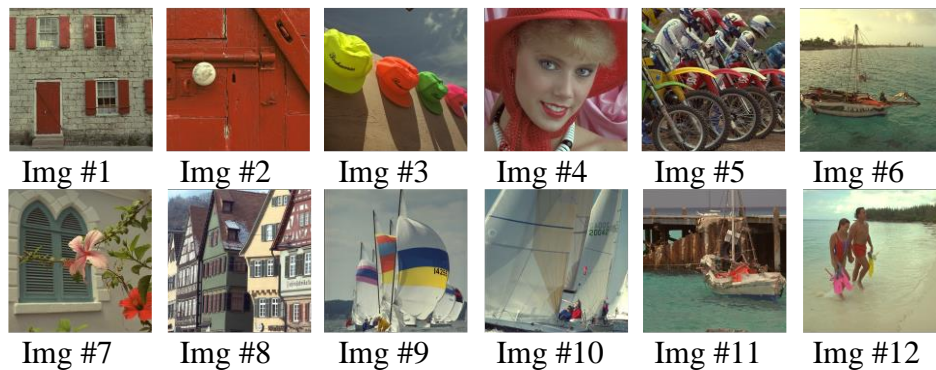


Figure. 4.1 - Basic image set using images ##1-12 of the Kodak test set

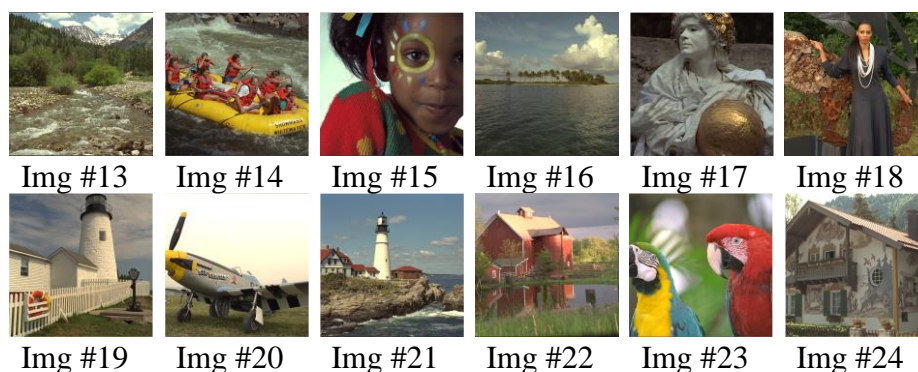


Figure. 4.2 - Verification image set using image ##13-24 of the Kodak test set

Other 12 images (see Fig.4.2) from Kodak image set are used to verify the proposed method, there are images of different scenes and people, the complexity of the images varies. This verification set can be treated as a tool to understand have we chosen the original image simulation set correctly or not.

2) Automatic optimizing method for BPG on color image

First, we need to choose a visual quality metric able to characterize well the human perception of color images. Of course, PSNR is the classic metric that can be left for further statistical analysis but it has been shown that large PSNR values do not consistently mean a high perceptual quality, some color banding and blackness are not well-captured by PSNR [136].

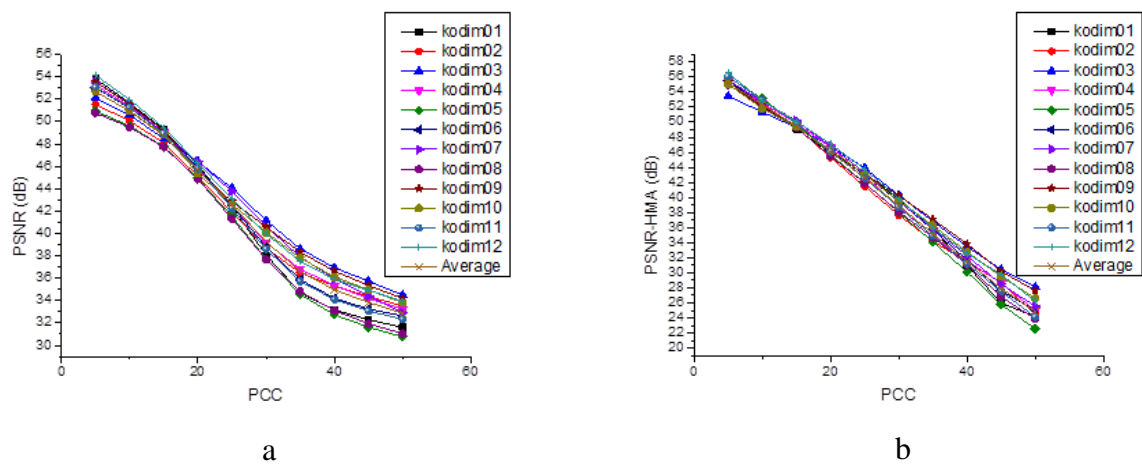


Figure 4.3 – The dependences of visual quality metrics on Q for 12 color basic images

PSNR-HMA [196, 197] is a modified image visual quality metric for HVS taking into account. In particular, it takes into consideration the perception of mean shift and contrast change. Advantages of PSNR-HMA are the following: a) it is expressed in dB and covers a wide range of values starting from 15-20 dB (annoying distortions, very bad quality) to >40 dB (with high probability, invisible distortions); b) its properties have been quite thoroughly studied [31]; c) it can be easily and quickly calculated, its open access code is available at [198]. One more advantage is rate-distortion curve behavior close to linear (Fig. 4.3.b) – this is important for the two-step procedure based on linear interpolations.

Fig. 4.3.b shows the rate-distortion curves PSNR-HMA vs. PCC (Q for the BPG coder) for 12 test images which come from Kodak image set. The plots are similar to those

of PSNR vs. PCC in Fig. 4.3.a. Despite of similar behavior, the plots are individual for each test image.

Similar to the implementation of the two-step method in BPG grayscale image compression, the calculation of the metric PSNR and PSNR-HMA is displayed in equations (4.1)-(4.4).

$$Q_{init} = Q_{est} + \frac{PSNR_{des} - PSNR_{ave}}{M'} \quad (4.1)$$

$$Q_{des} = Q_{init} + \frac{PSNR_{des} - PSNR_{init}}{M'} \quad (4.2)$$

$$Q_{init} = Q_{est} + \frac{PSNR-HMA_{des} - PSNR-HMA_{ave}}{M'} \quad (4.3)$$

$$Q_{des} = Q_{init} + \frac{PSNR-HMA_{des} - PSNR-HMA_{init}}{M'} \quad (4.4)$$

Here the PSNR is calculated as color version, a specific feature is that Q can be only integer. Thus, the determined values of Q in (4.1)-(4.4) have to be rounded-off to the nearest integer.

3) Analysis of the accuracy

We focus our attention on lossy compression with either invisible distortions or with distortions that are visible but are not annoying. This approximately corresponds to the range from 33dB to 38dB in terms of PSNR. Experiments performed for the database TID2013 have shown that PSNR about 35dB can be considered as distortion invisibility threshold. In our experiment, four values in this region have been set as desired ones, namely 40, 38, 35, and 33 dB. The same desired values have been also used for the metric PSNR-HMA. The obtained results are given in Table 4.1.

Table 4.1 -The obtained statistical data (training set).

Quality metric	M_{des}	VAR_{fir}	VAR_{sec}	ΔCR
PSNR	40	1,0067	0,0966	23,65
PSNR	38	1,298	0,1253	46,126
PSNR	35	1,511	0,0804	144,477
PSNR	33	1,0347	0,2058	737,175
PSNR-HMA	40	1,0615	0,0606	19,395
PSNR-HMA	38	1,1131	0,0651	28,45
PSNR-HMA	35	1,0861	0,0442	45,293
PSNR-HMA	33	1,1514	0,0575	72,542

It can be also interesting to look at compression ratio and its variation for a given quality. To get imagination, we present the values of ΔCR that denote the maximal difference between CR values corresponding to different images for the same M_{des} .

It can be seen from data in Table 4.1 that our optimization scheme can indeed provide visual quality for BPG color image compression, at least, in the human perception region under interest. The max VAR_{sec} go beyond 0,21; it has considerably (by about one order) reduced due to the second step. This indicates that the provided visual quality is enough close to the desired visual quality which can be set by a user. In some cases, even one (the first) step is enough to produce the desired quality for BPG applied to color image. (the examples are the test image #17 in Table 4.2 and test images #15,17,19 in Table 4.3).

Analysis of ΔCR shows that, for the same desired PSNR, the compression ratio can be sufficiently different, especially for small values (33 dB) of the desired metric. The differences ΔCR for PSNR are larger than for the same PSNR-HMA.

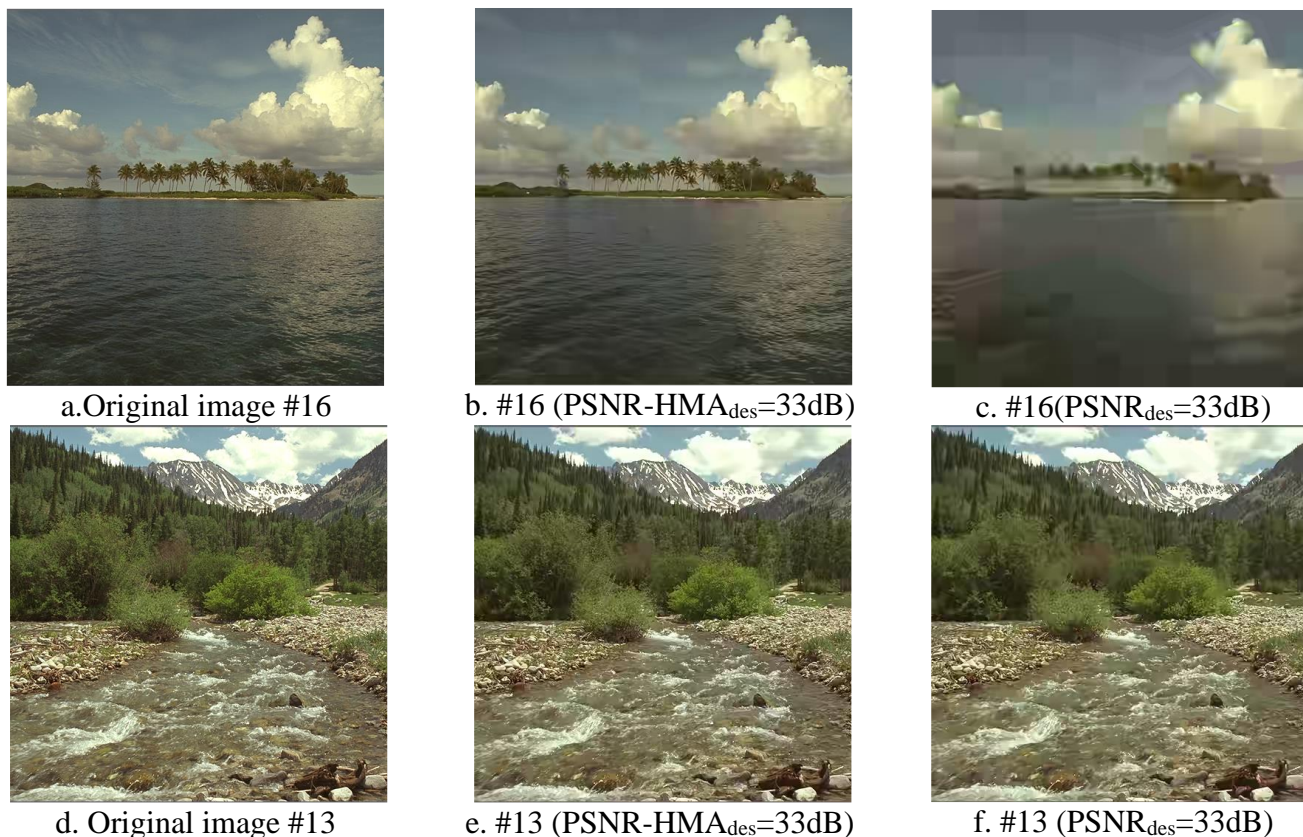


Figure 4.4 -Two examples with the desired metric value equal to 33dB for color image

BPG compression

Let us use the images #16 and #13 as examples. The original images are shown in Fig.4.4.a and Fig.4.4.d, the decompressed images with the desired PSNR-HMA=33dB are shown in Fig.4.4.b and Fig.4.4.e, and the reconstructed images with the desired PSNR=33dB are given in Fig.4.4.c and Fig.4.4.f. Comparing them, we can find that the distortions are practically invisible when the desired PSNR-HMA=33dB, the CR values are equal to 73,192 and 20,133, respectively. But when the desired PSNR=33dB, the distortions for the test image #13 are still invisible, while the distortions for the test image #16 are annoying, the CR values are equal to 768,731 and 31,575, respectively. The detailed data (for PSNR-HMA_{des}=33dB) are given in Table 4.2. In turn, the detailed data (for PSNR_{des}=33dB) are presented in Table 4.3.

In Table 4.3, the image #16 has the largest CR (768,731) while the highly textural image #13 has the smallest CR (31,575). The provided PSNR values are equal to 33,255dB and 32,408dB, respectively, both of them have the error less than 0,6 dB. According to CR for the test image #16, the obtained result is amazing. However, the visual quality of decompressed image is unsatisfactory (see Fig.4.4.c), the distortions are obvious compared to the original image. Meanwhile, the visual quality for decompressed image #13 is acceptable to human perception.

Table 4.2- Statistical data for the desired PSNR-HMA_{des}=33dB of color image on BPG coder (verification set)

Test image	Q _{init}	PSNR-HMA _{first}	Q _{des}	PSNR-HMA _{prov}	CR
kodim13	38	32,373	37	33,243	20,133
kodim14	38	31,493	36	33,086	26,605
kodim15	38	32,352	37	32,936	62,066
kodim16	38	33,882	39	33,088	73,192
kodim17	38	33,23	38	33,23	56,905
kodim18	38	31,893	36	33,384	29,041
kodim19	38	33,692	39	32,981	59,072
kodim20	38	34,965	41	32,686	92,675
kodim21	38	34,104	40	32,61	53,549
kodim22	38	32,212	37	32,776	49,855
kodim23	38	33,92	39	33,236	86,915
kodim24	38	32,284	37	32,965	32,213
Variance		1,1614		0,0575	

In Table 4.2, the analyzed metric is PSNR-HMA with the desired value equal to 33dB. The provided values of PSNR-HMA for the test images #16 and #13 are equal to 33,088dB and 33,243dB, respectively. And the visual quality of both images is satisfactory, the distortions are almost invisible. Analysis of variance values in Tables 4.2 and 4.3 shows that accuracy after the second step has radically improved and this shows expedience of using the two-step procedure. In addition, it shows that the simulation set has been chosen correctly.

Thus, some statements can be drawn. Firstly, there are distortions due to lossy compression that cannot be characterized by PSNR adequately even if a desired PSNR is provided and it is in the considered range. Because of this, we propose to use the metric PSNR-HMA to meet the human visual perception better and since it is more “reliable” than the metric PSNR. Secondly, by adopting the automatic optimization scheme for BPG, one can increase (maximize) the compression ratio while the provided visual quality meets the requirement.

Table 4.3- Statistical data for the desired $PSNR_{des}=33dB$ (verification set).

Test image	Q_{init}	$PSNR_{first}$	Q_{des}	$PSNR_{prov}$	CR
kodim13	49,000	31,267	40,000	32,408	31,575
kodim14	49,000	31,481	41,000	32,902	54,828
kodim15	49,000	33,034	49,000	33,034	385,121
kodim16	49,000	33,524	51,000	33,255	768,731
kodim17	49,000	33,078	49,000	33,078	301,304
kodim18	49,000	31,813	43,000	32,780	84,758
kodim19	49,000	33,030	49,000	33,030	241,745
kodim20	49,000	34,790	51,000	34,299	370,091
kodim21	49,000	33,194	50,000	33,041	303,233
kodim22	49,000	32,770	48,000	32,902	306,416
kodim23	49,000	33,905	51,000	33,426	367,731
kodim24	49,000	32,256	45,000	32,945	103,047
Variance		1,0347		0,2058	

The lower limit of $PSNR_{des}$ is about 33 dB whilst it is about 28 dB for the metric PSNR-HMA, this is because the PCC Q for BPG cannot be larger than 51 (there is the test image kodim20 in Table 4.3 for which the provided PSNR is equal to 34,3 dB although the maximal possible Q has been used). This also means that if the values of Q_{des} calculated

according to (4.2) or (4.4) are larger than 51, then compression with $Q=51$ has to be carried out at the second step.

As one can see, in some cases, the metric value provided at the first step is approximately equal to the desired value, then the procedure can quit after the first step. This makes it possible to skip the second step and, thus, to accelerate the two-step procedure. For this purpose, a precision (maximal allowed error) can be preset - if the error of metric providing at the first step is less than a certain threshold value, the procedure is supposed to quit. According to data in Tables 4.2 and 4.3, variance after the second stage is about $0,06\text{dB}^2$. Then, the standard deviation σ is about $0,25\text{dB}$. Then, supposing the Gaussian distribution of errors and requiring the provided metric values to be within the limits from $M_{\text{des}}-2*\sigma$ to $M_{\text{des}}+2*\sigma$, the following stopping rule is possible: if the error $|M_{\text{des}}-M_{\text{init}}|$ is less than $0,5\text{ dB}$, then quit.

4.1.2 Providing a desired quality of BPG compressed images for FSIM metric

It is desired to examine the performance of other adequate visual quality metrics in the two-step compression method.

Concerning adequate quality metrics, color version of FSIM [155] (denoted as FSIMc) is one of the best visual quality metric. Standard FSIM and its color version have high rank order correlation with mean opinion score (MOS) for images in databases that are corrupted by distortion types typical for remote sensing, medical and customer-oriented applications including lossy compression. This explains our interest just to this metric.

1) Rate-distortion curves for BPG color image compression according to the metric FSIMc

The compression technique BPG has two versions: one is oriented on color image compression and another – on grayscale image compression. Here we consider the former one where its realization has been taken from [113]. Dependences of quality characterized by a chosen quality metric (FSIMc in our case) on Q are individual for any compressed image although there are some general tendencies.

To illustrate them, consider several test images small copies of which are presented in Fig. 4.5.

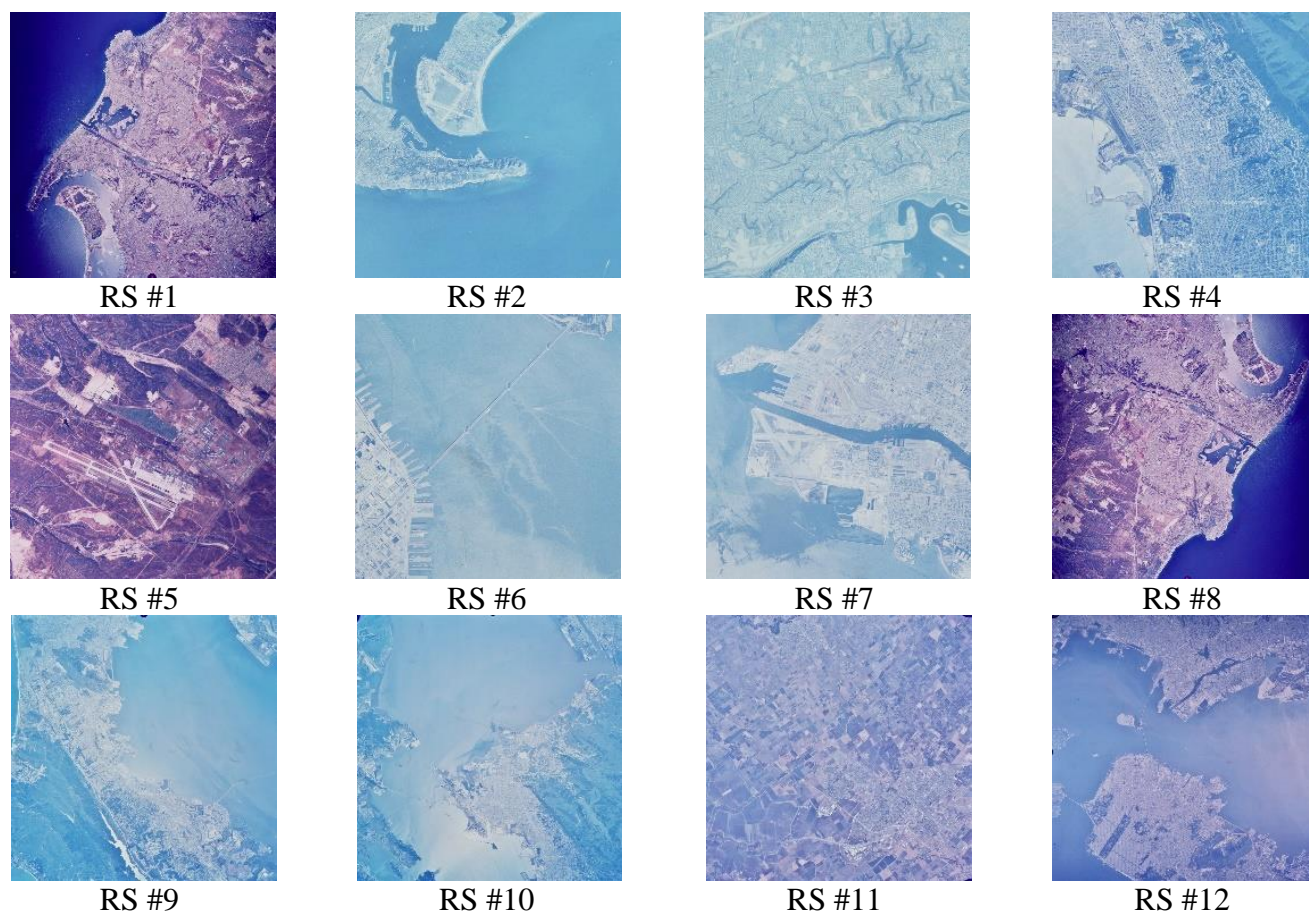


Figure 4.5 - Small copies of the test color image

The rate-distortion curves for these images are given in Fig. 4.6. As one can see, distortions cannot be observed for $Q < 20$ where all FSIMc values are close to unity (larger than 0,995). Then, starting from Q about 25, the curves start to diverge where for $Q > 35$ there is a considerable difference in compressed image visual quality depending on image complexity.

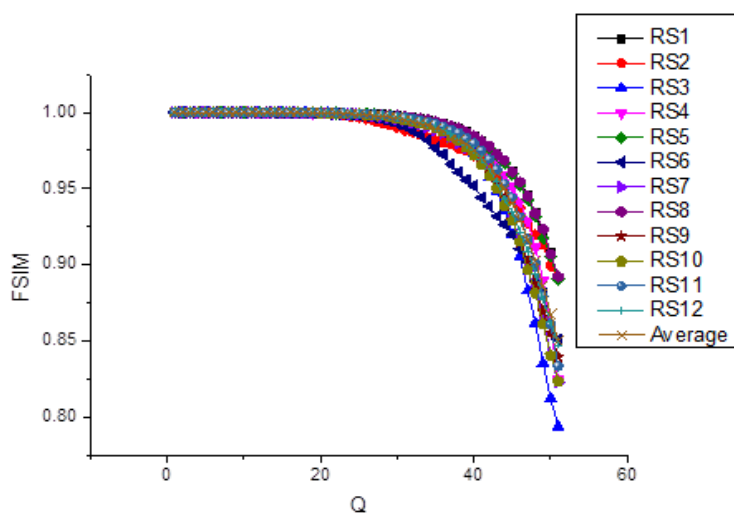


Figure. 4.6 - Examples of rate-distortion curves for the BPG coder (FSIMc).

For example, the test images RS03 and RS06 have the lowest quality whilst the test images RS01 and RS08 have the highest visual quality. This example clearly demonstrates that, for a given Q in the limits from about 30 and till $Q=51$ the quality of images can vary in rather wide limits. Thus, the task of providing a desired quality of compressed data according to the chosen quality metric is important.

Also note that it is stated that FSIMc varies from 0 to 1, but, in fact, for a wide range of Q variation the FSIMc values concentrate in the limits from 0,8 to 1. Another observation is that FSIMc is a monotonous and smooth function that can be locally approximated well by polynomials of the first or second order.

It can be also interesting to understand what differences in FSIMc values can be detected visually. This is important to understand what is an appropriate accuracy of providing a desired FSIMc. For this purpose, we have obtained compressed images with different values of FSIMc and compared them visually. Some results are given in Fig. 4.7. The image in Fig. 4.7,a cannot be distinguished from the original (uncompressed) image. Small changes can be detected (some small-sized objects become smeared) by comparing the images in Figures 4.7,a and 4.7,b. Meanwhile, it is difficult to find differences between the images in Figures 4.7,b and 4.7,c. This means that for FSIMc about 0,99 the errors of providing FSIMc about 0,003 are acceptable.

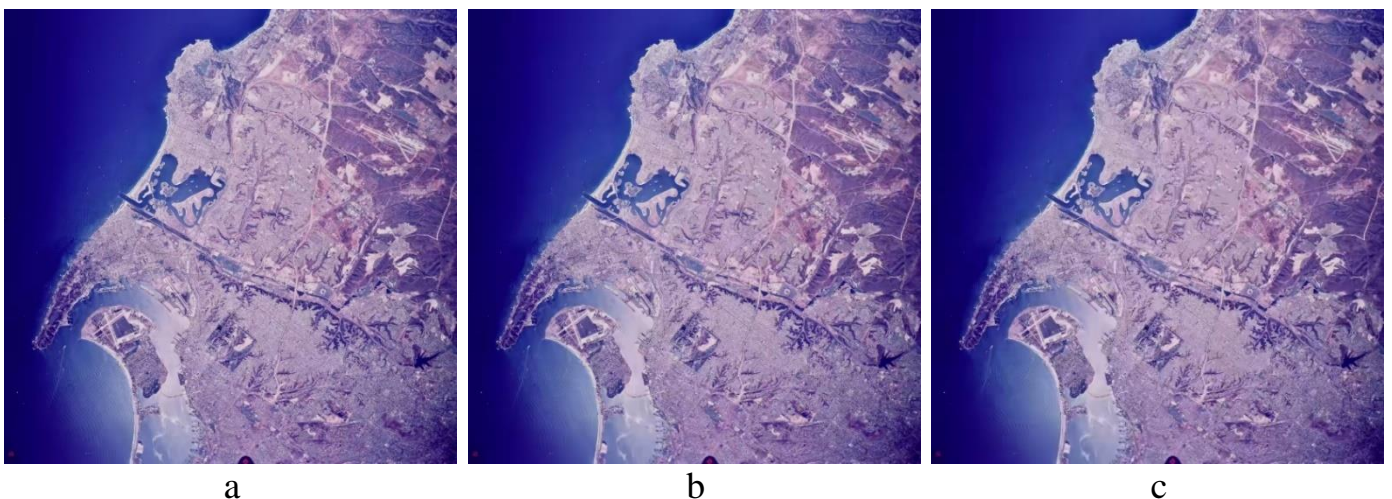


Figure. 4.7 - Compressed images with FSIMc equal to 0,995 (a), 0,989 (b), and 0,985 (c)

Similarly, Fig. 4.8 presents two compressed images with larger CR values (smaller FSIMc values). These images are very similar, but they both differ from the compressed

image in Fig. 4.7.c. This means that for FSIMc about 0,97 the errors of providing a desired FSIMc can be larger, about 0,005.

Thus, now we know practical requirements to accuracy of providing a desired FSIMc for practically important values of FSIMc and would like to see what is possible in practice.

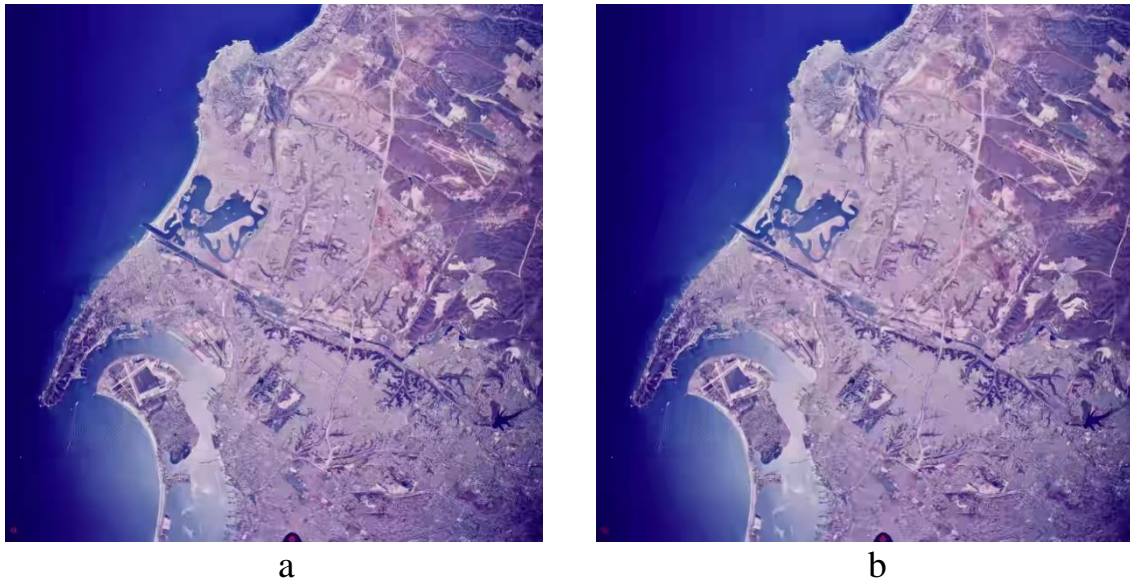


Figure 4.8 –Compressed images with FSIMc equal to 0,973(a),0,967 (b)

2) Two-step algorithm for BPG color image compression with the metric FSIMc

The two-step procedure is based on average rate/distortion curve, which is obtained in advance from a group data of a certain number of basic images. Here, twelve remote sensing images are chosen as the basic image set (shown in Fig.4.5), serial compression/decompression experiments are conducted by varying Q from 1 to 51, the visual quality data of decompressed images are recorded in tabular form as FSIMc values depending on Q . Subsequently, we have averaged the data of twelve images for each value of Q , and built the average distortion-ratio curve. This average rate/distortion curve is given in Fig. 4.6.

Two-step compression algorithm is realized in two stages. At the first stage, an image is compressed with initial value of Q , which is calculated from the obtained average rate/distortion data array. For BPG encoder, the initial Q equals to the estimated Q according to average rate/distortion array. There are two reasons for this. First, when simulating the average rate-distortion curve model, the interval of Q is set as 1. Second, the BPG only supports the integer form parameter, the rounding-off is essential for Q_{init}

before using in compression, the initial Q equals to the estimated Q . Therefore, the initial Q is replaced by estimated Q to be used in the first step compression. Let $FSIMc_{des}$ denote the desired visual quality value preset by a user, then $FSIMc_{ave}$ is the value closest to $FSIMc_{des}$ at the right end of the corresponding interval of the average rate/distortion curve data. Q_{est} is the value of Q corresponding to this $FSIMc_{ave}$ on average rate-distortion array. An image in hand is compressed the first time with Q_{init} , and followed by decompression.

Before the second stage, $FSIMc$ is calculated for the decompressed image to evaluate its visual quality with respect to the original image, the value is recorded as $FSIMc_{init}$. The visual quality value $FSIMc_{init}$ usually occurs to be quite close to the desired value, and, in some cases, this value is the closest the BPG coder provides to the desired visual quality, then the second step compression is not needed. But, in other cases, the accuracy can be further improved. Whether the second step compression is required is determined by the correction value of Q , which is calculated by equation (4.5).

$$Q_{des} = Q_{init} + \frac{FSIMc_{des} - FSIMc_{init}}{M'} \quad (4.5)$$

3) Result analysis

For testing the two-step compression method using $FSIMc$, four typical values are chosen, 0,99,0,98,0,97, and 0,96, respectively. Serial experiments are carried out for the basic image set, the data of each image are recorded in Tables, and the statistical results is shown in Table 4.4.

Table 4.4 - The statistical data of the basic color image set for BPG coder ($FSIMc$)

M_{des}	VAR_{fir}	VAR_{sec}	$MEAN_{sec}$
0,99	$2,86 \times 10^{-5}$	$2,99 \times 10^{-6}$	0,0007
0,98	$7,22 \times 10^{-5}$	$4,34 \times 10^{-6}$	0,0003
0,97	$1,04 \times 10^{-4}$	$2,67 \times 10^{-6}$	0,0016
0,96	$1,49 \times 10^{-4}$	$1,53 \times 10^{-5}$	0,0003

In this section, we calculated the mean bias of $FSIMc$ finally provided in the second step, expressed as $MEAN_{sec}$, the data gathered in Table 4.4 suggest that the two-step algorithm of $FSIMc$ works well in the considered conditions, since the variance after the second step compression has dropped by approximately an order of magnitude, and the mean error does not exceed 0,0016, i.e. the provided values are practically unbiased.

VAR_{fir} and VAR_{sec} values are considerably larger for smaller desired values of $FSIMc$. However, as we have shown in the previous Section, this is not the problem since requirements to accuracy become less strict when the desired $FSIMc$ reduces.

To further check whether or not the average curve model obtained from the basic image set works for other images, we have chosen other 12 remote sensing images to conduct the verification experiment. These images are numbered as RSI #13...#24. The statistical results for the validation image set are shown in Table 4.5.

Table 4.5 -The statistical data of validation color images set for the BPG coder ($FSIMc$)

M_{des}	VAR_{fir}	VAR_{sec}	$MEAN_{sec}$
0,99	$1,96 \times 10^{-5}$	$6,62 \times 10^{-7}$	0,0007
0,98	$3,32 \times 10^{-5}$	$2,97 \times 10^{-6}$	0,0003
0,97	$3,83 \times 10^{-5}$	$5,03 \times 10^{-6}$	0,0003
0,96	$6,12 \times 10^{-5}$	$6,54 \times 10^{-6}$	0,0007

As one can see, the tendencies and values for verification experiment are similar to Table 4.4, the accuracy is improved in the second step, and the mean error does not exceed 0,0007. This result proves that the basic image set was chosen correctly, and the average curve model works well for remote sensing images of the same size. More detail data for 12 validation images for $FSIMc_{des}=0,96$ are shown in Figure 4.9.

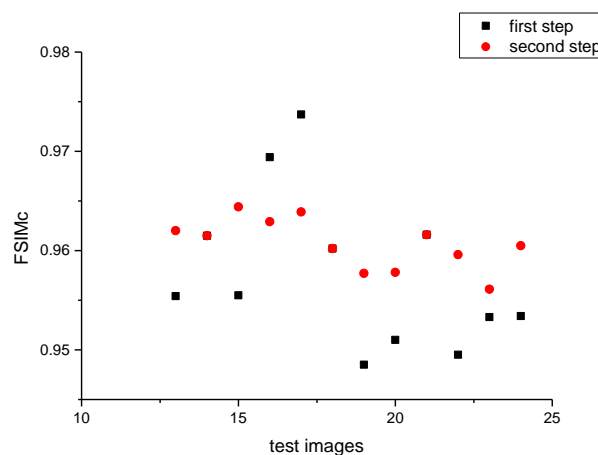


Figure4.9 - The result of validation image set ($FSIMc_{des}=0,96$).

In this example, two group data are displayed, the data obtained in the first step compression and the data gathered in the second step, respectively. On the basis of data observation in Fig4.9, it follows that the accuracy is improved due to the second step

compression, since visual quality provided in the second step is either closer to the desired value than the ones provided in the first step or remains the same (as, e.g., for the test images RSI # 18 and 21). Besides, the maximum error does not exceed 0.005, this error is acceptable in practice according to analysis in Fig 4.7 and 4.8.

In addition, as we have mentioned earlier, not all images need the second step compression. There are three images in Figure 5, which are marked as overlapping red dots and black squares. For these three images, the Q_{des} after rounding-off equals to Q_{init} , the two-step algorithm skips the second step compression to save time in such situations. In Figure 4.9, three out of twelve images need only one step compression to meet the requirement of accuracy, whilst other 9 images need the second step to improve the accuracy. For all verification experiments, 24% images need only one step of compression to provide the desired visual quality.

4.2 Quality control for the BPG lossy compression of three-channel remote sensing images

Considerable attention has been paid to metrics able to characterize the quality of RS data [36, 178, 199], including the artificial visible-like images based on SAR data generated with the use of deep CNNs [200]. In particular, special attention has been paid to the so-called visually lossless compression for RS image browsing and other applications [165, 201-204]. This is important since compressed RS images are often subject to visual inspection. The necessity to provide a desired compression ratio and quality quickly enough is important in practical applications where processing time and resources are limited [205].

It has been demonstrated recently [36] that the Mean Deviation Similarity Index (MDSI) [158] and some other elementary metrics can perform well in the characterization of three-channel RS images with distortions typical for remote sensing imagery, including distortions caused by lossy compression. It has also been shown that lossy compression, under certain conditions, can lead to practically the same or even better performance of image classification compared to the classification of original uncompressed data [54, 125, 142, 206, 207]. This may happen when noise suppression is observed or if distortions

cannot be detected visually [54]. This means that two benefits can be provided simultaneously—one obtains the CR that sufficiently differs from unity and an improved (or, at least, not worse) classification is observed. Additionally, sufficient work has been carried out towards accelerating lossy compression while attaining a predefined desired quality.

The first research goal of this Section is to check whether or not it is possible to apply the two-step method for compressing multi-channel RS images—more precisely, three-channel images that include color images and vision range data of multi-spectral imagery. The second goal of this Section is to investigate some important properties of the MDSI metric, pre-selected as the most appropriate, and to verify its usefulness for the proposed two-step approach. We also analyze the degree of accuracy of the MDSI that should and can be achieved in practice.

4.2.1 Metrics for the Assessment of the Visual Quality of Three-Channel RS Images

1) Properties of Three-Channel RS Images

Compared to conventional color images (photos), RS images have some specific features. First of all, they are usually more highly structured and each object has a semantic meaning [208], whereas natural images are more chaotic. These objects present in RS images need to be analyzed in further stages of data processing, particularly target recognition, classification, segmentation, and parameter estimation. For example, the main goal of the image segmentation process related to partitioning an image into a set of homogeneous segments, in terms of chromatography or texture, is highly important for remote sensing data [78]. Meanwhile, the fact that RS images often include large areas of background, which is much less important than foreground objects [87], may also be taken into account in compression.

Secondly, the correlation between component images of three-channel RS data can differ from the correlation of red, green, and blue components of color images. Whilst, for color images, the cross-correlation factor is usually around 0,7 [173], the correlation factor between components of three-channel RS data may be significantly higher [209]. This

might influence noticeably the final compression performance if 3D approaches are applied.

Furthermore, there are no commonly accepted databases of “pristine” (reference, distortion-free) RS images. Moreover, types of distortions inherent for RS images and color images partly coincide but are partly different. For example, image dithering is not met in practice in RS images. Meanwhile, speckle noise is not typical for color images but it might be an important factor for RS images of a special kind (synthetic aperture radar ones). This obstacle prevents the direct use of color image databases for making conclusions and recommendations for RS images. However, recently, the TID2013 dataset has been indirectly used to analyze the subsets of distortion types that might be present in RS images. This has allowed the determination of good elementary visual quality metrics for the adequate characterization of RS image quality and the design of combined metrics. Already known visual quality metrics, used in a combined metric [36] as one of the inputs, are referred to as elementary metrics. In particular, the MDSI [158] has been presented as one of the best elementary metrics [36]. Hence, more details about this metric are provided below, together with an explanation of why it has caught our attention.

It should be kept in mind that the visual quality metrics describe the quality of data from a specific viewpoint and the relation between visual quality metrics and, e.g., text recognition from document images or image classification accuracy is not fully known [206]. Nevertheless, preliminary results of the classification of compressed images have already demonstrated that visual quality has a high correlation with classification accuracy, especially for classes represented by small-sized, prolonged, and textural objects, i.e., for classes that are quite heterogeneous [36]. Since high-frequency information can be lost due to lossy compression for large CR, it might harm the classification as well.

2) The MDSI Metric and Its Properties

An important property of the MDSI metric is that, during its computation, a gradient magnitude is used to measure structural distortions, whereas chrominance features are used to measure color distortions (recalling that both these types of distortions are equally important for three-channel RS images). Subsequently, the two obtained similarity maps are combined to form a gradient-chromaticity similarity map. Differently than for SSIM

and FSIM, the deviation pooling strategy is used to compute the final quality score. In comparison to previous research, this new gradient similarity map is more likely to follow the human visual system (HVS).

Providing the desired visual quality in lossy compression is a challenging task; however, it would be possible if a metric value was associated with a certain level of quality. It could also be useful to know a range of metric values for which distortions are practically invisible. In lossy compression, the desired visual quality is often within a certain range, also for RS images. As illustrated by some already completed analyses based on other metrics [54], the lower limit is such that lossy compression has no negative impact on further image processing. Concerning the upper limit, the lossy compression should provide a higher CR than possible to achieve by the lossless compression (limited by entropy). A reasonable threshold should be set in such a way that the introduced distortions are invisible, so the visual quality of compressed data should be identical to lossless compressed images but higher CR can be achieved. This threshold is around 40 dB in terms of the PSNR-HVS-M metric multi-channel RS images [54].

To provide a reasonable range for the metric MDSI, we have tested 3000 color images of the database TID2013 to obtain the statistical data results [145] and put them into three categories, namely visible, just noticeable, and invisible distortions. Since MOS values have been provided for each image in the TID2013 dataset, the scatter plot for MDSI vs. MOS for the three mentioned classes of images is shown in Figure 4.10.

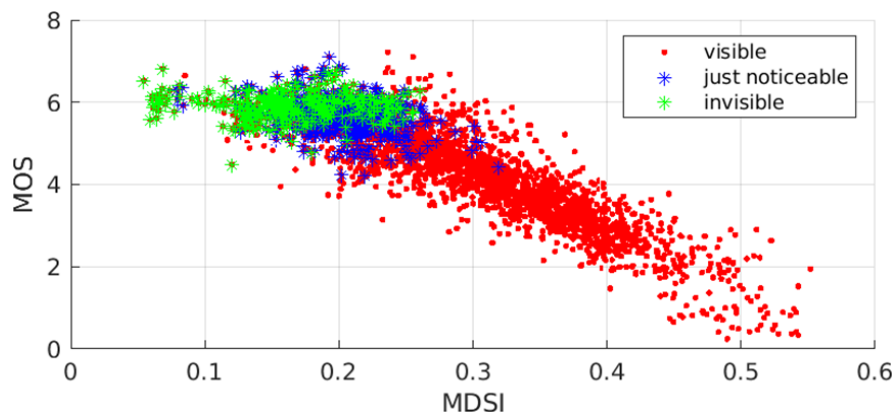


Figure 4.10 –The scatter plot of MDSI vs. MOS for the TID2013 dataset

Combining the statistical results and MOS values, it can be approximately stated that there are three gradations of image quality according to MDSI:

- excellent quality ($MDSI \leq 0,15$), the distortions are mostly invisible;
- good quality ($0,15 < MDSI \leq 0,25$), the distortions can be just noticeable;
- middle and bad quality ($MDSI > 0,25$), the distortions are visible or they can be annoying. Therefore, the reasonable range under interest is set as the range from 0,10 to 0,25. It is also worth noting that the relation between MDSI and MOS is almost linear and this should be considered as one more advantage of the MDSI metric.

4.2.2 The two-step method for lossy compression

First of all, a certain number of images are chosen to be compressed/decompressed assuming a series of CCP values, further referred to as the basic image set (sample images are shown in Fig 4.11).

The block diagram of the two-step compression method is illustrated in Fig4.12; in the first step, the initial CCP is determined using the desired visual quality and the average rate-distortion curve.

To improve the accuracy of provided visual quality, the CCP value needs to be corrected before the second step using the following equation:

$$CCP_{des} = CCP_{init} + \frac{M_{des} - M_{init}}{M'}, \quad (4.6)$$

where M_{des} is the desired visual quality preset by the user, and M_{init} denotes the quality calculated after the first step compression.

This corrected CCP value may be different for different images. Finally, the second step, compression, is carried out using the CCP_{des} , and the compressed image file obtained after the second step is considered as the final output with the desired quality.

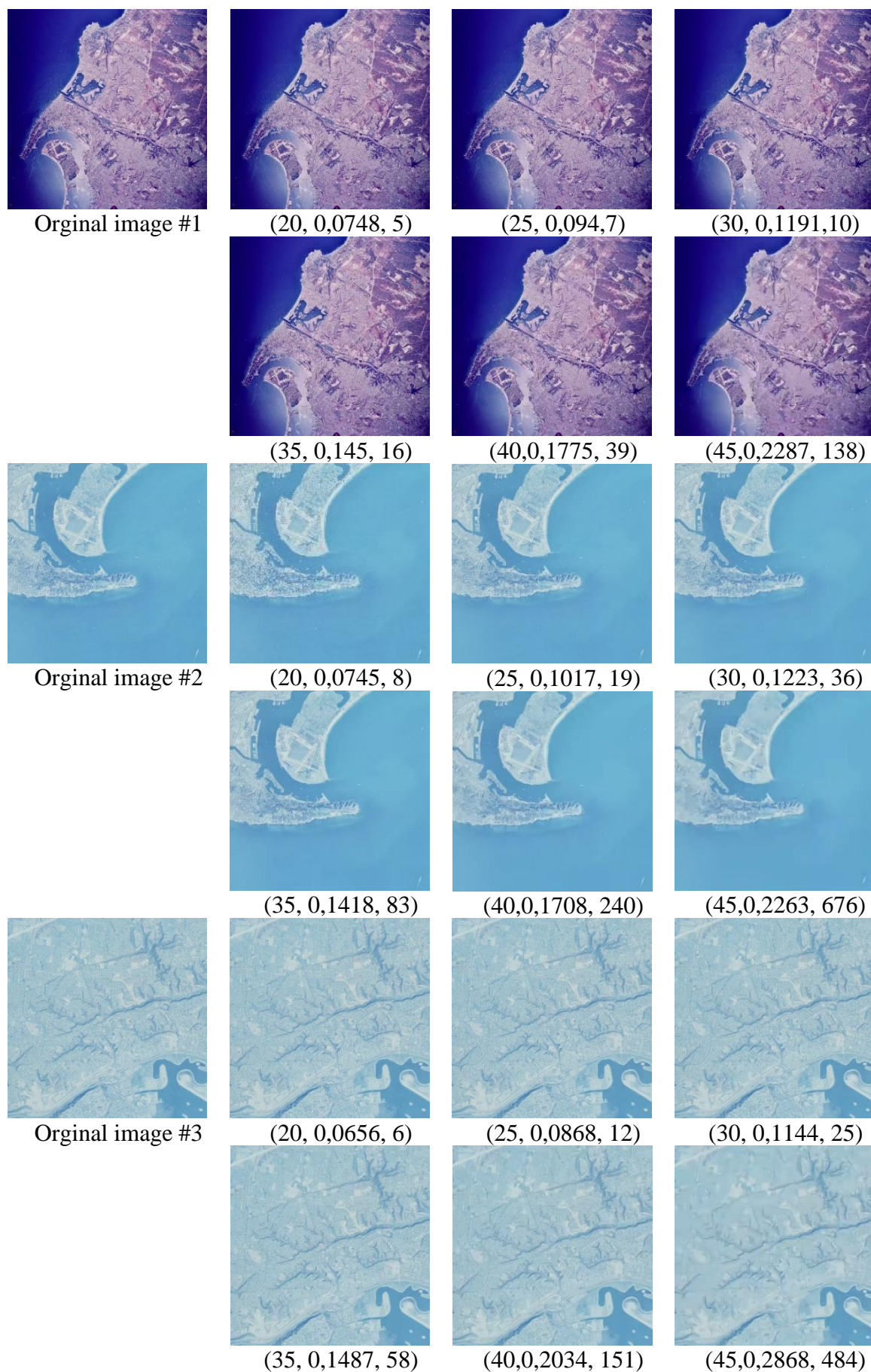


Figure. 4.11 -Samples of the basic RS images used in experiments. The values provided in brackets concern Q, MDSI, and CR values, respectively

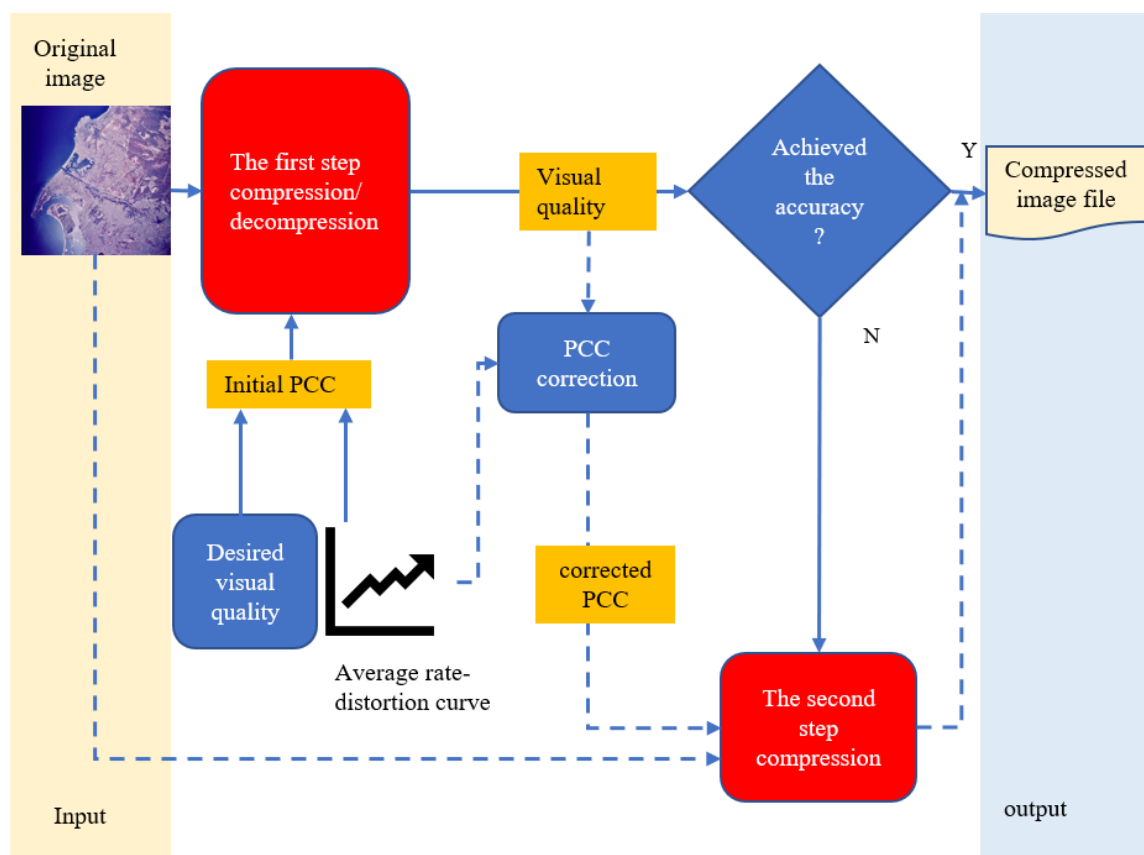


Figure. 4.12 - The block diagram of the two-step compression method

4.2.3 Implementation of the Two-Step Method for the BPG Coder

In this section, it is considered how the lossy compression coder is applied to compress three-channel RS images. The description of the implementation of the two-step compression method for the BPG coder is provided further.

1) The BPG coder on three-channel remote sensing images

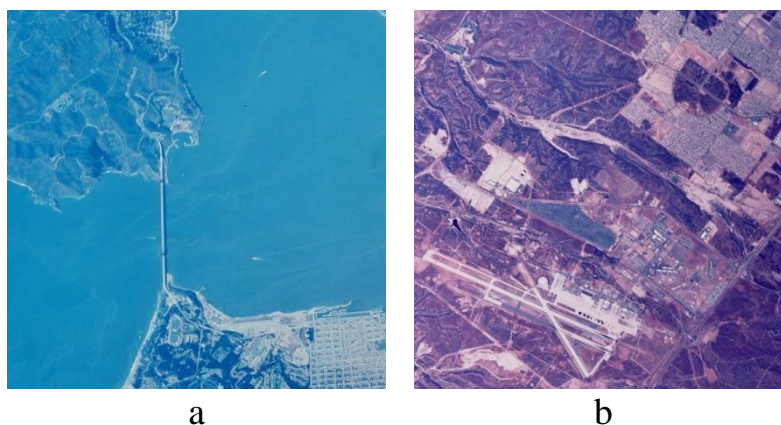


Figure 4.13 - Additional color test images from the USC-SIPI: (a) Frisco (image no. 2.1.03); (b) Diego (image no. 2.2.05).

A special comparison has been carried out for the considered case, i.e., for the MDSI metric and color RS images. Two 512×512 pixels color images for simple and complex structures (Frisco and Diego, respectively), shown in Figure 4.13, have been processed to produce three different values of MDSI: 0,1, 0,15, and 0,2.

The CR values have been obtained for the BPG (version 4:2:2), the 3D version of the AGU coder [210], JPEG, and JPEG2000 [124]. The data are collected in Table 4.6. Their analysis shows that the BPG slightly outperforms the 3D AGU and JPEG2000 for the middle values of the Q parameter (high quality of compressed images) and has obvious benefits for large Q values (low quality of the compressed image). Additional data on the comparison of coders can be found in [54, 207], showing that better performance of a coder usually leads to better classification.

Table 4.6 - Performance comparison of several coders: the CR values for two images and three MDSI values

Quality	Coder	CR	
		Frisco	Diego
MDSI=0.1	BPG	5,6	4,2
	3D AGU	5,2	4,0
	JPEG	4,9	2,7
	JPEG2000	3,8	3,2
MDSI=0.15	BPG	34,4	14,8
	3D AGU	21,2	9,9
	JPEG	20,1	8,7
	JPEG2000	10,2	6,5
MDSI=0.2	BPG	139,2	64,6
	3D AGU	60,4	22,0
	JPEG	38,9	21,8
	JPEG2000	37,9	15,2

2) Calculation of the Average Rate-Distortion Curve According to the MDSI

The prime task in the implementation of the two-step method for the BPG coder is to obtain the average rate-distortion curve for the considered visual quality metric (MDSI). Based on the analysis provided in Section 4.2.2, this average rate-distortion curve results from statistical data from a certain number of images. In this experiment, twelve 1024×1024 pixel three-channel RS images taken from the USC-SIPI dataset (volume 2: Aerials)

have been chosen as the basic image set, shown in Fig 4.5. Then, the compression and decompression have been carried out for each image with various CCP values, where all possible values of the Q parameter have been used. To obtain the average rate-distortion curve, MDSI values have been calculated for all 612 images (12 original images and 51 values of the Q parameter), and further averaged for the individual Q values. The obtained curves are presented in Fig 4.14 independently for each of the images used in the experiments. The obtained curves may be divided into four areas for further analysis:

- $Q \leq 3$ —the visual quality is constant and very high;
- $4 \leq Q \leq 25$ —the visual quality starts to change slowly, and the MDSI values for different images are very similar;
- $26 \leq Q \leq 40$ —the visual quality changes more significantly, but the deviations between quality indicators of different images are still small;
- $Q \geq 41$ —the visual quality starts to rapidly decline and the deviations of the MDSI values for different images become obvious.

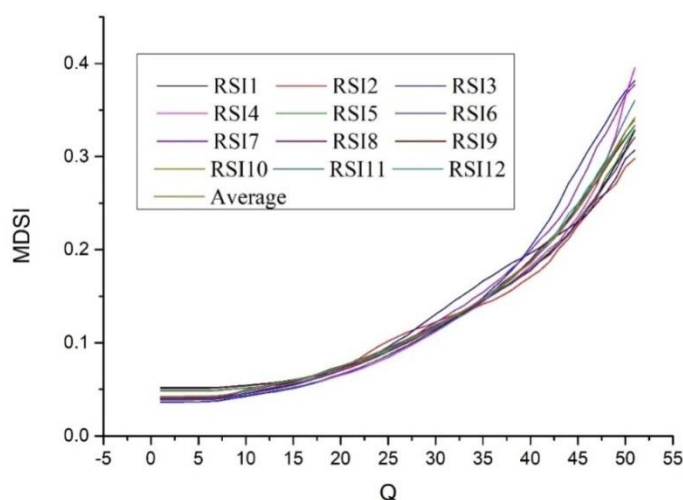


Figure 4.14 - Curves of the MDSI dependence on the Q parameter for individual basic images (RSI #1–#12) and the obtained average curve.

Consequently, the optimal work area of the Q parameter seems to be [26, 40], and the corresponding range of the MDSI values is [0,1, 0,25].

From Fig 4.14, we can also find that the difference in MDSI between images varies with the Q. Whilst the MDSI values are almost the same for small Q values (where the introduced distortions are anyway invisible), sufficient differences can be observed for large Q (e.g., data for Q equal to 40 and 45, which are of practical interest).

4.2.4 The experimental results

The two-step compression method has been applied firstly for the basic image set to provide four typical values for the MDSI metric, representing the four classes provided in Section 4.2.3. These four typical values have been set as 0,1, 0,15, 0,2, and 0,25, respectively, and the obtained statistical data are shown in Table 4.7, where M_{des} denotes the desired value of the MDSI metric, VAR_1 stands for the variance of MDSI provided in the first step, and VAR_2 is the variance of MDSI provided in the second step. For a better understanding of the data, the mean MDSI values finally provided in the second step are provided as well, denoted as $MEAN_2$.

Table 4.7 - The statistic data of basic images set for BPG coder (MDSI)

M_{des}	VAR_1	VAR_2	$MEAN_2$
0,1	$2,24 \times 10^{-5}$	$2,24 \times 10^{-6}$	0,1010
0,15	$4,31 \times 10^{-5}$	$6,73 \times 10^{-6}$	0,1495
0,2	$1,2 \times 10^{-4}$	$1,32 \times 10^{-5}$	0,1978
0,25	$3,32 \times 10^{-4}$	$1,85 \times 10^{-5}$	0,2466

The analysis of the data provided in Table 4.7 leads to the conclusion that the variance after the second step of compression has decreased by approximately one order of magnitude for each desired value. It proves that the proposed two-step procedure works well in the considered conditions. It can also be noticed that both variances VAR_1 and VAR_2 tend to increase if the desired MDSI increases. This means that the task of providing the desired MDSI is more important for larger M_{des} values, e.g., 0,2 or 0,25 (for $M_{des} = 0,1$, the distortions are invisible and they remain invisible if the desired MDSI is provided with the error of around 0,01). The mean absolute error of the desired quality, calculated as $|M_{des} - MEAN_2|$, does not exceed 0,034, and its value increases as the desired visual quality decreases, which is similar to the trend observed in previous works with the other coders [42, 45].

To verify the representativeness of the basic set, the other 12 RS images have been chosen as the test image set, shown in Fig 4.15 (further referred to as RSI #13–#24), which

is also a part of the USC-SIPI dataset. Then, the two-step compression method has been applied to these images to verify the correctness and universality of the previously obtained curve model, leading to the statistical data shown in Table 4.8.

Table 4.8 - The statistical data for the test image set for BPG coder (MDSI)

M_{des}	VAR_1	VAR_2	$MEAN_2$
0,1	$2,3 \times 10^{-5}$	$4,75 \times 10^{-6}$	0,1003
0,15	$5,48 \times 10^{-5}$	$6,33 \times 10^{-6}$	0,1490
0,2	$1,7 \times 10^{-4}$	$2,94 \times 10^{-6}$	0,1989
0,25	$3,01 \times 10^{-4}$	$2,04 \times 10^{-5}$	0,2465

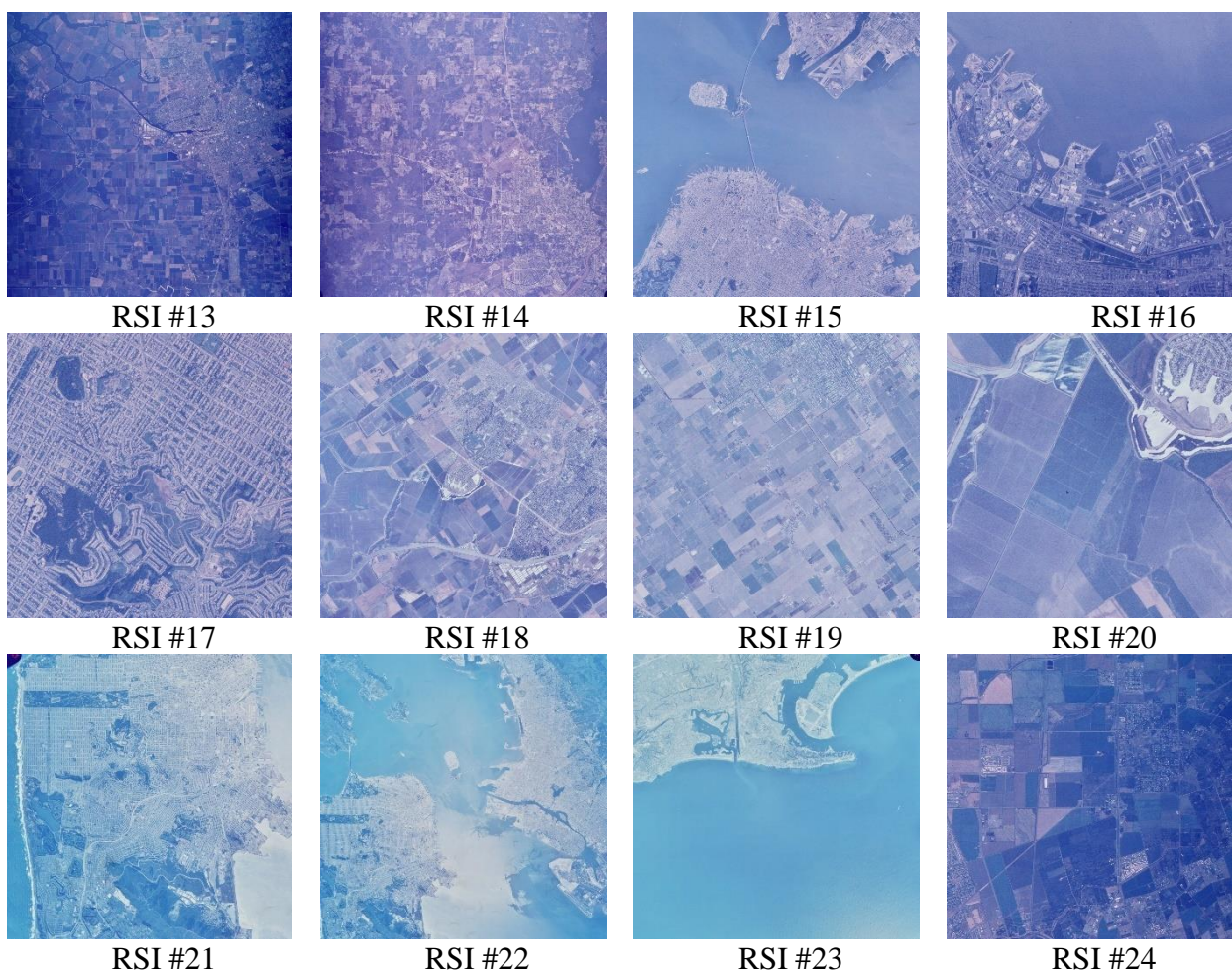


Figure 4.15 – The original images from the test remote sensing image set

As shown in Table 4.8, for each desired MDSI value, the variance after the second step of compression has also decreased by approximately an order of magnitude, and the mean error does not exceed 0,035. As one can see, the tendencies and values are similar to those observed in Table 4.7, so the basic image set has been chosen correctly to obtain the average rate-distortion curve and this model works well for other three-channel RS images.

To analyze the data for 12 test images (RSI #13–#24) in detail, the results obtained for $M_{des} = 0,25$ are presented in Table 4.9, where Q_{init} denotes the parameter Q used for the first step of compression. It comes from the average rate-distortion curve and equals 45 for all images. As may be seen in Table 4.9, although the initial MDSI values are different for individual test images, their variance after the second step has decreased significantly, being only 1/15 of its value after the first step. The mean value of the MDSI after the second step is also noticeably closer to 0,25; hence, its average relative error has also decreased (from 2,08% to 1,4%).

Table 4.9 - The statistical data for the test image set (the desired value of $MDSI=0,25$)

Test image	Q_{init}	$MDSI_{init}$	Q_{des}	$MDSI_{pro}$
Img# 13	45	0,2587	44	0,2427
Img# 14	45	0,2491	45	0,2491
Img# 15	45	0,2454	45	0,2454
Img# 16	45	0,2185	47	0,2447
Img# 17	45	0,2167	47	0,2394
Img# 18	45	0,2436	45	0,2436
Img# 19	45	0,2611	44	0,2445
Img# 20	45	0,2477	45	0,2477
Img# 21	45	0,2568	45	0,2568
Img# 22	45	0,2586	44	0,2443
Img# 23	45	0,2188	47	0,2507
Img# 24	45	0,2631	44	0,2489
VAR		$3,01 \times 10^{-4}$		$2,04 \times 10^{-5}$
MEAN		0,2448		0,2465

In general, the accuracy has radically improved due to the second step of compression. Meanwhile, there are cases when the MDSI after the second step is the same as for the first step, e.g., this happens for RSI #13. This means that there is no need to correct the Q parameter and apply the second step of compression in such cases. As shown in Table 4.8, 5 out of 12 images needed only one-step compression to meet the quality requirements, whilst the other seven images needed the second step to improve the accuracy. For all verification experiments (carried out for all 24 images and four desired MDSI values), 28,1% of images needed only the first step of the two-step compression to provide the desired visual quality.

4.2.5 Discussion

To analyze the accuracy of the provided visual quality for the BPG-based lossy compression of three-channel remote sensing images, three images (RSI #13, #14, and #16) have been selected as representative examples with the desired visual quality (MDSI) equal to 0,25. The decompressed images for the two-step compression method are shown in the middle (third) column in Fig 4.16. For the desired MDSI value equal to 0,25, the initial Q is equal to 45; the calculated Q_{des} values are different for different images (equal to 44, 45, and 47, respectively). The two images on the left from the third column are the images obtained if the parameter Q is set as $Q_{des} - 1$ and $Q_{des} - 2$, and the two images on the right are the images when Q is set as $Q_{des} + 1$ and $Q_{des} + 2$. For RSI #14, Q_{init} appears to be the appropriate value as for its change (increase or decrease), the error $\Delta MDSI = |MDSI_{final} - MDSI_{des}|$ increases. For RSI #13, the parameter Q is corrected to 44, and compared to the four other values, compression with the Q_{des} produces $MDSI_{final}$, which is the closest to the $MDSI_{des}$. In contrast, for RSI #16, the initial Q is corrected to 47, and compression with this Q_{des} produces the smallest error between $MDSI_{final}$ and $MDSI_{des}$. Concerning the error, for RSI #16 considered as an example, the provided $MDSI_{final}$ is 0,2447, and the error between $MDSI_{final}$ and $MDSI_{des}$ is 0,0053. Fig 4.16 shows five decompressed images resulting from RSI #16, where images compressed with two values of Q differing by unity seem to be practically identical, but if Q differs by 2 or more, e.g., for RSI #16 (45, 0,2185, 233) and RSI #16 (49, 0,2817, 527) compared to RSI #16 (47, 0,2447, 347), the difference is much easier to observe. Hence, if the difference $\Delta MDSI$ is approximately 0,015, it is difficult to observe the changes in the decompressed images. However, it becomes noticeable if $\Delta MDSI$ is approximately 0,03, e.g., for RSI #16 (46, 0,2303, 280) and RSI #16 (48, 0,2586, 424). Therefore, in practical applications, it is enough to ensure errors of providing the desired MDSI less than $\approx 0,01$. Consequently, it can be drawn that the accuracy of the two-step method for the BPG coder is good enough.

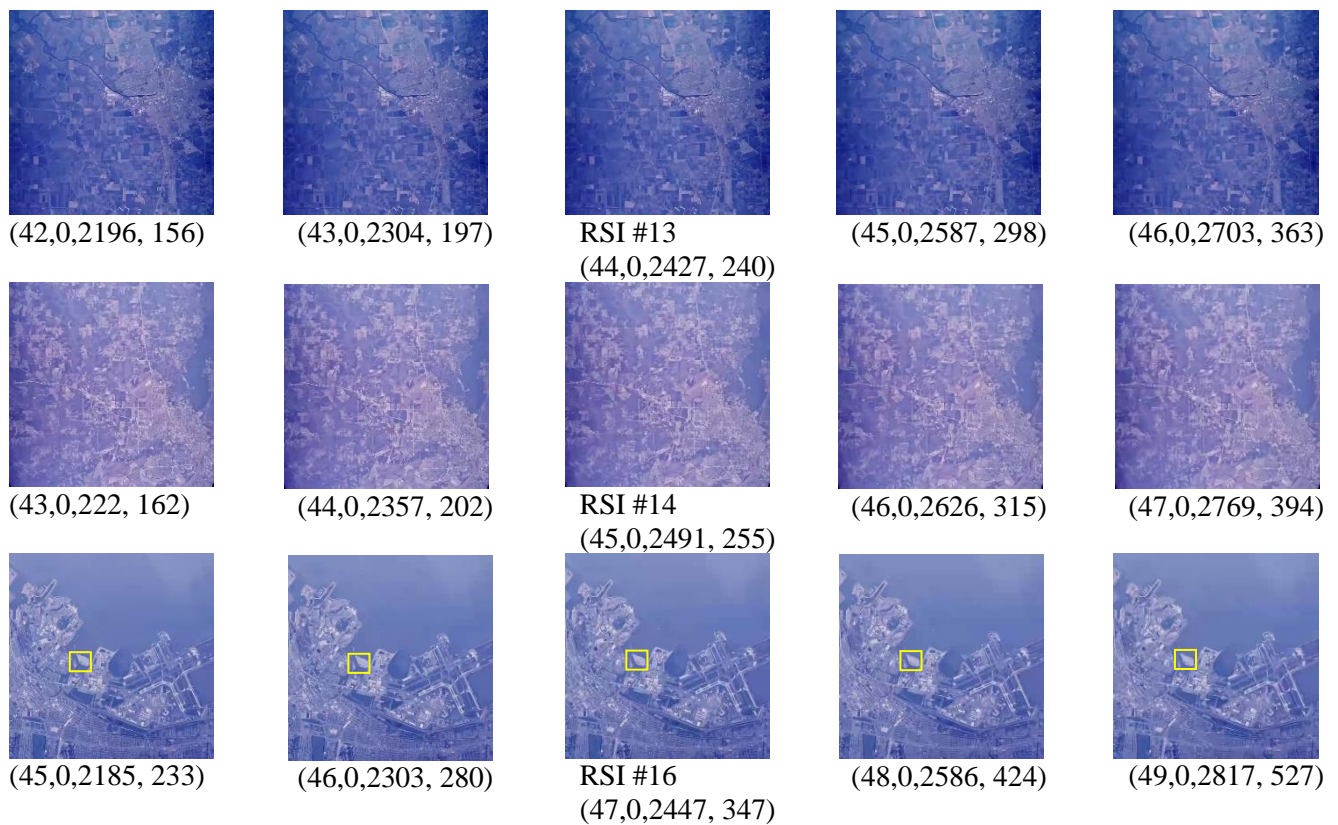


Figure 4.16 - The sample decompressed images obtained for the Q values around Q_{des} . The values provided in brackets concern Q , MDSI, and CR values, respectively

In summary, for images where the second step is necessary, regardless of whether the correction is forward or reverse (initial Q is increased or decreased), it gives a positive impact and eventually provides the visual quality that is the closest to the desired one. Additionally, the CR values provided for image lossy compression in the neighborhood of the distortion invisibility threshold are considerably higher than possible to achieve using a lossless compression.

To analyze the computational efficiency of the proposed approach, some tests have also been performed using a notebook with an Intel® Core™ i7-4710HQ CPU @2,50 GHz and 16,0 GB RAM, controlled by the 64-bit Windows 10 Pro operating system for the x64 processor architecture. For 512×512 pixels images, the compression time is from 0,02 s to 0,05 s depending on image complexity and the value of the parameter Q (a larger time is needed for more complex structure images). The decompression time is from 0,006 s to 0,019 s (more time is spent on the decompression of more complex structure images). For 1024×1024 pixels images, the compression time is from 0,06 s to 0,12 s; the decompression time is sufficiently smaller (from 0,02 s to 0,06 s). The MDSI values can

be calculated very quickly (the time for their calculation is only around 1,5 times longer than for the calculation of MSE).

4.3 Analysis of the effect of controlling visual quality on remote sensing classification

If lossy compression is applied to RS images, then distortions are introduced and this can lead to reduction of data classification accuracy [141, 211]. Thus, choosing a method of data compression and setting its parameters, one should keep in mind that lossy compression should not result in essential degradation of classification accuracy. It has been shown that if noise is present in RS images, lossy compression, under certain conditions, can even lead to improvement of classification accuracy due to specific noise filtering effect inherent for lossy compression. In turn, if RS images are practically noise-free (if their original quality is high), lossy compression results in decreasing the probability of correct classification.

4.3.1 Classification of compressed multichannel images and its improvement

In [8], it is shown that if compressed image quality is characterized by the metric PSNR-HVS-M [38] approximately equal to 43... 45 dB (this corresponds to invisible distortions) reduction of classification accuracy due to lossy compression is negligible. However, for smaller values of this visual quality metric, classification accuracy reduces. Meanwhile, the paper [211] presents a particular example and it is not clear how classification accuracy behaves when distortions are visible (PSNR-HVS-M is less than 40 dB). Besides, not a lot of image compression methods has been employed yet in analysis of classification accuracy. In addition, one can guess that influence of a classifier type can be essential as well as properties of classes (used features) might have impact.

Thus, the main goal of this Section is to analyze the influence of image compression on data classification accuracy using criteria typical just for classification [212]. Here, we consider three widely used classifiers, namely, support vector machine (SVM) [213], multilayer perceptron (MLP) [214, 215], and logistic regression (LGR) [216]. The

component-wise version of the coder AGU [108] is used for which a desired visual quality can be provided in different ways [81].

1)RS image component-wise compression

A multichannel RS image can be compressed in numerous ways[19, 211, 217, 218]. Different coders can be used [217, 219],it is possible to apply compression component-wise and in three-dimensional way [219], data can be pre-processed or not, etc. A choice of an optimal or appropriate compression method in practice depends upon many factors including priority of requirements to compression, restrictions, possible necessity to use standard tools, power consumption, etc.

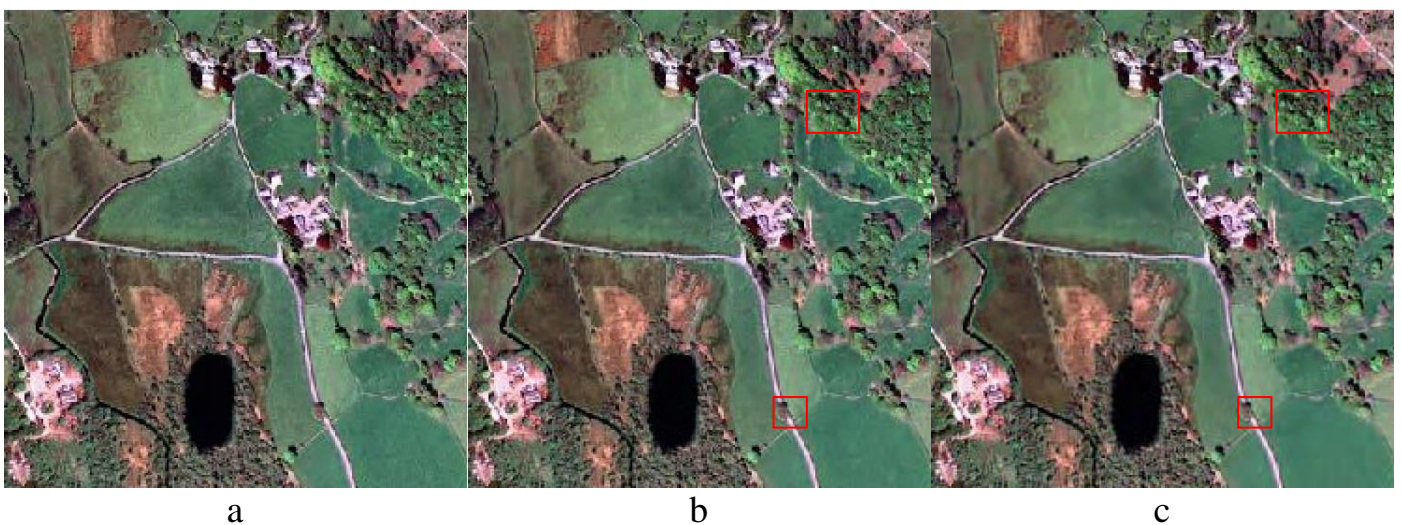


Figure 4.17 - Test image (Landsat TM image) before compression (a), after compression with providing PSNR-HVS-M equal to 36 dB (b), PSNR-HVS-M equal to 30 dB (c)

In our case, we have used component-wise compression using the coder AGU [108]. As a case study, we have used a three-channel image visualized in Fig. 4.17.a. This test image (represented as color image) has been acquired in three optical bands of Landsat TM imager where these bands relate to central wavelengths equal to $0,66 \mu\text{m}$, $0,56 \mu\text{m}$, and $0,49 \mu\text{m}$. These component images have been associated with R, G and B of synthesized color image of size 512×512 pixels. This image is “attractive” because of the following reasons. First, there are five visually distinguishable classes, namely, “Soil” (class 1), “Grass” (class 2), “Water” (class 3), “Urban (Roads and Buildings)” (class 4), and “Bushes” (class 5) where color features for many classes intersect [214], i.e. classification is not a simple task. Secondly, investigations of noise and other factors

influence on the classification accuracy were previously performed for this test image [214, 220].

Compression has been done in such a manner that PSNR-HVS-M values are practically the same for all components and there are six values of this metric considered in our experiments: 42, 39, 36, 33, and 30 dB. Figures 4.17.b and 4.17.c show compressed images with providing PSNR-HVS-M values equal to 36 and 30 dB, respectively.

Visual analysis of image in Fig. 4.17.b shows that there is a noticeable difference compared to the image in Fig. 4.17.a whilst difference is even larger for the image in Fig. 4.17.c (consider the fragments marked by red rectangles). Thus, we can study inter-connection between visual quality and classification results.

Table 4.10 - Compression data for three components of multichannel remote sensing image

Component	PSNR-HVS-M_{des}	QS	PSNR-HVS-M_{pro}	CR
Pseudo-red	45 dB	13,1221	44,996 dB	5,0373
	42 dB	17,3724	41,941 dB	6,0078
	39 dB	22,7182	38,943 dB	7,2387
	36 dB	29,8245	36,059 dB	8,9515
	33 dB	40,3305	32,993 dB	11,7343
	30 dB	55,5586	30,007 dB	16,5026
Pseudo-green	45 dB	12,246	45,6906 dB	4,8401
	42 dB	17,2849	41,9363 dB	5,9943
	39 dB	22,6586	38,9121 dB	7,2356
	36 dB	29,8559	35,9922 dB	8,9683
	33 dB	40,2208	33,0062 dB	11,7443
	30 dB	55,1339	29,9965 dB	16,4446
Pseudo-blue	45 dB	13,1779	44,8522 dB	5,0813
	42 dB	17,2265	41,847 dB	6,0157
	39 dB	22,5360	38,9052 dB	7,247
	36 dB	29,7293	35,9977 dB	8,9929
	33 dB	40,0144	32,9986 dB	11,7818
	30 dB	55,0136	30,0008 dB	16,5767

Besides, we would like to present some additional compression data. They are given in Table 4.10. We represent the data for three components of multichannel data, PSNR-HVS-M that is desired, QS that provides this PSNR-HVS-M and the obtained value. One can observe expected tendencies: since component images are highly correlated, QS and

CR values are practically the same for the same desired PSNR-HVS-M for all three components. The data also show that CR for PSNR-HVS-M is equal to 30 dB is about three times larger when PSNR-HVS-M is equal to 45 dB. Thus, there can be an essential gain in CR if reduction of probability of correct classification on QS (and, therefore, CR) is quite slow.

2) Image classification results

The proposed procedure for classifying multichannel images includes two stages. At the first stage, pixel-by-pixel classification is carried out with training using elementary classifiers.

For the formation of training samples (fragments of the initial noisy image containing objects of only one class), color masks were used, constructed by highlighting clearly distinguishable, homogeneous areas representing each class of objects of natural and anthropogenic origin [1, 4]: 1 (red) – "soil", 2 (green) – "grass", 3 (blue) – "water", 4 (yellow) – "urban", 5 (cyan) – "bushes"; classes 3 and 4 contained both areal and linear objects: "water" = [lake, river], "urban" = [buildings, roads]. The size (expressed by the number of pixels) of the training set and verification set is: class - [training set], [verification set]; soil - [1610], [12524]; grass - [1365], [25146]; water - [373], [4986]; urban - [188], [6230]; bushes - [720], [8692]; total - [4256], [57578]. An F-measure was used as a criterion for recognition quality. This is a harmonic mean of accuracy and completeness (accuracy shows how many of the objects identified by the classifier as positive are indeed positive; completeness shows how many of the positive objects were identified by the classifier). The harmonic mean has an important property –it is close to zero if at least one of the arguments is close to zero. In the multi-class case, this is the average of the F-measure of each class with weighting depending on the average parameter. The results of class recognition of the original (uncompressed) image and the image after compression with providing PSNR-HVS-M equal to 30 dB are shown in Fig. 4.18. Estimates of probabilities of correct classification for the multichannel image are given in Table 4.11.

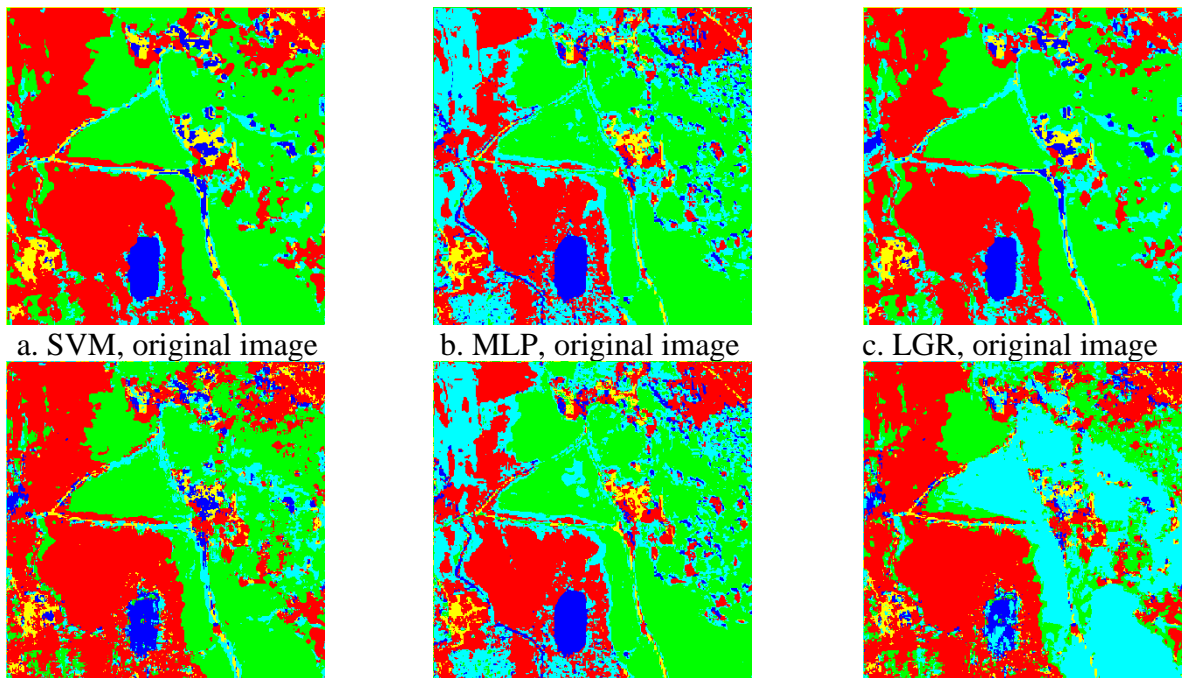
Analysis of the classification results of compressed images shows that an increase in the compression ratio leads to a deterioration in the reliability of objects recognition.

Moreover, compression had the greatest impact on the probability of recognizing such classes as 4 - “urban” and 5 - “bushes”. Wherein the logistic regression classifier turned out to be the most sensitive. A decrease in PNSR-HVS-M to 30 dB leads to a decrease in quality for these classes from 0,67 to 0,316 and from 0,42 to 0,042, respectively. Compression had less impact on the recognition results of classes 1 - “soil” and 2 - “grass”. For different classifiers, the decrease in quality was 4% and 2,5 - 41%, respectively. The best results for classifying objects in a compressed image were shown by the MLP. The overall recognition probability decreased by only 2,3% compared to the original image, while for LGR this decrease was 41,7%.

Table 4.11 estimates of correct recognition probability for objects in the images

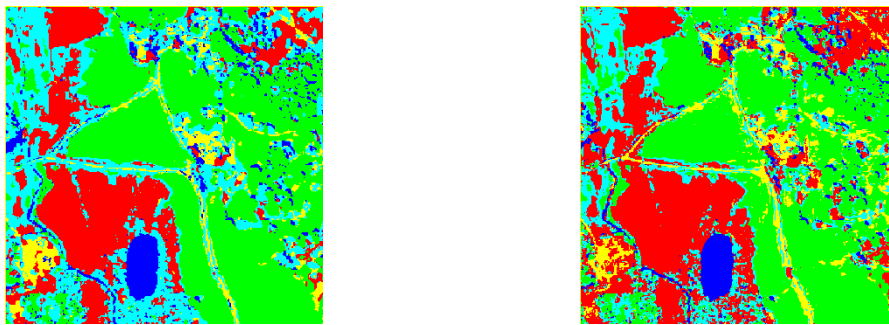
Classifier	Class					Overall probability
	<i>Soil</i>	<i>Grass</i>	<i>Water</i>	<i>Urban</i>	<i>Bushes</i>	
Original image						
SVM	0,73	0,99	0,84	0,71	0,25	0,8105
LGR	0,76	0,99	0,85	0,67	0,42	0,8261
MLP	0,78	0,96	0,97	0,60	0,54	0,8345
PSNR-HVS-M=36 dB						
SVM	0,73	0,973	0,84	0	0,386	0,773
LGR	0,76	0,98	0,83	0,531	0,376	0,821
MLP	0,78	0,97	0,936	0,58	0,46	0,813
PSNR-HVS-M=30 dB						
SVM	0,73	0,965	0,76	0,396	0,15	0,761
LGR	0,73	0,406	0,527	0,316	0,042	0,481
MLP	0,78	0,94	0,97	0,448	0,382	0,815

At the second stage of the procedure, a multilayer classification (analysis "in depth") is carried out based on the recognition results by the considered classifiers. Each of the matrices of recognition results for elementary classifiers can be considered as a classification layer. Thus, for each pixel we get a vector of ambiguous decisions and select a specific decision. It is necessary to determine the relationship between different combinations of the classifiers' outcomes, and the set of classes. Then, the decision is made in favor of the class for which the resulting combination of outcomes is the most likely. The results $Y\Sigma$, obtained by joint statistical processing of SVM, LGR, MLP outcomes are presented in Fig. 4.19 and in Tables 4.12 – 4.14.



e. SVM, PSNR-HVS-M=30dB f. MLP, PSNR-HVS-M=30dB g. LGR, PSNR-HVS-M=30dB

Figure 4.18 – Class recognition results



a. layers convolution, original images b. logical convolution, PSNR-HVS-M=30 dB

Figure 4.19 – Results of the classification

Table 4.12 - Confusion matrices for original image

Decisions	Classes					P_{corr}
	<i>Soil</i>	<i>Grass</i>	<i>Water</i>	<i>Urban</i>	<i>Bushes</i>	
Soil	0,978	0	0	0	0,022	0,909
Grass	5,97E-04	0,998	0	1,19E-04	1,35E-03	
Water	1,00E-03	0	0,989	0	0,01	
Urban	0,163	3,21E-04	0,018	0,713	0,106	
Bushes	0,293	0,013	0,04	5,18E-03	0,649	

Table 4.13 Confusion matrices for PSNR-HVS-M=36dB

Decisions	Classes					P_{corr}
	<i>Soil</i>	<i>Grass</i>	<i>Water</i>	<i>Urban</i>	<i>Bushes</i>	
Soil	0,979	0	0	0	0,021	0,873
Grass	0,006	0,988	0	0,002	0,004	
Water	0,005	0	0,993	0	0,002	
Urban	0,19	0	0	0,625	0,185	
Bushes	0,455	0,016	0,039	0,003	0,487	

Table 4.14 Confusion matrices for PSNR-HVS-M=30dB

Decisions	Classes					P_{corr}
	<i>Soil</i>	<i>Grass</i>	<i>Water</i>	<i>Urban</i>	<i>Bushes</i>	
Soil	0,979	0	0	0	0,021	0,853
Grass	0,001	0,978	0	0,018	0,003	
Water	0,007	0	0,991	0	0,001	
Urban	0,289	0,001	0	0,683	0,026	
Bushes	0,552	0,05	0,034	0,017	0,347	

4.3.2 Lossy compression of multichannel remote sensing images with quality control

In this Section, a dependence between classification accuracy of maximum likelihood and neural network classifiers applied to three-channel test and real-life images and quality of compressed images characterized by standard and visual quality metrics is studied.

We consider the impact of lossy compression on multichannel RS image classification [103, 125, 142, 206, 211]. Specifically, we focus on the case of a limited number of channels, e.g., color, multi-polarization, or multispectral images, due to the following reasons: (a) it is simpler to demonstrate the effects that take place in images due to lossy compression and how these effects influence classification just for the case of a small number of image components; (b) an accurate classification of multichannel images with a small number of components is usually a more difficult task than the classification of hyper spectral data (images with a relatively large number of components) because of a limited number of available features and the necessity to reliably discriminate classes in feature space.

As we need some metrics that should be good enough and applicable to single component images, we apply the metrics PSNR-HVS and PSNR-HVS-M.

1) Considered Approaches to Multichannel Image Classification

In this subsection, we consider two of them. The first one is based on the maximum likelihood method (MLM) [221, 222] and the second one relies on a neural network (NN) training [221]. In both cases, the pixel-wise classification are studied. There are many reasons behind using the pixel-wise approach and just these classifiers: (a) to simplify the classification task and to use only Q features ($q = 1, \dots, Q$), i.e., the values of a given multichannel image in each pixel; (b) to show the problems of pixel-wise classification; (c)

MLM and NN based classifiers are considered to be among the best ones [221]. The note is that classifiers of RS data classification can be trained in different ways.

A. Maximum Likelihood Classifier

Maximum likelihood (ML) classification is based on several main principles. First, it is assumed that there is some feature space. For three-channel images, these features can be just pixel values and/or their ratios. It is also possible to exploit some texture features derived from a group of neighbor pixels. However, in this section, we consider the simplest version of the ML classification which is pixel-wise and employs only pixel values. Second, classification is based on preliminary training. At this stage, sample data for chosen classes are obtained and distributions for them are estimated or modeled. Decision rule is chosen. Then, the classifier is applied to RS images to be classified.

B. NN Classifier Description

We have chosen a simple but efficient NN for our application that can be easily implemented or placed on different platforms and devices. The employed classifier is an all-to-all connected perception combined with a self-organizing map to treat the obtained weights or probabilities of the pixel belonging to one of the classes.

For NN training, we have taken 70% of the produced set, the other 30% of data is used for validation. The chosen architecture was NN with one hidden layer with a fully optimal number of training epochs equal to 50 for the given NN. The proposed classifier is easy to use, and it is fast. The overall training process was repeated 100 times with full permutation of the dataset and the obtained classifier has been applied to test images.

2) Analysis of real-life three-channel images

Our experiments with real-life images have been done using a three-channel Landsat TM image earlier used in our studies [53, 223]. This test image (shown in pseudo-color representation in Figure 4.20.a). The image fragments used in classifier training are shown in Figure 4.20.b whilst the pixel used for verification of classifiers are marked by the same colors in Figure 4.20.c. Details concerning the numbers of pixels are given in [53]. Class 1 is marked by red color, Class 2—by green, Class 3—by dark blue; Class 4—by yellow, Class 5—by azure.

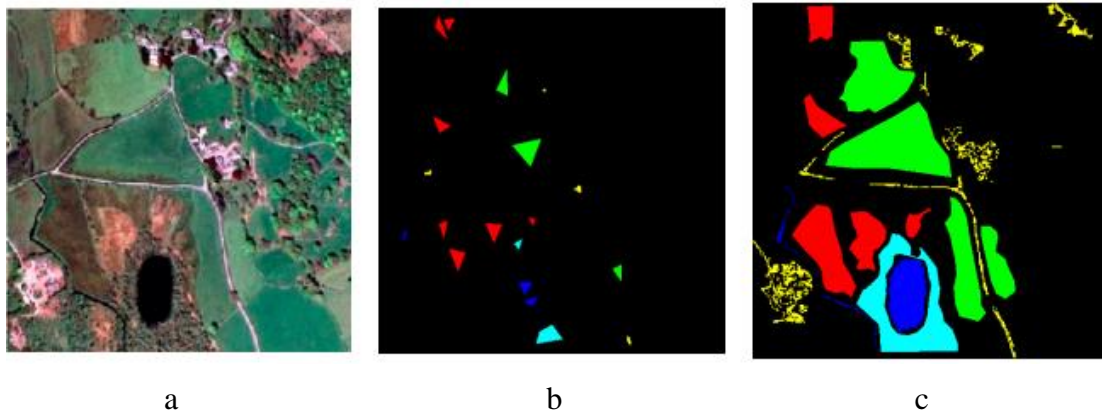


Figure 4.20 - The three-channel image, a) pseudo-color representation; b) pixel groups used for training; c) pixel groups used for verification.

A. MLM Classifier Results

Let us start by considering the results of applying the MLM classifier trained for the original image. Tables 4.16–4.19 present the obtained data in the form of confusion matrices. As it is seen, probabilities of correct classification for particular classes vary from 0,75 to 0,99 and the smallest probabilities take place for rather heterogeneous classes “Soil” and “Bushes”. If the same classifier is applied to a compressed image (PSNR-HVS-M=42 dB), the results are slightly worse (see data in Table 4.20). Reduction of probabilities of correct classification by 0,002 . . . 0,033 takes place. The largest reduction is observed for the most heterogeneous class “Bushes”.

Consider now the data obtained if the MLM classifier is applied to the image compressed with visible distortions (PSNR-HVS-M = 36 dB). The confusion matrix is given in Table 4.21. Most probabilities of correct classification for particular classes continue to decrease although this reduction is not essential—up to 0,021 compared to the previous case (Table 4.20). Finally, let us analyze data for a compressed image (PSNR-HVS-M = 30 dB) to which the MLM classifier has been applied. The data are presented in Table 4.22. The tendency is the same—most probabilities of correct classification for particular classes continue to decrease but, again, the reduction is not large. Classification maps for three images are presented in Fig 4.22. They are quite similar although some differences can be found. Compression does not lead to radical degradation of classification accuracy.

Table 4.19 - Classification probabilities for the MLM method trained for the original image and applied to it

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,747	0,027	0	0,037	0,189
Grass	0,173	0,812	0	0,008	0,007
Water	0	0	0,967	0	0,033
Urban	0,004	0	0	0,989	0,007
Bushes	0,091	0,068	0,013	0,016	0,812

Table 4.20 - Classification probabilities for the MLM method trained for the original image and applied to the compressed image (PSNR-HVS-M=42dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,732	0,025	0	0,043	0,199
Grass	0,177	0,808	0	0,008	0,006
Water	0	0	0,963	0	0,037
Urban	0,005	0	0	0,987	0,008
Bushes	0,118	0,07	0,014	0,018	0,779

Table 4.21- Classification probabilities for the MLM method trained for the original image and applied to the compressed image (PSNR-HVS-M=36dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,744	0,028	0	0,043	0,185
Grass	0,185	0,799	0	0,009	0,007
Water	0	0	0,959	0	0,041
Urban	0,005	0	0	0,985	0,010
Bushes	0,133	0,074	0,016	0,018	0,758

Table 4.22 - Classification probabilities for the MLM method trained for the original image and applied to the compressed image (PSNR-HVS-M=30dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,724	0,026	-	0,044	0,206
Grass	0,178	0,809	0	0,006	0,007
Water	-	0	0,942	0	0,058
Urban	0,007	-	0	0,982	0,011
Bushes	0,145	0,079	0,016	0,019	0,741

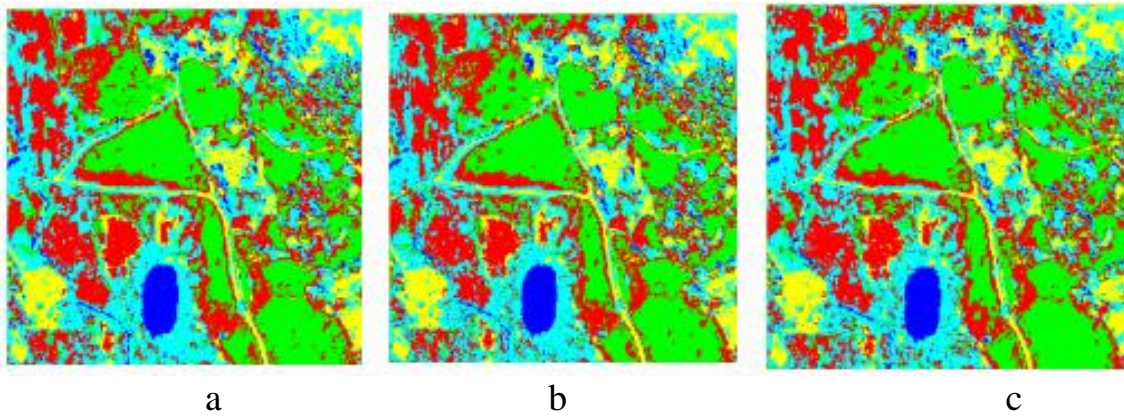


Figure 4.22 - Classification results: a) for original image, b) for compressed image (PSNR-HVS-M = 36 dB), c) for compressed (PSNR-HVS-M = 30 dB)

For original image $P_{cc} = 0,865$; for compressed images it equals to 0,853, 0,854, 0,853, 0,849, 0,842, and 0,839 for PSNR-HVS-M equal to 45, 42, 39, 36, 33, and 30 dB, respectively. So, it is possible to state that lossy compression providing PSNR-HVS-M about 40 dB does not lead to sufficient reduction of P_{cc} for the considered case. Moreover, if training is done for the image compressed with the same conditions as an image subject to classification, classification results can improve. For example, if training has been done for the compressed image (PSNR-HVS-M = 36 dB) and then applied to this image (to verification set of pixels), P_{cc} increases to 0,855. We have analyzed distributions of features before and after compression. One reason why P_{cc} does not radically reduce with CR increase is that the corresponding distributions do not differ a lot. The largest differences are observed for the classes “Soil” and “Bushes”. It might be slightly surprising that the class “Urban” is recognized so well. The reason is that, for the considered image, features for this class do not sufficiently overlap with features for other classes.

B. NN Classifier Results

The ideal case data (NN trained for original image is applied to the original image) are given in Table 4.23. The results can be compared to the data in Table 4.19. The NN classifier better recognizes the classes “Soil” and “Grass”, the results for the class “Water” are approximately the same.

Table 4.23 - Classification probabilities for the NN-based method trained for the original image and applied to it

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,950	0	0	0	0,05
Grass	0	1	0	0	0
Water	0	0	0,952	0	0,048
Urban	0,098	0	0	0,767	0,135
Bushes	0,187	0,09	0,02	0,009	0,774

Table 4.24 - Classification probabilities for the NN-based method trained for the original image and applied to the compressed image (PSNR-HVS-M = 42 dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,942	0	0	0	0,058
Grass	0	0,999	0	0	0,001
Water	0	0	0,942	0	0,058
Urban	0,01	0	0	0,765	0,136
Bushes	0,207	0,016	0,018	0,009	0,751

Table 4.25 - Classification probabilities for the NN-based method trained for original image and applied to the compressed image (PSNR-HVS-M = 36 dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,941	0	0	0	0,057
Grass	0	0,999	0	0	0
Water	0,001	0	0,919	0	0,059
Urban	0,099	0	0	0,747	0,131
Bushes	0,247	0,028	0,014	0,004	0,718

Table 4.26 - Classification probabilities for the NN-based method trained for the original image and applied to the compressed image (PSNR-HVS-M = 30 dB)

Class	Probability of decision				
	Soil	Grass	Water	Urban	Bushes
Soil	0,941	0	0	0	0,059
Grass	0	0,999	0	0	0,001
Water	0,001	0	0,919	0	0,008
Urban	0,099	0	0	0,747	0,153
Bushes	0,247	0,028	0,014	0,004	0,707

Suppose now that this classifier (NN trained for the original image) is applied to compressed images. The results for different qualities of compressed data are presented in Tables 4.24–26. If PSNR-HVS-M is equal to 42 dB or 36 dB, the results keep practically the same. Only the probability of correct classification for “Bushes” steadily decreases. Reduction of classification accuracy occurs to be larger for the image compressed with PSNR-HVS-M equal to 30 dB. Mainly, reduction takes place for the classes “Water” and “Urban”.

Some classification maps are presented in Figure 4.23. They do not differ a lot from each other. Some pixels that belong to the class “Water” for the narrow river (see the left low corner in Figure 4.23.c) “disappear” (become misclassified). This is because of the effects of smearing the prolonged objects due to lossy compression.

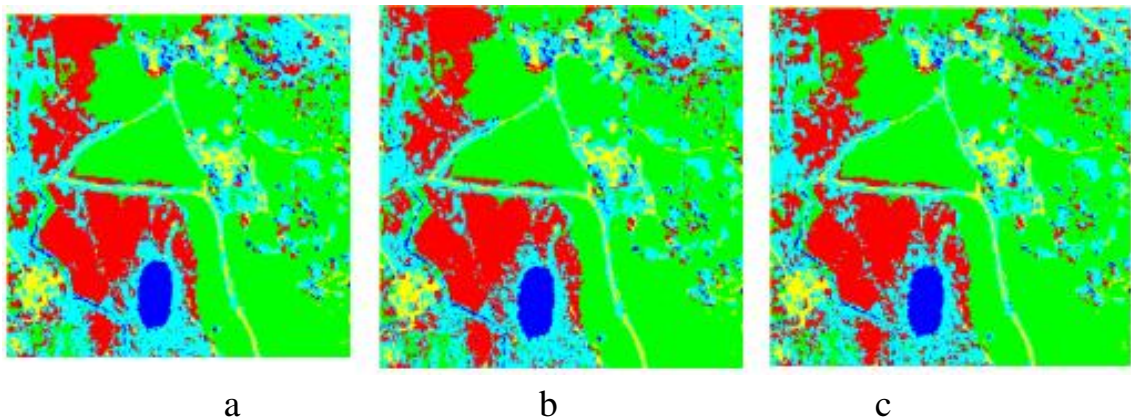


Figure 4.23 - Classification results provided by NN-based method: a) for original image, b) for compressed image (PSNR-HVS-M = 36 dB), c) for compressed (PSNR-HVS-M=30dB)

The MLM classifier training for compressed images was performed using training samples for fragments shown in Figure 4.23.b. Due to the usage of the same data for both training and validation, noticeable classification improvement can be observed (at least, for several classes).

4) Brief Analysis for Sentinel-2 Three-Channel Images

Concerning other images or their fragments, other imagers, and other compression techniques, we have taken three-channel images of the Kharkiv region (Ukraine) from Sentinel-2, in visible range acquired on 30 August 2019 when there were practically no clouds (images are available at <https://apps.sentinel-hub.com/eo-browser/?lat=46.45&lng=34.12&zoom=6&time=2019-11->

03&preset=3_NDVI&datasource=Sentinel-2%20L1C). The analyzed 512×512 pixel fragments are for the neighborhood of StaryiSaltiv (45 km north-east from Kharkiv, Ukraine, set 1) and north part of Kharkiv (set 2)—see Fig.4.24. The main reason for choosing these fragments is the availability of ground truth data that allow easy marking of four typical classes: Urban, Water, Vegetation, and Bare Soil. One more reason is that the image in Fig. 4.24.a is considerably less complex (textural) than the image in Fig 4.24.b.



Figure 4.24 - Fragments of Sentinel-2 images of a)StaryiSaltiv, b) north part of Kharkiv

First, these fragments have been compressed component-wise with providing a set of PSNR-HVS-M values using three coders: AGU, SPIHT [5], and Advanced DCT coder (ADCTC)[119]. Basic data on compression performance are given in Table 4.27. The observed dependences are predictable. CR for all coders increases if desired PSNR-HVS-M reduces. CR values for different components for the same desired PSNR-HVS-M and set differ but not considerably. CR for AGU is usually slightly larger than for SPIHT (for the same conditions), ADCTC outperforms both coders. CR values for Set 2 images are several times smaller than for the corresponding Set 1 images due to the higher complexity of Set 2 RS data.

Second, the classifier training has been performed and probabilities of correct classification for all four classes as well as the total probability of correct classification have been obtained. For the NN classifier and AGU coder, the data are presented in Tables 4.28 and 4.29. Here and below the abbreviation SS1 means that classification has been done for an original (uncompressed) image 1 whilst, e.g., SS1_45 means that classification

has been done for the image compressed with providing PSNR-HVS-M = 45 dB. Training has been done for the original image. As one can see, compression results for the Set 1 image practically do not depend on compressed image quality. Even for PSNR-HVS-M = 30 dB the probabilities of correct classification are practically the same as for original data. Meanwhile, for the Set 2 image, the situation is another. There is a tendency for classification accuracy degradation if compressed image quality becomes worse. This is mainly due to the reduction of correct classification probabilities for more heterogeneous classes, e.g., vegetation. For this image, it is possible to recommend compression with providing PSNR-HVS-M about 42 dB to avoid considerable reduction of classification accuracy.

Table 4.27 - CR comparison for real-life data for three coders, both sets.

Color	PSNR-HVS- M _{des}	CR		
		AGU	SPIHT	ADCTC
Pseudo-Red, set 1	45,000	10,338	9,467	11,306
	39,000	18,383	16,631	20,362
	33,000	32,929	29,303	36,465
Pseudo-Red, set 2	45,000	3,317	3,283	3,498
	39,000	4,520	4,762	4,886
	33,000	7,121	7,360	7,986
Pseudo-Green, set 1	45,000	7,179	5,957	7,513
	39,000	12,666	10,914	13,513
	33,000	23,907	20,833	26,030
Pseudo-Green, set 2	45,000	3,214	3,150	3,363
	39,000	4,356	4,564	4,684
	33,000	6,784	7,105	7,573
Pseudo-Green, set1	45,000	6,665	5,727	7,216
	39,000	11,942	10,336	12,740
	33,000	24,914	22,346	27,191
Pseudo-Green, set1	45,000	3,212	3,252	3,418
	39,000	4,314	4,698	4,785
	33,000	6,882	7,380	7,817

We have also analyzed the possibility of using compressed images for training. The probability of correct classification has improved by about 0.01 for Set 1 image and remained practically the same for Set 2 image. Another part of our study relates to classification by MLM (the same maps have been used for training the NN and MLM). MLM has been applied to data compressed by three aforementioned coders. For the AGU coder, the obtained data are presented in Tables 4.30 and 31.

Table 4.28- Probabilities of correct classification depending on compressed image quality for Set 1, AGU coder, NN classifier

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0,774	0,777	0,775	0,777	0,778	0,773	0,786
Water	0,998	0,998	0,998	0,999	0,999	0,999	0,998
Vegetation	0,915	0,913	0,914	0,915	0,918	0,921	0,927
Bare soil	0,809	0,811	0,809	0,809	0,812	0,798	0,801
P _{total}	0,874	0,875	0,874	0,875	0,877	0,873	0,878

Table 4.29 - Probabilities of correct classification depending on compressed image quality for Set 2, AGU coder, NN classifier

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0,877	0,872	0,871	0,865	0,861	0,861	0,864
Water	0,654	0,659	0,662	0,661	0,654	0,644	0,647
Vegetation	0,857	0,825	0,815	0,801	0,788	0,797	0,788
Bare soil	0,890	0,867	0,865	0,869	0,866	0,860	0,870
P _{total}	0,820	0,810	0,803	0,799	0,792	0,791	0,792

Table 4.30. Probabilities of correct classification depending on compressed image quality for Set 1, AGU coder, ML classifier

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0,695	0,727	0,727	0,735	0,745	0,743	0,771
Water	0,979	0,99	0,992	0,992	0,993	0,992	0,988
Vegetation	0,9	0,896	0,896	0,895	0,899	0,901	0,902
Bare soil	0,85	0,852	0,851	0,849	0,852	0,831	0,826
P _{total}	0,856	0,866	0,866	0,868	0,892	0,867	0,872

Table 4.31. Probabilities of correct classification depending on compressed image quality for Set 1, AGU coder, ML classifier

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0,917	0,914	0,913	0,913	0,91	0,913	0,918
Water	0,923	0,902	0,893	0,885	0,873	0,873	0,857
Vegetation	0,691	0,667	0,647	0,645	0,634	0,631	0,619
Bare soil	0,856	0,819	0,815	0,817	0,812	0,802	0,812
P _{total}	0,847	0,826	0,817	0,815	0,807	0,805	0,801

Table 4.32 - Probabilities of correct classification depending on compressed image quality for Set 1, SPIHT coder, ML classifier

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0,895	0,721	0,724	0,728	0,735	0,742	0,759
Water	0,979	0,991	0,991	0,993	0,993	0,992	0,985
Vegetation	0,9	0,899	0,901	0,9	0,904	0,908	0,906
Bare soil	0,85	0,856	0,864	0,858	0,858	0,848	0,816
P _{total}	0,856	0,866	0,87	0,87	0,872	0,873	0,867

Table 4.33. Probabilities of correct classification depending on compressed image quality for Set 2, SPIHT coder, ML classifier

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0,917	0,915	0,915	0,917	0,916	0,92	0,924
Water	0,923	0,918	0,913	0,903	0,892	0,884	0,875
Vegetation	0,691	0,686	0,688	0,692	0,697	0,699	0,707
Bare soil	0,856	0,847	0,858	0,853	0,853	0,861	0,859
P _{total}	0,847	0,841	0,843	0,839	0,839	0,841	0,841

Table 4.34 - Probabilities of correct classification depending on compressed image quality for Set 1, ADCTC coder, ML classifier

Classes	SS1	SS1_45	SS1_42	SS1_39	SS1_36	SS1_33	SS1_30
Urban	0,695	0,723	0,725	0,734	0,736	0,751	0,76
Water	0,979	0,992	0,994	0,995	0,995	0,995	0,995
Vegetation	0,9	0,898	0,9	0,901	0,901	0,91	0,913
Bare soil	0,85	0,863	0,855	0,85	0,864	0,84	0,827
P _{total}	0,856	0,869	0,869	0,87	0,874	0,874	0,874

Table 4.35. Probabilities of correct classification depending on compressed image quality for Set 2, ADCTC coder, ML classifier

Classes	SS2	SS2_45	SS2_42	SS2_39	SS2_36	SS2_33	SS2_30
Urban	0,917	0,914	0,911	0,91	0,911	0,912	0,915
Water	0,923	0,918	0,912	0,906	0,909	0,901	0,898
Vegetation	0,691	0,647	0,688	0,682	0,664	0,67	0,693
Bare soil	0,856	0,848	0,848	0,847	0,843	0,859	0,868
P _{total}	0,847	0,838	0,84	0,836	0,832	0,836	0,844

It is seen that lossy compression leads to a positive effect for the Set 1 image. For all classes except Bare Soil, the probabilities of correct classification improve or remain the same. Total probability also improves and remains at approximately the same level for all considered qualities of the compressed image (Table 4.30). Meanwhile, for the Set 2 image, larger CR leads to a steady reduction of total probability and decreasing of probabilities for most classes (Table 4.31). We have also checked whether or not it is worth carrying out training for compressed images instead of uncompressed ones. The answer, as earlier, is yes. In particular, for images compressed with providing PSNR-HVS-M = 42 dB. The total probabilities equal to 0,893 for the Set 1 data and 0,858 for the Set 2 image. Thus, it is worth using those data in training that have been obtained with the same compression conditions as images subject to classification. Data obtained for the SPIHT coder are presented in Tables 4.32 and 4.33. Their analysis shows the following. Again, lossy compression has a small impact on the classification of the Set 1 data. An optimum is observed for PSNR-HVS-M about 40 dB. The impact of compression for the Set 2 data is also small. For some classes, probabilities of correct classification improve, for others become slightly worse. In aggregate, the total probability remains almost the same.

Consider now the data obtained for ADCTC. They are presented in Tables 4.34 and 4.35. Their analysis shows the same tendencies as those observed for the SPIHT coder.

5) Discussion

Above, we have considered tendencies for three real-life three-channel images presented as 8-bit 2D data for each component. In practice, images can be presented differently, for example, by 16-bit data [219] or by 10-bit data after certain normalization [224]. Then, a question arises how to provide the desired PSNR-HVS-M (e.g., 40 dB) for a given multichannel image to be compressed. To answer this question, let us consider some data. First, Fig.4.25 taken from [42] presents the dependences of PSNR-HVS-M on QS for nine 8-bit grayscale test images of different complexity for AGU. The average curve obtained for these nine test images is also given. It is seen that this average curve allows the approximate setting of QS to provide the desired PSNR-HVS-M. For example, to

provide PSNR-HVS-M ≈ 43 dB, one has to set $QS_{rec} \approx 15$. To provide PSNR-HVS-M ≈ 40 dB, it is possible to set $QS_{rec} \approx 20$. If D is not equal to 255, then the recommended QS is

$$QS_{recD} = \frac{QS_{rec}D}{255} \quad (4.23)$$

where QS_{rec} is determined from the average curve in Fig.4.25.

As it follows from the analysis of data in Fig.4.25, the use of QS_{rec} or QS_{recD} provides PSNR-HVS-M **only** approximately, errors can be up to 1 . . . 2 dB depending upon the complexity of an image to be compressed. Such accuracy can be treated as acceptable, since, as it is shown in the previous Section, change of PSNR-HVS-M by even 2 dB does not lead to radical changes of P_{cc} and probabilities of correct classification for particular classes. Second, if errors in providing the desired PSNR-HVS-M are inappropriate, accuracy can be improved by applying a two-step procedure proposed in [42]. As it has been shown in Table 4.15, QS that should be used in component-wise compression is practically the same for all components. This means that it is enough to determine QS_{rec} or QS_{recD} for one component image and then apply it for compressing other components (this can save time and resources at the data compression stage). Moreover, QS_{rec} or QS_{recD} determined according to recommendations given above can be used in joint compression of all components of multichannel images or groups of components [224] by the 3D version of AGU. In this case, the positive effect is twofold. First, a larger CR is provided compared to the component-wise compression. Second, a slightly larger quality of the compressed image can be ensured. Fig2.11.c demonstrates the dependences PSNR-HVS-M on QS for different test images.

In summary, in order to achieve safer efficiency in remote sensing images, acquired images (e.g., on-board) are subject to “careful” lossy compression in Quality Control Compression Unit where PCC (e.g., QS) is determined using the desired threshold for a chosen quality metric (e.g., 40 dB for PSNR-HVS-M) and rate/distortion curves obtained in advance (like those in Fig 2.11.c). PCC corrections can be done if, e.g., images are normalized before compression. Finally, classification or other operations on the RS data are allowed.

4.3.3 Lossy compression of the three-channel remote sensing images with controllable quality

In this section, we test whether it is possible to control quality for 3D compression. Compared to component-wise compression, 3D approach leads to two important benefits: 1) compression ratio can be almost twice larger; 2) probability of correct classification can be slightly better. These benefits are confirmed for real-life three-component data acquired by Sentinel sensor using maximum likelihood-based classifier.

1) Used coders, quality metrics, and classification

There are numerous lossy compression techniques proposed so far. Some of them can be applied only component-wise, other are specially designed for three-dimensional (3D) case, i.e. for color and multichannel images. In addition, advanced methods usually exploit inter-channel correlation of data inherent for multichannel images[209]to improve compression ratio without losing image quality [110, 218, 219]. Consequently, it is worth analyzing is the 3D version of AGU able to provide them simultaneously with one more benefit – providing of a larger CR compared to component-wise processing. Note that 3D AGU performs 1D three-point DCT to decorrelate data as preliminary operation before 2D compression of obtained data. Decompression is performed in reverse order.

In this Section, we consider the simplest version of ML classification [54] which is pixel-wise and employs only pixel values. Second, classification is based on preliminary training. At this stage, sample data for chosen classes are obtained and distributions for them are estimated or modeled. Decision rule is chosen. Then, the classifier is applied to RS images to be classified.

To describe a class of spectral features, we have applied histogram approximation by Johnson SB-distribution [225] that has four adjustable parameters that allow performing approximation quite accurately. Four classes are supposed to be present in RS data, namely, Water, Vegetation, Bare Soil, and Urban. Training samples are of size from 4000 to 20000 pixels. Verification samples have by several times larger sizes. Examples of images to be compressed and classified are given in Fig. 4.25. These are three-channel Sentinel-2 multispectral data in color representation. These are images of Kharkov

(Ukraine) and country-side region near it collected at the end of August 2019. Both fragments are of size 512x512 pixels. The examples are selected to represent simple structure (a) and complex-structure (b) RS data.



Figure 4.25. Fragments of the considered three-channel images for a) country-side and b) urban regions

2) Characteristics of the considered coders

First, let us analyze the conventional performance characteristics for component-wise and 3D compression, i.e., dependence of CR and PSNR-HVS-M on QS. Since performance depends on image complexity, data for both image fragments in Fig. 4.25 are given in Tables 4.34 and 4.35, respectively.

Analysis of data in Tables 4.34 and 4.35 allows concluding the following: 1) CR really depends not only on QS, but on image complexity; for the same QS, CR is about two times larger for the simpler structure image (Table 4.34) than for highly textural image (Table 4.35); 2) CR and PSNR-HVS-M are quite close for a given QS for all three components for the case of component-wise compression of three-channel images: 3) CR for 3D compression for a given image and given QS is almost two times larger than for the component-wise compression (where CR is approximately equal to CR averaged for three components); 4) PSNR-HVS-M for a given QS for 3D and component-wise compression are almost equal as well; this means that methodology of QS setting for providing a desired quality earlier designed for component-wise (grayscale) image compression [54] is

valid for the 3D case as well; in particular, to provide lossy compression at the threshold of distortion invisibility (about 41 dB), it is necessary to set QS about 17.

Table 4.34 - Performance characteristics for 3D and component-wise compression of the image fragment in Fig 4.25.a

QS	3D compression		Red component		Green component		Blue component	
	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M
5	6,49	51,0	5,3	51,8	4,1	52,8	3,4	53,5
10	12,34	44,5	8,48	45,9	6,93	45,6	5,82	45,5
15	18,51	41,1	11,37	42,7	9,53	41,8	8,37	41,2
20	25,33	38,9	14,15	40,4	12,01	39,4	10,84	38,4
25	32,36	37,6	16,92	38,5	14,41	37,4	13,39	36,4
30	39,95	35,9	19,54	36,9	16,88	35,9	16,03	34,8
35	47,94	34,8	22,14	35,7	19,35	34,5	18,59	33,5

Table 4.35 - Performance characteristics for 3D and component-wise compression of the image fragment in Fig 4.25.b

QS	3D compression		Red component		Green component		Blue component	
	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M	CR	PSNR-HVS-M
5	3,41	51,9	2,30	53,9	2,14	55,5	2,21	53,9
10	5,65	44,9	3,14	47,3	2,89	47,2	3,02	46,7
15	8,43	41,1	3,93	43,5	3,62	42,7	3,78	32,8
20	11,62	38,6	4,71	40,5	4,39	39,5	4,52	40,0
25	14,41	36,8	5,55	38,2	5,22	37,2	5,30	37,8
30	17,33	35,4	6,42	36,4	6,10	35,3	6,13	36,0
35	20,39	34,3	7,39	34,7	7,07	33,8	7,02	34,5



a



b

Figure 4.26 - Test image compressed in a) 3D and b) component-wise manner

Thus, for a given QS, there is visually no essential difference between images compressed in 3D and component-wise manner. To prove this, Fig. 4.26 shows compressed images for QS=20 (for the image in Fig. 4.25.b).

Some minor differences compared to the image in Fig. 4.25.b can be noticed, but, in general, visual quality of images in Fig. 4.26.a and 26.b is of the same order. Meanwhile, it may be interesting to know if these differences affect the accuracy of the three-channel classification by the ML classifier.

3) Classification results

As it has been already mentioned, the first stage is the ML classifier training. Fig. 4.27.a illustrates the parts of the test image in Fig. 4.26. a that have been used as training samples. Here blue color corresponds to Water, green – to Vegetation, black – to Bare Soils. Yellow – to Urban. Fig. 4.26.b shows the image fragments for which verification characteristics have been determined. As one can see, there is no coincidence between fragments employed in training and verification.

To provide equal conditions for comparison of classification results for component-wise and 3D compression, we have obtained images with quality characterized by PSNR-HVS-M equal to 45, 42, 39, 36, 33, and 30 dB. Confusion matrix have been obtained for each case to seen transformations happening for each class and their pairs. Let us present some examples of the obtained results.

The confusion matrix for the fragment in Fig. 4.25.a is given in Table 4.36, for the fragment in Fig.4.25.b – in Table 4.37 (in other words, the classification has been carried out for original (uncompressed) data). As one can see, a simpler structure image is, in general, classified better. Urban is a wide class which is misclassified in the largest extent. Bare Soil class is not detected well too.

Table 4.36 - Confusion matrix for the fragment in Fig. 4.25.a

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,883	1,24E-03	0,031	0,084
Water	1,69E-03	0,997	1,42E-03	0
Vegetation	1,36E-03	0,051	0,911	0,037
Bare soil	0,303	0	6,28E-04	0,696

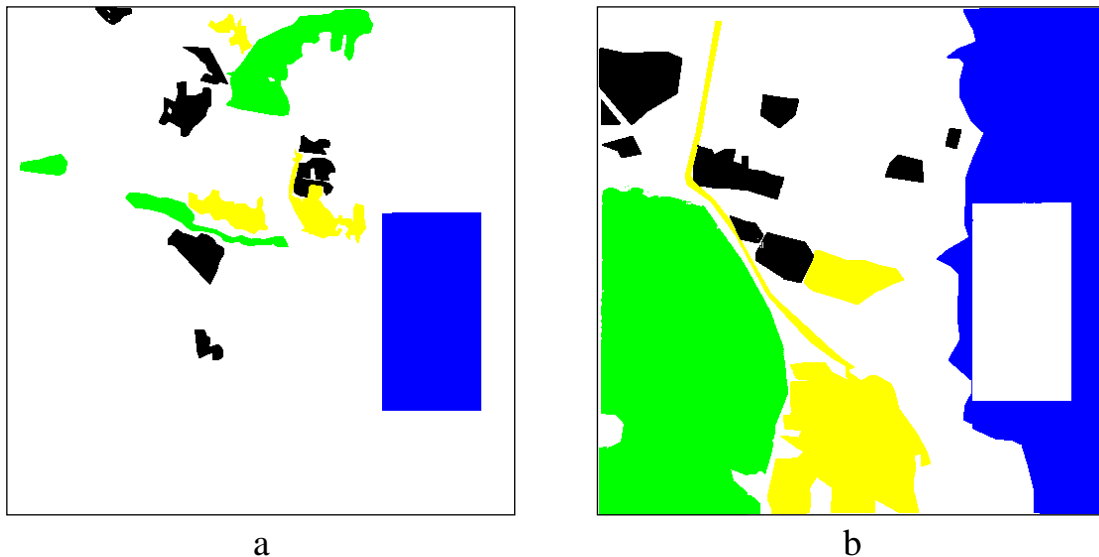


Figure 4.27. Fragments used for a) training and b) verification

Table 4.37 - Confusion matrix for the fragment in Fig.4.25.b

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,423	2,30E-02	0,177	0,377
Water	2,60E-02	0,863	9,20E-02	0,018
Vegetation	7,21E-03	2,20E-02	0,97	6,18E-04
Bare soil	0,174	6,47E-03	0,044	0,775

Table 4.38. Confusion matrix for the fragment in Fig. 4.25.a compressed component-wise with providing PSNR-HVS-M equal to 39 dB

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,891	7,71E-04	0,03	0,078
Water	2,09E-03	0,995	3,07E-03	0
Vegetation	9,04E-04	0,029	0,965	5,44E-03
Bare soil	0,282	0	6,28E-04	0,717

Table 4.39. Confusion matrix for the fragment in Fig. 4.25.a compressed by 3D coder with providing PSNR-HVS-M equal to 39 dB

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,90	1,63E-03	0,026	0,072
Water	1,90E-03	0,997	1,50E-03	0
Vegetation	4,72E-04	0,019	0,978	2,79E-03
Bare soil	0,178	0	7,07E-04	0,821

Data in Tables 4.38 and 4.39 are confusion matrices for the case PSNR-HVSM-39 dB for two ways of lossy compression for the image shown in Fig. 4.25.a.

As one can see from comparison of data in Tables 4.36 and 4.38, the compressed image is classified even better than the original one. Comparison of data in Table 4.39 to the corresponding data in Tables 4.36 and 4.38 shows that the image compressed by the 3D coder is classified better than for two previous cases. Note that CR for 3D compression is about 25 whilst for component-wise compression CR is about 13. Thus, we have two benefits for 3D compression for the considered case – a considerably larger CR and a noticeably better classification.

Let us present some classification examples. Fig. 4.28.a shows classification results for the original (uncompressed) image. Figures 4.28.b, 4.28.c, and 4.28.d present classification results for images compressed with providing PSNR-HVS-M equal to 42 dB (invisible distortions), 36 dB (visible distortions), and 30 dB (sufficient distortions). As one can see, some separate class pixels are lost if CR increases (and image quality reduces). Meanwhile, classification results for the original image and the image compressed with invisible distortions (see classification maps in Figures 4.28.a and 4.28.b) are quite similar.

It can be also interesting to analyze aggregate probabilities of correct classification. They are equal to 0,872, 0,914, 0,931, and 0,948, respectively. It is possible to compare these probabilities to the case of component-wise compression. The probabilities are equal to 0,891, 0,891, and 0,885 for images compressed with PSNR-HVS-M equal to 42, 36, and 30 dB, respectively. In other words, probability of correct classification for all compressed images is higher than for the original one although the training has been done for the original image. Meanwhile, classification of the images compressed in 3D way is more efficient.

Consider now the results for another test image (Fig. 4.25.b). Confusion matrices for this image compressed with PSNR-HVS-M=39 dB are given in Tables 4.40 and 4.41 for the cases of component-wise and 3D compression.

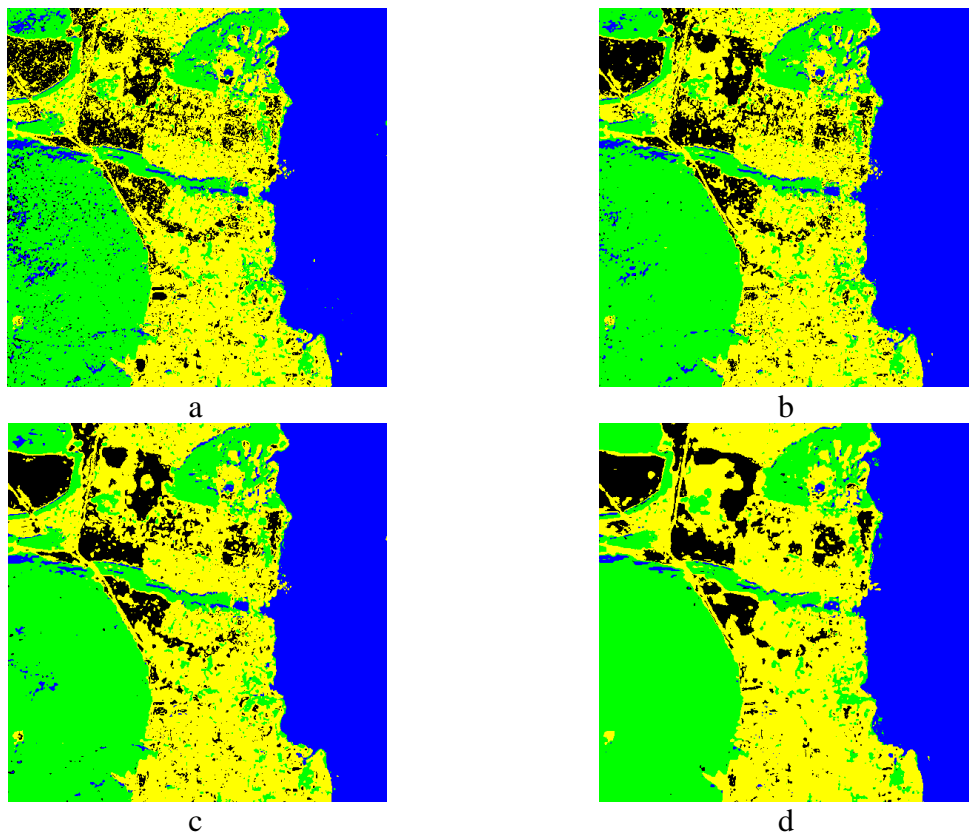


Figure 4.28 - Classification results, a) original image, b) compressed images with PSNR-HVS-M=42 dB, c) PSNR-HVS-M=36 dB (c), and PSNR-HVS-M=30 dB (d)

Table 4.40 - Confusion matrix for the fragment in Fig.4.25.b compressed component-wise with providing PSNR-HVS-M equal to 39 dB.

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,463	0,021	0,147	0,369
Water	0,06	0,793	0,132	0,016
Vegetation	3,17E-03	4,25E-03	0,992	2,83E-04
Bare soil	0,201	4,80E-03	0,045	0,749

Table 4.41. Confusion matrix for the fragment in Fig. 4.25.b compressed by 3D coder with providing PSNR-HVS-M equal to 39 dB.

Class	Probability of Decision			
	Urban	Water	Vegetation	Bare soil
Urban	0,406	0,027	0,165	0,401
Water	0,022	0,793	0,165	0,019
Vegetation	0,022	0,018	0,96	4,38E-04
Bare soil	0,181	3,40E-03	0,043	0,773

The choice of these images for comparison is due to their different statistical characteristics, in particular, the type and complexity of textures, the number of small-sized objects, etc. Considering these differences, compression algorithms have different effects on the classification results. For the image in Fig. 4.25.a with a relatively simple structure, which represents mainly area objects, compression leads to the suppression of high-frequency spatial noise, which can improve the classification accuracy. If the image contains many linear and small-sized objects, as, for example, in Fig. 4.25.b, then compression can lead to a decrease in the reliability of their recognition.

As one can see, there are certain differences for data in Tables 4.37, 4.40, and 4.41. In this case, the results for 3D compression are slightly worse. Compare the aggregate probabilities of correct classification. They are equal to 0,758, 0,731, 0,726, and 0,715 for the original image and the images compressed by 3D coder with providing PSNR-HVS-M equal to 42, 36, and 30 dB, respectively. If component-wise compression is applied, the probabilities are equal to 0,747, 0,753, and 0,754 for images compressed with PSNR-HVS-M equal to 42, 36, and 30 dB, respectively. In other words, probability of correct classification for all compressed images is smaller than for the original one. Classification of the images compressed in 3D way is less efficient. To avoid sufficient reduction of aggregate probability of correct classification, it is worth providing PSNR-HVS-M not less than 36 dB.

It might be interesting to consider the classification results for separate classes. Table 4.42 and 4.43 give data for the image in Fig. 4.25.a for component-wise and 3D compression, respectively. Abbreviation 1a_45 means that the image has been compressed with providing PSNR-HVS-M=45 dB. Dependences for classes are different. The class Urban is recognized slightly better for the compressed images. Detection of Water becomes slightly worse if CR increases and image quality reduces. The class Vegetation is considerably better recognized in the compressed images. The class Bare Soil is recognized with approximately the same probability of correct classification for component-wise compression but it is better recognized in the compressed images for the case of 3D compression. Tables 4.44 and 4.45 present probabilities of correct classification for the image in Fig. 4.25.b.

Table 4.42 - Probabilities of correct classification for classes and aggregate for the test image in Fig. 4.25.a compressed component-wise

Classes	Original	1a_45	1a_42	1a_39	1a_36	1a_33	1a_30
Urban	0,883	0,89	0,89	0,891	0,892	0,894	0,893
Water	0,997	0,996	0,996	0,995	0,994	0,992	0,990
Vegetation	0,911	0,955	0,961	0,965	0,972	0,978	0,984
Bare soil	0,696	0,715	0,716	0,717	0,705	0,698	0,675
P_{total}	0,872	0,889	0,891	0,892	0,891	0,891	0,885

Table 4.43- Probabilities of correct classification for classes and aggregate for the test image in Fig. 4.25.a compressed in 3D manner

Classes	Original	1a_45	1a_42	1a_39	1a_36	1a_33	1a_30
Urban	0,883	0,894	0,896	0,90	0,908	0,919	0,938
Water	0,997	0,996	0,996	0,997	0,997	0,997	0,994
Vegetation	0,911	0,965	0,97	0,978	0,981	0,989	0,994
Bare soil	0,696	0,769	0,793	0,821	0,839	0,87	0,866
P_{total}	0,872	0,906	0,914	0,924	0,931	0,944	0,948

Table 4.44- Probabilities of correct classification for classes and aggregate for the test image in Fig. 4.25.b compressed component-wise

Classes	Original	1b_45	1b_42	1b_39	1b_36	1b_33	1b_30
Urban	0,423	0,437	0,448	0,463	0,486	0,499	0,509
Water	0,863	0,81	0,799	0,793	0,783	0,774	0,764
Vegetation	0,97	0,99	0,992	0,992	0,993	0,994	0,994
Bare soil	0,775	0,754	0,75	0,749	0,751	0,753	0,751
P_{total}	0,758	0,748	0,747	0,749	0,753	0,755	0,754

Table 4.45. Probabilities of correct classification for classes and aggregate for the test image in Fig.4.25.b compressed in 3D manner

Classes	Original	1b_45	1b_42	1b_39	1b_36	1b_33	1b_30
Urban	0,423	0,431	0,418	0,406	0,396	0,379	0,367
Water	0,863	0,816	0,798	0,793	0,771	0,744	0,749
Vegetation	0,97	0,931	0,943	0,96	0,97	0,974	0,97
Bare soil	0,775	0,764	0,765	0,773	0,766	0,766	0,773
P_{total}	0,758	0,736	0,731	0,733	0,726	0,716	0,715

Here, i.e., abbreviation 1b_42 means that the image in Fig. 4.25.b has been compressed with providing PSNR-HVS-M=42 dB. Dependences for classes and compression methods are different. The class Urban is recognized slightly better for the images compressed component-wise (Table 11) but worse for 3D compressed images (Table 4.45). Water classification becomes worse if CR increases and image quality reduces for both compression techniques. The class Vegetation is better recognized in the images compressed component-wise. If 3D compression is applied, classification accuracy is approximately at the same level. The class Bare Soil is better recognized in images compressed in 3D manner. In general, probability of correct classification for this class is almost the same as for original data.

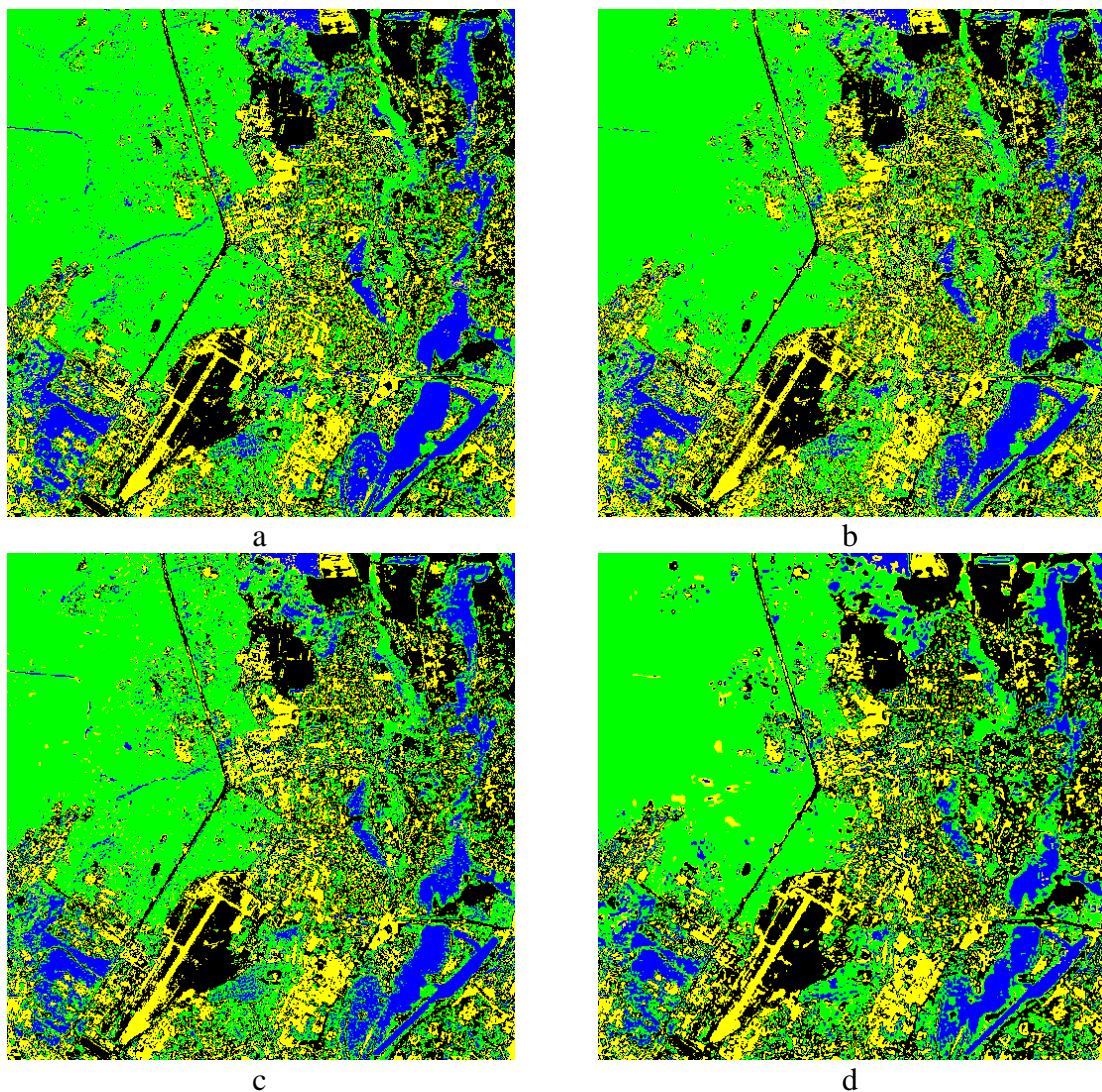


Figure 4.29 - Classification results for original (a) and compressed images with PSNR-HVS-M=39 dB for component-wise (b) and 3D (c) compression, and for the image compressed with PSNR-HVS-M=30 dB (d) in 3D manner

We would like to present some classification maps for the image in Fig. 4.25.b. Fig. 4.29.a represents the classification map for original image. Figures 4.29.b and 4.29.c show the classification results for images compressed with providing PSNR-HVS-M=39 dB for component-wise and 3D manners, respectively. Finally, Fig.4.29.d present the map for the image compressed in 3D manner with providing PSNR-HVS-M=30 dB, i.e. with sufficient distortions. As one can see, the results in Figures 4.29.a, 4.29.b, and 4.29.c are quite similar whilst some details are lost for the classification map in Fig. 4.29.d. Meanwhile, since we know this region well, it is possible to state that there are quite many cases when asphalt roads, clearings in the forest and shadows are classified as bare soil. So, in fact, more classes could be used in analysis. The problem is that a limited number of pixels relate to these classes and it is quite difficult to carry out training for them.

With pixel-wised statistical classification, recognition errors are inevitable if distributions describing different classes intersect. Compression leads to some blurring of the image, the appearance of noise halos around the sharp edges of objects, color noise, therefore, the original distributions of class features change their shape. In turn, this affects the classification results. In general, more compression of a complex-structure image leads to a decrease in total probability of correct classification. It can be noted that for such classes as Urban and Vegetation for the test image in Fig. 4.25.b compressed component-wise, the classification accuracy increases, presumably due to smoothing of inhomogeneities.

4.4 Conclusion

In this chapter, methods for providing desired visual quality values in lossy compression are extended. Remote sensing images, including multichannel ones, are studied more thoroughly.

First, an automatic optimization method for the BPG coder for color images has been proposed for the classic metric PSNR and the HVS-based metric PSNR-HMA. A comparison of decompressed images shows that the use of the metric PSNR-HMA is

preferable since it takes into account peculiarities of human perception of images better. If the error is less than a certain value (e.g., 0,5dB), the procedure quits automatically.

Second, the improved two-step compression method is utilized to compress three-channel remote sensing images. An analysis is carried out for two visual metrics, FSIM and MDSI, which are highly correlated with human subjective scoring results. Three visual quality levels have been proposed corresponding to the MDSI values appropriate for excellent quality, good quality, as well as middle and bad quality, respectively. It has been demonstrated that the two-step procedures work well for the aforementioned metric, providing improvement of accuracy by one order of magnitude for the second step compared to the first one.

Finally, the effect of controlling visual quality on remote sensing data classification is studied. Analysis of classification accuracy is performed for the overall probability of correct recognition of classes in the real satellite images compressed with AGU coder, including both component-wise version and 3D version. It is shown that possible to control quality for 3D AGU compression to obtain considerably (by about two times) higher CR and a better probability of correct classification than in the component-wise case. The results show that classification accuracy usually starts to decrease faster when compressed image quality due to CR increasing reach the distortion visibility threshold that can be characterized by PSNR-HVS-M. There is no sufficient decrease in classification accuracy if images are compressed without visible distortions. These conclusions are drawn based on experiments performed for Landsat and Sentinel RS data.

CONCLUSION

The dissertation research solves several actual tasks related to the development of efficient methods of lossy compression, which are focused on providing a desired visual quality for decompressed image that impacts image perception and/or further image processing. However, the analysis of the existing methods of quality (introduced distortion) control shows that the accuracy of quality providing and time efficiency are difficult to balance. In this regard, there is an urgent need to predict the quality of decompressed images and calculate appropriate CCP at a reasonable time cost to guarantee the distortion introduced by lossy compression does not influence the further image processing or terminal user's perception.

According to the results of the dissertation, we can draw the following conclusions:

1. Analysis of the requirements for the lossy compression techniques, shows that fast compression and high compression ratios are generally desired. Meanwhile, distortion control is very important, even primary.

It has been shown that how much loss can be allowed is application-dependent, but no matter what, the distortions are desired to be controlled carefully. Moreover, selection of the image quality metric is also important since different metrics use different HVS models, so that the reliability of the indication of image quality based on different final tasks is also different.

2. The method of predicting the visual quality for lossy compression has been improved for AGU coder, and extended to SPIHT coder. It has facilitated a fast and considerably better parameter setting.

3. The two-step lossy compression method is developed for gray-scale images to provide a desired visual quality. The experiments conducted for AGU coder and BPG coder showed that the two-step method could achieve a balance between accuracy and time efficiency and only requires at most two compressions and one decompression to achieve appropriate accuracy.

The performance of method for the BPG coder is better than for AGU, particularly in terms of the metric PSNR, the variances of the former are about $1/4 \sim 1/10$ of the latter. For

BPG, even a single step can meet the accuracy requirements according to the metrics PSNR and PSNR-HVS-M for some images.

4. The two-step method for gray-scale images to avoid the occurrence of parameter “over-correction” has been proposed and it has ensured the positive effect of parameter correction in the second-step compression on visual quality error control, thus extending the use of the two-step method to the DCT-based ADCT coder and the DWT-based SPIHT coder. Besides, the pre-classification and composite schemes were adopted to further improve the accuracy of the two-step method. For the SPIHT encoder, the overall variance drops by 45%.

5. Extension of the method on BPG coder for providing desired visual quality to color images and remote sensing images, including multichannel ones has been done. It has been shown that the HVS-based metric PSNR-HMA is more reliable than PSNR. Moreover, the result according to the other two metrics, namely FSIM and MDSI, which are highly correlated with human subjective scoring results, were analyzed. For the first time, three visual quality levels have been proposed corresponding to the MDSI values appropriate for excellent quality, good quality, as well as middle and bad quality.

6. The effect of controlling visual quality on remote sensing data classification has been studied; the overall classification accuracy and probability of correct recognition of class have been analyzed for real satellite images compressed with AGU coder, including both component-wise version and 3D version. It has been shown that it is possible to control quality for 3D AGU compression to obtain considerably higher CR and a better probability of correct classification than in the component-wise case. The results also show that classification accuracy usually starts to decrease faster when compressed image quality due to CR increasing reaches the distortion visibility threshold that can be characterized by PSNR-HVS-M. There is no sufficient decrease in classification accuracy if images are compressed without visible distortions (PSNR-HVS-M exceeds 40 . . . 42 dB). Thus, providing the desired visual quality in remote sensing image compression can maximize the compression ratio without compromising classification accuracy.

Reference

1. N. H. Barnouti, "Fingerprint recognition improvement using histogram equalization and compression methods," *International Journal of Engineering Research and General Science*, vol. 4, no. 2, pp. 685-692, 2016.
2. Y. Ding, *Visual quality assessment for natural and medical image*. Springer, 2018.
3. M. P. Reddy, "The lossless medical image compression for telemedicine applications with delimiter," *Jour of Adv Research in Dynamical Control Systems*, vol. 10, no. 3, pp. 74-79, 2018.
4. J. Jinju, N. Santhi, K. Ramar, and B. S. Bama, "Spatial frequency discrete wavelet transform image fusion technique for remote sensing applications," *Engineering Science Technology, an International Journal* vol. 22, no. 3, pp. 715-726, 2019.
5. S. Doss, S. Pal, D. Akila, S. Jeyalakshmi, T. N. Jabeen, and G. Suseendran, "Satellite image remote sensing for identifying aircraft using SPIHT and NSCT," *Journal of Critical Reviews* vol. 7, no. 5, pp. 631-634, 2020.
6. N. Joshi *et al.*, "A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring," *Remote Sensing*, vol. 8, no. 1, p. 70, 2016.
7. B. AlShehri, A. Alotaibi, H. Alqhtani, A. AlAli, and H. Kurdi, "A Mobile Platform for Social Media Filters," *Procedia Computer Science*, vol. 170, pp. 297-302, 2020.
8. A. A. Ghani, S. I. Hisham, N. A. A. Usop, and N. B. A. Warif, "Watermarking Techniques for Mobile Application: A Review," in *International Conference of Reliable Information and Communication Technology*, 2020, pp. 796-806.
9. M. O'Connor, "Tested: Huawei P30 Pro," *Australian Photography*, no. Jun 2019, pp. 66-68, 2019.
10. M. Zappavigna, "Social media photography: construing subjectivity in Instagram images," *Visual Communication*, vol. 15, no. 3, pp. 271-292, 2016.

11. Z. Men, P. Wang, J. Chen, C. Li, W. Liu, and W. Yang, "Advanced high-order nonlinear chirp scaling algorithm for high-resolution wide-swath spaceborne SAR," *Chinese Journal of Aeronautics*, vol. 34, no. 2, pp. 563-575, 2021.
12. J. Yu, Y. Li, X. Zheng, Y. Zhong, and P. He, "An effective cloud detection method for Gaofen-5 Images via deep learning," *Remote Sensing*, vol. 12, no. 13, p. 2106, 2020.
13. В. Р. України. (2018). *Про схвалення Концепції Загальнодержавної цільової науково-технічної космічної програми України на 2018-2022 роки*. Available: <https://zakon.rada.gov.ua/laws/show/629-2018-%D1%80#Text>
14. J. Xie *et al.*, "Design and data processing of China's first spaceborne laser altimeter system for earth observation: GaoFen-7," *IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing*, vol. 13, pp. 1034-1044, 2020.
15. X. Tong, W. Zhao, J. Xing, and W. Fu, "Status and development of china high-resolution earth observation system and application," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 3738-3741
16. B. Ryan and O. Ochiai, "The New 10-Year GEOSS Strategy for 2016 and Beyond," in *Satellite Earth Observations and Their Impact on Society and Policy*: Springer, Singapore, 2017, pp. 129-136.
17. Y. Ma *et al.*, "Remote sensing big data computing: Challenges and opportunities," *Future Generation Computer Systems*, vol. 51, pp. 47-60, 2015.
18. F. Li, S. Krivenko, and V. Lukin, "A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality," in *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2020, pp. 307-312, doi: 10.1109/DESSERT50317.2020.9125000.
19. I. Blanes, E. Magli, and J. Serra-Sagrasta, "A tutorial on image compression for optical space imaging systems," *IEEE Geoscience Remote Sensing Magazine*, vol. 2, no. 3, pp. 8-26, 2014.
20. B. Penna, T. Tillo, E. Magli, and G. Olmo, "Transform coding techniques for lossy hyperspectral data compression," *IEEE Transactions on Geoscience Remote Sensing*, vol. 45, no. 5, pp. 1408-1421, 2007.

21. C. Shi, L. Wang, J. Zhang, F. Miao, and P. He, "Remote Sensing Image Compression based on direction lifting-based block transform with content-driven Quadtree coding adaptively," *Remote Sensing*, vol. 10, no. 7, p. 999, 2018.
22. E. Syahrul, "Lossless and nearly-lossless image compression based on combinatorial transforms," Université de Bourgogne, 2011.
23. A. J. Hussain, A. Al-Fayadh, and N. Radi, "Image compression techniques: A survey in lossless and lossy algorithms," *Neurocomputing*, vol. 300, pp. 44-69, 2018.10.1016/j.neucom.2018.02.094.
24. B. Aiazzi, L. Alparone, S. Baronti, C. Latri, and M. Selva, "Spectral distortion in lossy compression of hyperspectral data," *Journal of Electrical Computer Engineering*, vol. 2012, 2012.10.1155/2012/850637.
25. V. Fedoseev and T. Androsova, "Watermarking algorithms for JPEG 2000 lossy compressed images," *Information Technology Nanotechnology*, pp. 366-370, 2019.
26. B. Chen, Y. Wu, G. Coatrieux, X. Chen, and Y. Zheng, "JSNet: A simulation network of JPEG lossy compression and restoration for robust image watermarking against JPEG attack," *Computer Vision and Image Understanding*, vol. 197, p. 103015, 2020.
27. R. Matsumoto, "Introducing the perception-distortion tradeoff into the rate-distortion theory of general information sources," *IEICE Communications Express*, vol. 7, no. 11, pp. 427-431, 2018.
28. S. Krivenko, O. Krylova, E. Bataeva, and V. Lukin, "Smart lossy compression of images based on distortion prediction," *Telecommunications Radio Engineering*, vol. 77, no. 17, 2018.
29. H. Jiang, K. Yang, T. Liu, and Y. Zhang, "Remote sensing image compression assessment based on multilevel distortions," *Journal of Applied Remote Sensing*, vol. 8, no. 1, p. 083680, 2014.
30. V. Lukin, A. Zemliachenko, S. Krivenko, B. Vozel, and K. Chehdi, "Lossy compression of remote sensing images with controllable distortions," in *Satellite Information Classification Interpretation* London: IntechOpen, 2018, pp. 27-43.

31. H. L. Tan, Z. Li, Y. H. Tan, S. Rahardja, and C. Yeo, "A perceptually relevant MSE-based image quality metric," *IEEE Transactions on Image Processing*, vol. 22, no. 11, pp. 4447-4459, 2013.
32. N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, and M. Carli, "Modified image visual quality metrics for contrast change and mean shift accounting," in *2011 11th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*, 2011, pp. 305-311.
33. A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *2010 20th international conference on pattern recognition*, 2010, pp. 2366-2369.
34. U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study," *Journal of Computer Communications*, vol. 7, no. 3, pp. 8-18, 2019.
35. D. Tao, S. Di, X. Liang, Z. Chen, and F. Cappello, "Fixed-PSNR Lossy Compression for Scientific Data," presented at the 2018 IEEE International Conference on Cluster Computing (CLUSTER), Belfast, UK, 2018.
36. O. Ieremeiev, V. Lukin, K. Okarma, and K. Egiazarian, "Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images," *Remote Sensing*, vol. 12, no. 15, p. 2349, 2020.
37. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "TID2008—a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30-45, 2009.
38. N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," in *Proceedings of the third international workshop on video processing and quality metrics*, 2007, vol. 4
39. F. Li, "Adaptive two-step method for providing the desired visual quality for SPIHT," *RADIOELECTRONIC COMPUTER SYSTEMS*, no. 1, pp. 195-205, 2022..
40. S. Krivenko, F. Li, V. Lukin, B. Vozel, and O. Krylova, "Prediction of visual quality metrics in lossy image compression," in *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, 2020, pp. 478-483.

41. F. Li, S. Krivenko, and V. Lukin, "A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT," in *Conference on Integrated Computer Technologies in Mechanical Engineering–Synergetic Engineering*, 2020, pp. 17-29.
42. F. Li, S. Krivenko, and V. Lukin, "A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression," in *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, 2020, pp. 502-506.
43. F. Li, S. Krivenko, and V. Lukin, "An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality," in *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, 2020, pp. 13-17.
44. F. Li, S. Krivenko, and V. Lukin, "Analysis of two-step approach for compressing texture images with desired quality," *Aerospace Technic and Technology*, vol. 161, no. 1, pp. 50-58, 2020.
45. F. Li, S. Krivenko, and V. Lukin, "Two-step providing of desired quality in lossy image compression by SPIHT," *Radio electronic and computer systems*, vol. 94, no. 2, pp. 22-32, 2020.
46. F. Li, S. Krivenko, and V. Lukin, "Adaptive two-step procedure of providing desired visual quality of compressed image," in *Proceedings of the 2020 4th International Conference on Electronic Information Technology and Computer Engineering*, 2020, pp. 407-414.
47. F. Li and V. Lukin, "Intelligent lossy compression method of providing a desired visual quality for images of different complexity," presented at the Applied Mathematics, Modeling and Computer Simulation, 2021.
48. F. Li and V. Lukin, "A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images," presented at the 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2021.

49. F. Li, S. Krivenko, and V. Lukin, "An Automatic Optimization Method for BPG Compression Based on Visual Perception," presented at the 2021 International scientific-practical conference, 2021.
50. S. K. a. V. L. F. Li, "Remote Sensing Image Lossy Compression Based on Controlled Visual Quality," in *SPIE*, 2021.
51. F. Li, V. Lukin, O. Ieremeiev, and K. Okarma, "Quality control for the BPG lossy compression of three-channel remote sensing images," *Remote Sensing*, vol. 14, no. 8, p. 1824, 2022.
52. F. Li and V. Lukin, "Providing a Desired Quality of BPG Compressed Images for FSIM Metric," presented at the 2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT), 2021.
53. G. Proskura, I. Vasilyeva, L. Fangfang, and V. Lukin, "Classification of Compressed Multichannel Images and Its Improvement," in *2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA)*, 2020, pp. 1-6
54. V. Lukin *et al.*, "Lossy Compression of Multichannel Remote Sensing Images with Quality Control," *Remote Sensing*, vol. 12, no. 22, p. 3840, 2020.
55. I. Vasilyeva, F. Li, S. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, "Lossy compression of three-channel remote sensing images with controllable quality," in *Image and Signal Processing for Remote Sensing XXVII*, 2021, vol. 11862, pp. 199-210
56. J. Segarra, M. L. Buchaillet, J. L. Araus, and S. C. Kefauver, "Remote sensing for precision agriculture: Sentinel-2 improved features and applications," *Agronomy*, vol. 10, no. 5, p. 641, 2020.
57. R. P. Sishodia, R. L. Ray, and S. K. Singh, "Applications of remote sensing in precision agriculture: A review," *Remote Sensing*, vol. 12, no. 19, p. 3136, 2020.
58. J. Li, Y. Pei, S. Zhao, R. Xiao, X. Sang, and C. Zhang, "A review of remote sensing for environmental monitoring in China," *Remote Sensing*, vol. 12, no. 7, p. 1130, 2020.
59. Z. Shao, H. Fu, D. Li, O. Altan, and T. Cheng, "Remote sensing monitoring of multi-scale watersheds impermeability for urban hydrological evaluation," *Remote Sensing of Environment*, vol. 232, p. 111338, 2019.

60. R. Padmanaban, A. K. Bhowmik, and P. Cabral, "A remote sensing approach to environmental monitoring in a reclaimed mine area," *ISPRS international journal of geo-information*, vol. 6, no. 12, p. 401, 2017.
61. M. M. Nielsen, "Remote sensing for urban planning and management: The use of window-independent context segmentation to extract urban features in Stockholm," *Computers, Environment Urban Systems*, vol. 52, pp. 1-9, 2015.
62. K. Kaku, "Satellite remote sensing for disaster management support: A holistic and staged approach based on case studies in Sentinel Asia," *International Journal of Disaster Risk Reduction*, vol. 33, pp. 417-432, 2019.
63. T. P. P. Sharma, J. Zhang, U. A. Koju, S. Zhang, Y. Bai, and M. K. J. I. j. o. d. r. r. Suwal, "Review of flood disaster studies in Nepal: A remote sensing perspective," vol. 34, pp. 18-27, 2019.
64. M. Shimoni, R. Haelterman, and C. Perneel, "Hypersectral imaging for military and security applications: Combining myriad processing and sensing techniques," *IEEE Geoscience Remote Sensing Magazine*, vol. 7, no. 2, pp. 101-117, 2019.
65. M. J. McAuliffe, F. M. Lalonde, D. McGarry, W. Gandler, K. Csaky, and B. L. Trus, "Medical image processing, analysis and visualization in clinical research," in *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, 2001, pp. 381-386
66. S. Bhattacharya *et al.*, "Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey," *Sustainable cities society*, vol. 65, p. 102589, 2021.
67. J. Priesnitz, C. Rathgeb, N. Buchmann, C. Busch, and M. Margraf, "An overview of touchless 2D fingerprint recognition," *EURASIP Journal on Image Video Processing*, vol. 2021, no. 1, pp. 1-28, 2021.
68. G. Guo and N. Zhang, "A survey on deep learning based face recognition," *Computer vision and image understanding*, vol. 189, p. 102805, 2019.
69. Y. Tong, H. Konik, and A. Tremeau, "Color face-tuned salient detection for image quality assessment," in *2010 2nd European Workshop on Visual Information Processing (EUVIP)*, 2010, pp. 253-260

70. W. Zhu *et al.*, "A multiple attributes image quality database for smartphone camera photo quality assessment," in *2020 IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 2990-2994
71. M. Babic, M. A. Farahani, and T. Wuest, "Image Based Quality Inspection in Smart Manufacturing Systems: A Literature Review," *Procedia CIRP*, vol. 103, pp. 262-267, 2021.
72. H. L. de Moura, V. O. Silva, G. A. Guarneri, T. A. Passarin, D. R. Pipa, and G. P. Pires, "Image-based ultrasound speed estimation for NDT in homogeneous media," *Review of Progress in Quantitative Nondestructive Evaluation*, 2019.
73. C. Li *et al.*, "An underwater image enhancement benchmark dataset and beyond," *IEEE Transactions on Image Processing*, vol. 29, pp. 4376-4389, 2019.
74. R. Bazine, H. Wu, and K. Boukhechba, "Spectral DWT multilevel decomposition with spatial filtering enhancement preprocessing-based approaches for hyperspectral imagery classification," *Remote Sensing*, vol. 11, no. 24, p. 2906, 2019.
75. W. Wang, X. Wu, X. Yuan, and Z. Gao, "An experiment-based review of low-light image enhancement methods," *IEEE Access*, vol. 8, pp. 87884-87917, 2020.
76. S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Transactions on Pattern Analysis Machine Intelligence*, 2021.
77. M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: achievements and challenges," *Journal of digital imaging*, vol. 32, no. 4, pp. 582-596, 2019.
78. A. Z. Chitade and S. Katiyar, "Colour based image segmentation using k-means clustering," *International Journal of Engineering Science Technology*, vol. 2, no. 10, pp. 5319-5325, 2010.
79. X. Wan, H. Song, L. Luo, Z. Li, G. Sheng, and X. Jiang, "Pattern recognition of partial discharge image based on one-dimensional convolutional neural network," in *2018 Condition Monitoring and Diagnosis (CMD)*, 2018, pp. 1-4

80. S. Fang, X. Pan, S. Xiang, and C. Pan, "Meta-msnet: Meta-learning based multi-source data fusion for traffic flow prediction," *IEEE Signal Processing Letters*, vol. 28, pp. 6-10, 2020.
81. A. Zemliachenko, V. Lukin, N. Ponomarenko, K. Egiazarian, and J. Astola, "Still image/video frame lossy compression providing a desired visual quality," *Multidimensional Systems Signal Processing: Image Communication*, vol. 27, no. 3, pp. 697-718, 2016.
82. K. Shinoda, Y. Murakami, M. Yamaguchi, and N. Ohyama, "Multispectral image compression for spectral and color reproduction based on lossy to lossless coding," in *Image Processing: Algorithms and Systems VIII*, 2010, vol. 7532, p. 75320H
83. J. Telles and G. Kemper, "A Multispectral Image Compression Algorithm for Small Satellites Based on Wavelet Subband Coding," in *Brazilian Technology Symposium*, 2019, pp. 181-191, doi: 10.1007/978-3-030-57548-9_17.
84. X. Zhang, "Lossy compression and iterative reconstruction for encrypted image," *IEEE transactions on information forensics security*, vol. 6, no. 1, pp. 53-58, 2010.
85. A. Skodras, C. Christopoulos, and T. Ebrahimi, "The jpeg 2000 still image compression standard," *IEEE Signal processing magazine*, vol. 18, no. 5, pp. 36-58, 2001.
86. M. Singha and K. Hemachandran, "Color Image Segmentation for Satallite Images," *International Journal on Computer Science Engineering Science Technology*, vol. 3, no. 12, p. 3756, 2011.
87. G. Kazakeviciute-Januskeviciene, E. Janusonis, R. Bausys, T. Limba, and M. Kiskis, "Assessment of the Segmentation of RGB Remote Sensing Images: A Subjective Approach," *Remote Sensing*, vol. 12, no. 24, p. 4152, 2020.
88. S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: a review," *Journal of medical systems*, vol. 42, no. 11, pp. 1-13, 2018.
89. S. Dhawan, "A review of image compression and comparison of its algorithms," *International Journal of electronics Communication technology*, vol. 2, no. 1, pp. 22-26, 2011.

90. D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098-1101, 1952.
91. J. J. Rissanen, "Generalized Kraft inequality and arithmetic coding," *IBM Journal of research development*, vol. 20, no. 3, pp. 198-203, 1976.
92. W.-H. Chen and W. Pratt, "Scene adaptive coder," *IEEE Transactions on Communications*, vol. 32, no. 3, pp. 225-232, 1984.
93. G. K. Wallace, "The JPEG still picture compression standard," *IEEE transactions on consumer electronics*, vol. 38, no. 1, pp. xviii-xxxiv, 1992.
94. A. S. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Transactions on image Processing*, vol. 1, no. 2, pp. 244-250, 1992.
95. A. Descampe *et al.*, "JPEG XS--A New Standard for Visually Lossless Low-Latency Lightweight Image Coding," *Proceedings of the IEEE*, vol. 109, no. 9, pp. 1559-1577, 2021.
96. G. Toderici *et al.*, "Full resolution image compression with recurrent neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 5306-5314
97. L. Zhao, H. Bai, A. Wang, and Y. Zhao, "Multiple description convolutional neural networks for image compression," *IEEE Transactions on Circuits Systems for Video Technology*, vol. 29, no. 8, pp. 2494-2508, 2018.
98. K. Marlapalli, R. S. Bandlamudi, R. Busi, V. Pranav, and B. Madhavrao, "A Review on Image Compression Techniques," in *Communication Software and Networks*: Springer, 2021, pp. 271-279.
99. V. S. Thakur, K. Thakur, and S. Gupta, "An improved symbol reduction technique based Huffman coder for efficient entropy coding in the transform coders," *IET Image Processing*, vol. 15, no. 4, pp. 1008-1022, 2021.
100. S. Yuan and J. Hu, "Research on image compression technology based on Huffman coding," *Journal of Visual Communication Image Representation*, vol. 59, pp. 33-38, 2019.

101. J. Minguillon and J. Pujol, "JPEG standard uniform quantization error modeling with applications to sequential and progressive operation modes," *Journal of electronic imaging*, vol. 10, no. 2, pp. 475-485, 2001.
102. M. S. Mahdi and N. F. Hassan, "A proposed lossy image compression based on multiplication table," *Kurdistan Journal of Applied Research*, vol. 2, no. 3, pp. 98-102, 2017.
103. A. Zabala and X. Pons, "Impact of lossy compression on mapping crop areas from remote sensing," *International Journal of Remote Sensing*, vol. 34, no. 8, pp. 2796-2813, 2013.
104. N. N. Ponomarenko, V. V. Lukin, M. Zriakhov, A. Kaarna, and J. Astola, "An automatic approach to lossy compression of AVIRIS images," in *2007 IEEE International Geoscience and Remote Sensing Symposium*, 2007, pp. 472-475
105. G. Mathur, R. Mathu, and M. K. Mathur, "A Comparative Study of Various Lossy Image Compression Techniques," *International journal of Engineering Research & Technology (ETRASCT)*, vol. 2, p. 4, 2014.
106. A. AbuBaker, M. Eshtay, and M. AkhoZahia, "Comparison Study of Different Lossy Compression Techniques Applied on Digital Mammogram Images," (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 12, pp. 149-155, 2016.
107. A. Moon, J. Kim, J. Zhang, and S. W. Son, "Lossy compression on IoT big data by exploiting spatiotemporal correlation," in *2017 IEEE High Performance Extreme Computing Conference (HPEC)*, 2017, pp. 1-7.
108. N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, "DCT based high quality image compression," in *Scandinavian Conference on Image Analysis*, 2005, pp. 1177-1185.
109. N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, "ADCT: a new high quality DCT based coder for lossy image compression," *CD ROM Proceedings of LNLA*, vol. 6, 2008.

110. F. Khelifi, A. Bouridane, and F. Kurugollu, "Joined spectral trees for scalable SPIHT-based multispectral image compression," *IEEE Transactions on Multimedia*, vol. 10, no. 3, pp. 316-329, 2008.
111. D. Yee, S. Soltaninejad, D. Hazarika, G. Mbuyi, R. Barnwal, and A. Basu, "Medical image compression based on region of interest using better portable graphics (BPG)," in *2017 IEEE international conference on systems, man, and cybernetics (SMC)*, 2017, pp. 216-221, doi: 10.1109/SMC.2017.8122605.
112. U. Albalawi, S. P. Mohanty, and E. Kougianos, "A hardware architecture for better portable graphics (BPG) compression encoder," in *2015 IEEE International Symposium on Nanoelectronic and Information Systems*, 2015, pp. 291-296.
113. F. Bellard. (2018, 2021.7.25). *BPG Image format*. Available: <https://bellard.org/bpg/>
114. M. Gupta and A. K. Garg, "Analysis of image compression algorithm using DCT," *International Journal of Engineering Research Applications*, vol. 2, no. 1, pp. 515-521, 2012.
115. A. Raid, W. Khedr, M. A. El-Dosuky, and W. Ahmed, "Jpeg image compression using discrete cosine transform-A survey," *International Journal of Computer Science & Engineering Survey*, vol. 5, no. 2, p. 9, 2014. 10.5121/ijcses.2014.5204.
116. X. Wang, X. Feng, N. He, Q. Jia, and Y. Zou, "JPEG image compression algorithm analysis and local optimization," in *2010 IEEE International Conference on Advanced Management Science (ICAMS 2010)*, 2010, vol. 1, pp. 5-7
117. K. Egiazarian, M. Helsingius, P. Kuosmanen, and J. Astola, "Removal of blocking and ringing artifacts using transform domain denoising," in *1999 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1999, vol. 4, pp. 139-142
118. N. N. Ponomarenko, K. O. Egiazarian, V. V. Lukin, and J. T. Astola, "High-quality DCT-based image compression using partition schemes," *IEEE Signal Processing Letters*, vol. 14, no. 2, pp. 105-108, 2007.
119. N. Ponomarenko, V. Lukin, K. Egiazarian, and J. Astola, "ADCTC: advanced DCT-based image coder," *Proceedings of LNLA, Switzerland*, 2008.

120. N. Brahim, T. Bouden, T. Brahim, and L. Boubchir, "Lossy image compression based on efficient multiplier-less 8-points DCT," *Multimedia Systems*, vol. 28, no. 1, pp. 171-182, 2022.
121. N. Brahim, T. Bouden, T. Brahim, and L. Boubchir, "A novel and efficient 8-point DCT approximation for image compression," *Multimedia Tools Applications*, vol. 79, no. 11, pp. 7615-7631, 2020.
122. H. Kanagaraj and V. Muneeswaran, "Image compression using Haar discrete wavelet transform," in *2020 5th International Conference on Devices, Circuits and Systems (ICDCS)*, Coimbatore, India, 2020, pp. 271-274
123. S. Mallat, "Wavelets for a vision," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 604-614, 1996.
124. T. Acharya and P.-S. Tsai, *JPEG2000 standard for image compression: concepts, algorithms and VLSI architectures*. John Wiley & Sons, 2005.
125. A. Zabala, X. Pons, R. Díaz-Delgado, F. García, F. Aulí-Llinàs, and J. Serra-Sagrà, "Effects of JPEG and JPEG2000 lossy compression on remote sensing image classification for mapping crops and forest areas," in *2006 IEEE International Symposium on Geoscience and Remote Sensing*, 2006, pp. 790-793, doi: 10.1109/IGARSS.2006.203.
126. S. Triantaphillidou, E. Allen, and R. Jacobson, "Image quality comparison between JPEG and JPEG2000. II. Scene dependency, scene analysis, and classification," *Journal of Imaging Science Technology*, vol. 51, no. 3, pp. 259-270, 2007.
127. R. Chourasiya and A. Shrivastava, "A Study of image compression based transmission algorithm Using SPIHT for low bit rate application," *Bulletin of Electrical Engineering Informatics*, vol. 2, no. 2, pp. 117-122, 2013.
128. A. L. E. M. Jamel, "Efficiency Spiht in compression and quality of image," *Journal of the College of Education for Women*, vol. 22, no. 3, pp. 627-637, 2011.
129. Y. Dua, R. S. Singh, K. Parwani, S. Lunagariya, and V. Kumar, "Convolution Neural Network based lossy compression of hyperspectral images," *Signal Processing: Image Communication*, vol. 95, p. 116255, 2021.
130. H. Fu, F. Liang, and B. Lei, "An Extended Hybrid Image Compression Based on Soft-to-Hard Quantification," *IEEE Access*, vol. 8, pp. 95832-95842, 2020.

131. J. Lee, S. Cho, and S.-K. Beack, "Context-adaptive Entropy Model for End-to-end Optimized Image Compression," *arXiv preprint*, p. arXiv:1809.10452 2018.
132. H. Ma, D. Liu, N. Yan, H. Li, and F. Wu, "End-to-end optimized versatile image compression with wavelet-like transform," *IEEE Transactions on Pattern Analysis Machine Intelligence*, vol. 44, no. 3, pp. 1247-1263, 2020.
133. C. Lin, J. Yao, F. Chen, and L. Wang, "A Spatial RNN Codec for End-to-End Image Compression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 13269-13277
134. F. Dufaux, G. J. Sullivan, and T. Ebrahimi, "The JPEG XR image coding standard [Standards in a Nutshell]," *IEEE Signal Processing Magazine*, vol. 26, no. 6, pp. 195-204, 2009.
135. S. Donapati and H. Yagain, "A Comparative study of effects of CSC on Image Compression Ratios while using JPEG-XR," in *2013 International Symposium on Electronic System Design*, Singapore, 2013, pp. 158-161.
136. J. C. Mier, E. Huang, H. Talebi, F. Yang, and P. Milanfar, "Deep Perceptual Image Quality Assessment for Compression," *arXiv preprint*, p. arXiv:2103.01114, 2021.
137. Y. Patel, S. Appalaraju, and R. Manmatha, "Human Perceptual Evaluations for Image Compression," *arXiv preprint*, p. arXiv:1908.04187 2019.
138. F. Mentzer, L. V. Gool, and M. Tschannen, "Learning better lossless compression using lossy compression," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, 2020, pp. 6638-6647.
139. A. Pandey, B. S. Saini, B. Singh, and N. Sood, "Quality controlled ECG data compression based on 2D discrete cosine coefficient filtering and iterative JPEG2000 encoding," *Measurement*, vol. 152, p. 107252, 2020.10.1016/j.measurement.2019.107252.
140. W. L. Lau, Z. L. Li, and K. K. Lam, "Effects of JPEG compression on image classification," *International journal of remote sensing*, vol. 24, no. 7, pp. 1535-1544, 2003.

141. F. García-Vílchez *et al.*, "On the impact of lossy compression on hyperspectral image classification and unmixing," *IEEE Geoscience remote sensing letters*, vol. 8, no. 2, pp. 253-257, 2010.
142. Z. Chen, Y. Hu, and Y. Zhang, "Effects of compression on remote sensing image classification based on fractal analysis," *IEEE Transactions on Geoscience Remote Sensing*, vol. 57, no. 7, pp. 4577-4590, 2019.
143. S. Athar and Z. Wang, "A comprehensive performance evaluation of image quality assessment algorithms," *IEEE Access*, vol. 7, pp. 140030-140070, 2019.
144. H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, "LIVE image quality assessment database release 2 (2005)," ed, 2005.
145. N. Ponomarenko *et al.*, "Image database TID2013: Peculiarities, results and perspectives," *Signal processing: Image communication*, vol. 30, pp. 57-77, 2015.
146. Y. Horita, K. Shibata, Y. Kawayoke, and Z. P. Sazzad, "MICT image quality evaluation database," ed, 2011.
147. Y. Tong, H. Konik, F. Cheikh, and A. Tremeau, "Full reference image quality assessment based on saliency map analysis," *Journal of Imaging Science Technology*, vol. 54, no. 3, pp. 1-15, 2010.
148. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "A comprehensive evaluation of full reference image quality assessment algorithms," in *2012 19th IEEE International Conference on Image Processing*, 2012, pp. 1477-1480.
149. D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE transactions on image processing*, vol. 16, no. 9, pp. 2284-2298, 2007.
150. K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS," in *Proceedings of the Second International Workshop on Video Processing and Quality Metrics*, 2006, vol. 4
151. Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE signal processing letters*, vol. 9, no. 3, pp. 81-84, 2002.

152. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600-612, 2004.
153. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, 2003, vol. 2, pp. 1398-1402.
154. Z. Wang and E. P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," in *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, vol. 2, pp. ii/573-ii/576.
155. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, 2011.
156. L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Transactions on Image processing*, vol. 23, no. 10, pp. 4270-4281, 2014.
157. A. Balanov, A. Schwartz, Y. Moshe, and N. Peleg, "Image quality assessment based on DCT subband similarity," in *2015 IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 2105-2109
158. H. Z. Nafchi, A. Shahkolaei, R. Hedjam, and M. Cheriet, "Mean deviation similarity index: Efficient and reliable full-reference image quality evaluator," *IEEE Access* vol. 4, pp. 5579-5590, 2016.
159. G. Chen, Y. Zhang, and S. Wang, "Hyperspectral remote sensing IQA via learning multiple kernels from mid-level features," *Signal Processing: Image Communication*, vol. 83, p. 115804, 2020.
160. T. Kyono, J. Lucas, M. Werth, B. Calef, I. McQuaid, and J. Fletcher, "Machine learning for quality assessment of ground-based optical images of satellites," *Optical Engineering*, vol. 59, no. 5, p. 051403, 2020.
161. J. Zhang, J. Zhou, M. Li, H. Zhou, and T. Yu, "Quality Assessment of SAR-to-Optical Image Translation," *Remote Sensing*, vol. 12, no. 21, p. 3472, 2020.

162. E. Christophe, "Hyperspectral data compression tradeoff," in *Optical remote sensing*, S. Prasad, L. M. Bruce, and J. Chanussot, Eds. Berlin, Heidelberg: Springer, 2011, pp. 9-29.
163. N. Ponomarenko, V. Lukin, J. Astola, and K. Egiazarian, "Analysis of HVS-metrics' properties using color image database TID2013," in *International Conference on Advanced Concepts for Intelligent Vision Systems*, 2015, pp. 613-624
164. D. Wu, D. M. Tan, and H. R. Wu, "Visually lossless adaptive compression of medical images," in *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, 2003, vol. 1, pp. 458-463
165. H. Oh, A. Bilgin, and M. Marcellin, "Visually lossless JPEG 2000 for remote image browsing," *Information* vol. 7, no. 3, p. 45, 2016.10.3390/info7030045.
166. N. Ye, M. Pérez-Ortiz, and R. K. Mantiuk, "Visibility Metric for Visually Lossless Image Compression," in *2019 Picture Coding Symposium (PCS)*, 2019, pp. 1-5
167. J. Y. Lin *et al.*, "Experimental design and analysis of JND test on coded image/video," in *Applications of digital image processing XXXVIII*, 2015, vol. 9599, p. 95990Z
168. X. Zhang, W. Lin, S. Wang, J. Liu, S. Ma, and W. Gao, "Fine-grained quality assessment for compressed images," *IEEE Transactions on Image Processing*, vol. 28, no. 3, pp. 1163-1175, 2018.
169. B. Bondžulić, N. Stojanović, V. Petrović, B. Pavlović, and Z. Miličević, "Efficient prediction of the first just noticeable difference point for JPEG compressed images," *Acta Polytechnica Hungarica*, vol. 18, no. 8, pp. 201-220, 2021.
170. J. Tichonov, O. Kurasova, and E. Filatovas, "Quality prediction of compressed images via classification," in *International Conference on Image Processing and Communications*, 2016, pp. 35-42
171. S. Krivenko, M. Zriakhov, V. Lukin, and B. Vozel, "MSE and PSNR prediction for ADCT coder applied to lossy image compression," in *2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, 2018, pp. 613-618

172. M. Dejean-Servières, K. Desnos, K. Abdelouahab, W. Hamidouche, L. Morin, and M. Pelcat, "Study of the impact of standard image compression techniques on performance of image classification with a convolutional neural network," INSA Rennes; Univ Rennes; IETR; Institut Pascal, 2017.
173. R. Kozhemiakin, V. Lukin, and B. Vozel, "Image quality prediction for DCT-based compression," in *2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*, 2017, pp. 225-228
174. S. Krivenko, V. Lukin, and B. Vozel, "Prediction of Introduced Distortions Parameters in Lossy Image Compression," in *2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T)*, 2018, pp. 447-451
175. A. C. Cameron and F. A. Windmeijer, "An R-squared measure of goodness of fit for some common nonlinear regression models," *Journal of econometrics*, vol. 77, no. 2, pp. 329-342, 1997.
176. T.-Y. Sung, Y.-S. Shieh, C.-W. Yu, and H.-C. Hsin, "High-efficiency and low-power architectures for 2-D DCT and IDCT based on CORDIC rotation," in *2006 Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'06)*, 2006, pp. 191-196
177. A. N. Zemliachenko *et al.*, "Prediction of compression ratio for DCT-based coders with application to remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing*, vol. 11, no. 1, pp. 257-270, 2017.
178. O. A. Agudelo-Medina, H. D. Benitez-Restrepo, G. Vivone, and A. Bovik, "Perceptual quality assessment of pan-sharpened images," *Remote Sensing*, vol. 11, no. 7, p. 877, 2019.
179. V. Lukin, A. Zemliachenko, S. Abramov, B. Vozel, and K. Chehdi, "Automatic lossy compression of noisy images by spiht or jpeg2000 in optimal operation point neighborhood," in *2016 6th European Workshop on Visual Information Processing (EUVIP)*, 2016, pp. 1-6
180. V. V. Lukin, M. S. Zriakhov, N. N. Ponomarenko, S. S. Krivenko, and M. Zhenjiang, "Lossy compression of images without visible distortions and its application,"

in *IEEE 10th INTERNATIONAL CONFERENCE ON SIGNAL PROCESSING PROCEEDINGS*, 2010, pp. 698-701

181. S. Krivenko, D. Demchenko, I. Dyogtev, and V. Lukin, "A two-step approach to providing a desired quality of lossy compressed images," in *Integrated Computer Technologies in Mechanical Engineering*: Springer, 2020, pp. 482-491.

182. M. Alhihi, A. Zemliachenko, S. Abramov, B. Vozel, K. O. Egiazarian, and V. Lukin, "Pre-requisites for smart lossy compression of noisy remote sensing images," *Telecommunications Radio Engineering*, vol. 77, no. 3, pp. 225-241, 2018.

183. L. Armi and S. Fekri-Ershad, "Texture image analysis and texture classification methods-A review," *arXiv preprint*, p. arXiv:.06554, 2019.

184. K. Okarma and J. Fastowicz, "Computer vision methods for non-destructive quality assessment in additive manufacturing," in *International Conference on Computer Recognition Systems*, 2019, pp. 11-20

185. D. Salomon and G. Motta, *Handbook of data compression*. London; New York: Springer, 2010.

186. P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 895-911, 2009.

187. USC-SIPI. (2021). *The USC-SIPI image database*. Available: <http://sipi.usc.edu/database/database.php?volume=aerials>

188. L. Best-Rowden and A. K. Jain, "Learning face image quality from human assessments," *IEEE Transactions on Information Forensics Security*, vol. 13, no. 12, pp. 3064-3077, 2018.

189. H. Jiang, K. Yang, T. Liu, and Y. Zhang, "Quality prediction of DWT-based compression for remote sensing image using multiscale and multilevel differences assessment metric," *Mathematical Problems in Engineering*, vol. 2014, p. 593213, 2014.

190. D. G. Bonett and T. A. Wright, "Sample size requirements for estimating Pearson, Kendall and Spearman correlations," *Psychometrika*, vol. 65, no. 1, pp. 23-28, 2000.

191. R. A. Kozhemiakin, S. K. Abramov, V. V. Lukin, B. Vozel, and K. Chehdi, "Output MSE and PSNR prediction in DCT-based lossy compression of remote sensing

images," in *Image and Signal Processing for Remote Sensing XXIII*, 2017, vol. 10427, p. 1042721

192. (2021, 2021.01.08). *Fractal coding and analysis group*. Available: <https://links.uwaterloo.ca/Repository/TIF/>

193. D. P. Bavirisetti. (2021, 2021.01.08). *Image Fusion Datasets*. Available: <https://sites.google.com/view/durgaprasadbavirisetti/datasets>

194. K. Okarma and J. Fastowicz, "Improved quality assessment of colour surfaces for additive manufacturing based on image entropy," *Pattern Analysis Applications*, vol. 23, no. 3, pp. 1035-1047, 2020.

195. A. Koschan and M. Abidi, "Detection and classification of edges in color images," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 64-73, 2005.10.1109/MSP.2005.1407716.

196. R. Xiao and M. Yeh, "A new method with saliency detection for image quality assessment," in *2017 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, 2017, pp. 75-76

197. M. Prabhushankar, D. Temel, and G. AlRegib, "Ms-unique: Multi-model and sharpness-weighted unsupervised image quality estimation," *Electronic Imaging*, vol. 2017, no. 12, pp. 30-35, 2017.

198. N. Ponomarenko. (2021/3/27). *PSNR-HMA download page*. Available: <http://www.ponomarenko.info/psnrhma.htm>

199. S. G. Dellepiane and E. Angiati, "Quality assessment of despeckled SAR images," *IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing*, vol. 7, no. 2, pp. 691-707, 2013.

200. P. Wang and V. M. Patel, "Generating high quality visible images from SAR images using CNNs," in *2018 IEEE Radar Conference (RadarConf18)*, 2018, pp. 0570-0575

201. K. Yang and H. Jiang, "Optimized-SSIM based quantization in optical remote sensing image compression," in *2011 Sixth International Conference on Image and Graphics*, 2011, pp. 117-122

202. K.-K. Huang, H. Liu, C.-X. Ren, Y.-F. Yu, and Z.-R. Lai, "Remote sensing image compression based on binary tree and optimized truncation," *Digital Signal Processing*, vol. 64, pp. 96-106, 2017.
203. X. Liang *et al.*, "Significantly improving lossy compression quality based on an optimized hybrid prediction model," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, 2019, pp. 1-26
204. D. M. Chandler, "Seven challenges in image quality assessment: past, present, and future research," *International Scholarly Research Notices*, vol. 2013, 2013.
205. J. Chen *et al.*, "Understanding performance-quality trade-offs in scientific visualization workflows with lossy compression," in *2019 IEEE/ACM 5th International Workshop on Data Analysis and Reduction for Big Scientific Data (DRBSD-5)*, 2019, pp. 1-7
206. N. Ozah and A. Kolokolova, "Compression improves image classification accuracy," in *Canadian Conference on Artificial Intelligence*, 2019, pp. 525-530, doi: 10.1007/978-3-030-18305-9_55.
207. M. Giannopoulos, A. Aidini, A. Pentari, K. Fotiadou, and P. Tsakalides, "Classification of Compressed Remote Sensing Multispectral Images via Convolutional Neural Networks," *Journal of Imaging*, vol. 6, no. 4, p. 24, 2020.
208. I. Demir *et al.*, "A challenge to parse the earth through satellite images," *ArXiv preprint*, p. arXiv:1805.06561
2018.
209. D. Manolakis, R. Lockwood, and T. Cooley, "On the spectral correlation structure of hyperspectral imaging data," in *2008-2008 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Boston, MA, USA, 2008, vol. 2, pp. II-581-II-584
210. N. N. Ponomarenko, V. V. Lukin, M. Zriakhov, A. Kaarna, and J. Astola, "Automatic approaches to on-land/on-board filtering and lossy compression of AVIRIS images," in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, 2008, vol. 3, pp. III-254-III-257

211. M. A. Popov, S. A. Stankevich, L. P. Lischenko, V. V. Lukin, and N. N. Ponomarenko, "Processing of hyperspectral imagery for contamination detection in urban areas," in *Environmental Security and Ecoterrorism*: Springer, 2011, pp. 147-156.
212. R. G. J. I. J. o. W. F. Congalton, "Accuracy assessment and validation of remotely sensed and other spatial information," vol. 10, no. 4, pp. 321-328, 2001.
213. C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A Practical Guide to Support Vector Classification," Department of computer science 2016.
214. I. Vasilyeva and V. Lukin, "Multichannel images post-classification processing techniques analysis," *Radio electronic and computer systems*, vol. 89, no. 1, pp. 17-28, 2019.
215. R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *The Journal of machine Learning research*, vol. 9, pp. 1871-1874, 2008.
216. D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013.
217. G. Yu, T. Vladimirova, and M. N. Sweeting, "Image compression systems on board satellites," *Acta Astronautica*, vol. 64, no. 9-10, pp. 988-1005, 2009.
218. B. Aiazzi, L. Alparone, and S. Baronti, "Near-lossless compression of 3-D optical data," *IEEE Transactions on Geoscience Remote Sensing*, vol. 39, no. 11, pp. 2547-2557, 2001.
219. A. N. Zemliachenko *et al.*, "Lossy compression of hyperspectral images based on noise parameters estimation and variance stabilizing transform," *Journal of applied remote sensing*, vol. 8, no. 1, p. 083571, 2014.
220. V. Lukin, G. Proskura, and I. Vasilyeva, "Improvement of multichannel image classification by combining elementary classifiers," in *2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T)*, 2019, pp. 660-664
221. J. Sun, J. Yang, C. Zhang, W. Yun, and J. Qu, "Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method," *Mathematical Computer Modelling*, vol. 58, no. 3-4, pp. 573-581, 2013.

222. P. S. Sisodia, V. Tiwari, and A. Kumar, "Analysis of supervised maximum likelihood classification for remote sensing image," in *International conference on recent advances and innovations in engineering (ICRAIE-2014)*, 2014, pp. 1-4
223. V. Lukin, "Processing of Multichannel RS data for Environment Monitoring," *Processing and Visual Analytics for Environmental Security*, p. 129, 2009.
224. A. Zemliachenko, O. Ieremeiev, V. Lukin, and B. Vozel, "Peculiarities of hyperspectral image lossy compression for sub-band groups," in *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2019, pp. 918-923
225. B. R. Parresol, *Recovering parameters of Johnson's SB distribution*. US Department of Agriculture, Forest Service, Southern Research Station, 2003.

APPENDIX A.

LIST OF APPLICANT PUBLICATIONS

Articles in scientific periodicals:

1. F. Li, V. Lukin, O. Ieremeiev, and K. Okarma, "Quality control for the BPG lossy compression of three-channel remote sensing images," *Remote Sensing*, vol.14, no. 8, pp.1824, 2022, doi: 10.3390/rs14081824.
2. F. Li, S. Krivenko, and V. Lukin, "Two-step providing of desired quality in lossy image compression by SPIHT," *Radioelectronic and computer systems*, vol. 2, no. 2020, pp. 22-32, 2020, doi: 10.32620/reks.2020.2.02.
3. F. Li, "Adaptive two-step method for providing the desired visual quality for SPIHT," *Radioelectronic and computer systems*, vol.1, no. 2022, pp. 195-205, 2022, doi: 10.32620/reks.2022.1.15.
4. V. Lukin, I. Vasilyeva, S. Krivenko, F. Li, S. Abramov, O. Rubel, B. Vozel, K. Chehdi, and K. Egiazarian, "Lossy Compression of Multichannel Remote Sensing Images with Quality Control," *Remote Sensing*, vol. 12, no. 22: 3840. doi: 10.3390/rs12223840.
5. F. Li, S. Krivenko, and V. Lukin, "Analysis of two-step approach for compressing texture images with desired quality," *Aerospace technic and technology*, vol. 1, no. 2020, pp. 50-58, 2020, doi:10.32620/aktt.2020.1.08.

Scientific works certifying the approbation of the dissertation materials:

6. F. Li, S. Krivenko and V. Lukin, "A Two-step Procedure for Image Lossy Compression by ADCTC With a Desired Quality," *2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Kyiv, Ukraine, 2020, pp. 307-312, doi: 10.1109/DESSERT50317.2020.9125000.
7. S. Krivenko, F. Li, V. Lukin, B. Vozel and O. Krylova, "Prediction of Visual Quality Metrics in Lossy Image Compression," *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, 2020, pp. 478-483, doi: 10.1109/ELNANO50318.2020.9088819.

8. F. Li, S. Krivenko and V. Lukin, “A Fast Method for Visual Quality Prediction and Providing in Image Lossy Compression by SPIHT,” In: Nechyporuk M., Pavlikov V., Kritskiy D. (eds) *Integrated Computer Technologies in Mechanical Engineering- ICTM 2020. Lecture Notes in Networks and Systems*, vol 188. Springer, Cham. doi:10.1007/978-3-030-66717-7_2.
9. F. Li, S. Krivenko and V. Lukin, “A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression,” *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Slavske, Ukraine,2020, pp. 502-506, doi: 10.1109/TCSET49122.2020.235483.
10. F. Li, S. Krivenko and V. Lukin, “An Approach to Better Portable Graphics (BPG) Compression with Providing a Desired Quality,” *2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine,2020, pp. 13-17, doi: 10.1109/ATIT50783.2020.9349289.
11. F. Li, S. Krivenko and V. Lukin, “Adaptive two-step procedure of providing desired visual quality of compressed image,” *2020 4th International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, Xiamen, China, 2020, pp. 407-414, doi:10.1145/3443467.3443791.
12. G. Proskura, I. Vasilyeva, F. Li and V. Lukin, “Classification of Compressed Multichannel Images and Its Improvement,” *2020 30th International Conference Radioelektronika (RADIOELEKTRONIKA)*, Bratislava, Slovakia, 2020, pp. 1-6, doi: 10.1109/RADIOELEKTRONIKA49387.2020.9092371.
13. I. Vasilyeva, F. Li, S. Abramov, V. Lukin, B.Vozel, and K.Chehdi, “Lossy compression of three-channel remote sensing images with controllable quality, ”*Proc. SPIE 11862, Image and Signal Processing for Remote Sensing XXVII*, pp. 118620R, doi: 10.1117/12.2599902.
14. F. Li, V. Lukin, K. Okarma and Y. Fu, “Providing a Desired Quality of BPG Compressed Images for FSIM Metric,”*2021 IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine,2021, pp. 10-14, doi: 10.1109/ATIT54053.2021.9678522.

15. F. Li, S. Krivenko and V. Lukin, “An Automatic Optimization Method for BPG Compression Based on Visual Perception,”*2021 Mathematical Modeling and Simulation of Systems, Selected Papers of 16th International Scientific-practical Conference (MODS)*. Chernihiv, Ukraine, 2021, pp. 213-225. doi:10.1007/978-3-030-89902-8_17.
16. F. Li, V. Lukin, K. Okarma, Y. Fu, and Duan J. “Intelligent lossy compression method of providing a desired visual quality for images of different complexity,”*Applied Mathematics, Modeling and Computer Simulation (AMMCS)*, Wuhan, China, 2021, pp.500-505. doi:10.3233/ATDE220050.
17. F. Li, S. Krivenko and V. Lukin, “A Modified Algorithm of Visual Quality Providing in Lossy Compression of Remote Sensing Images,” *2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo)*, Odesa, Ukraine,2021, pp. 16-21, doi: 10.1109/UkrMiCo52950.2021.9716684.

APPENDIX B.
IMPLEMENTATION ACTS



«CONFIRM»

Dean of the School of Measuring and Optical Engineering
of Nanchang Hangkong University
Professor Song Kai

« — » *[Signature]* 2022 p.

Song Kai

ACT

about implementing the results of scientific work of Li Fangfang
«Design and analysis of efficient methods for providing a desired quality in image lossy
compression»
into educational process of College of Testing and Optoelectronic Engineering of Nanchang
Hangkong University

We, the commission, consisting of:
Chairman - head of the department, Professor Liu Bin,
commission members:

Deputy Director of the department, Associate Professor Jiang Guangyu
Professor Gong Yingqing,

made this act that the results of the scientific work "Design and analysis of efficient methods
for providing a desired quality in image lossy compression" are used in the educational
process of the department, namely:

- methods of lossy image compression are considered in lectures in the course "Digital signal processing";
- the main approaches and method of evaluating the image visual quality are considered in the laboratory work "Three dimensional measurement and detection"
- algorithms and appropriate software for analysis and controlling of distortion are considered in lectures in the course "MATLAB programming".

In addition, scientific research is constantly conducted with the participation of students and graduate students of the department, the results of which are reflected in course and diploma theses and projects, as well as in scientific articles and theses of reports.

The head of the Commission,
Head of the Department of Electronic
Science and Technology, Professor

[Signature] Liubin

Liu Bin

Members of the Commission:

Deputy Director of the department, Associate
Professor

[Signature] Guangyu Jiang

Jiang Guangyu

Professor

[Signature]
Gong Yingqing

Gong Yingqing

“Затверджую”

Директор Державної установи
“Науковий центр аерокосмічних
досліджень Землі ІГН НАН України”
член-кореспондент НАН України

“17” жовтня 2022 р.



АКТ

впровадження результатів наукових досліджень аспірантки кафедри інформаційно-комунікаційних технологій ім. О.О. Зеленського Національного аерокосмічного університету ім. М. Є. Жуковського “ХАІ” **Лі Фангфанг**

Комісія в складі: голова – доктор технічних наук, професор Станкевич С. А., члени комісії – доктор технічних наук, с. н. с. Якимчук В. Г., кандидат технічних наук Хижняк А. В., склала цей акт про те, що такі наукові положення і результати, одержані Лі Фангфанг, а саме

- метод контролю та прогнозування рівня спотворень, внесених різними кодерами у зображення дистанційного зондування Землі, що дозволяє забезпечити бажану якість зображень, стиснутих із втратами;
- результати аналізу впливу втрат, внесених внаслідок стиснення, на ймовірність правильної класифікації багатоспектральних зображень та рекомендації щодо вибору параметрів стиснення, що дозволяє суттєво підвищити коефіцієнт стиснення без суттєвого погіршення надійності класифікації;

були використані при виконанні науково-дослідної роботи №: ПП-4-18 “Міждисциплінарний підхід в дистанційних аерокосмічних дослідженнях Землі та створення на його основі науково-методичної бази вирішення тематичних завдань природокористування” для підвищення ефективності вторинної обробки, візуалізації та аналізу багатоспектральних зображень в задачах дистанційного зондування Землі.

Голова комісії

Члени комісії

С. А. Станкевич

В. Г. Якимчук

А. В. Хижняк