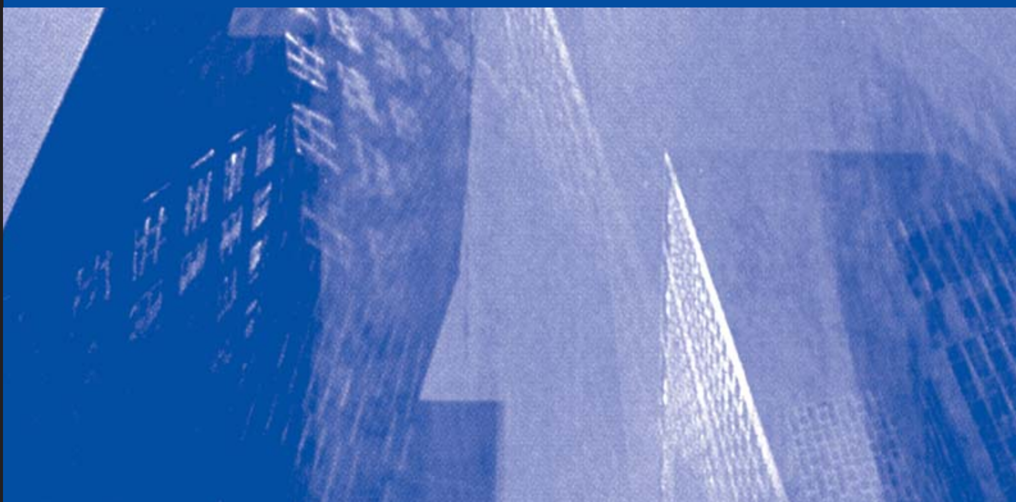
Advanced Topics in Science
and Technology in China 60

Yong Ding
Guangming Sun

Stereoscopic Image Quality Assessment



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

Advanced Topics in Science and Technology in China

Volume 60

Zhejiang University is one of the leading universities in China. In Advanced Topics in Science and Technology in China, Zhejiang University Press and Springer jointly publish monographs by Chinese scholars and professors, as well as invited authors and editors from abroad who are outstanding experts and scholars in their fields. This series will be of interest to researchers, lecturers, and graduate students alike.

Advanced Topics in Science and Technology in China aims to present the latest and most cutting-edge theories, techniques, and methodologies in various research areas in China. It covers all disciplines in the fields of natural science and technology, including but not limited to, computer science, materials science, the life sciences, engineering, environmental sciences, mathematics, and physics.

This book series is indexed by the SCOPUS database.

If you are interested in publishing your book in the series, please contact Dr. Mengchu Huang(Email: mengchu.huang@springer.com).

More information about this series at http://www.springer.com/series/7887

Yong Ding · Guangming Sun

Stereoscopic Image Quality Assessment

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

Yong Ding
College of Information Science
and Electronic Engineering
Zhejiang University
Hangzhou, Zhejiang, China

Guangming Sun
College of Information Science
and Electronic Engineering
Zhejiang University
Hangzhou, Zhejiang, China

ISSN 1995-6819 ISSN 1995-6827 (electronic)
Advanced Topics in Science and Technology in China
ISBN 978-981-15-7763-5 ISBN 978-981-15-7764-2 (eBook)
https://doi.org/10.1007/978-981-15-7764-2

Jointly published with Zhejiang University Press, China
The print edition is not for sale in China (Mainland). Customers from China (Mainland) please order the print book from: Zhejiang University Press.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

With the rapid development of digital image and video acquisition, transmission and display techniques in the last few decades, the demands of high-quality images and videos are growing amazingly fast in both people's everyday lives and specific application scenarios such as the fields of academy and engineering. Over the past years, the wave of stereoscopic display technology has an exponential increase, which can be mainly witnessed by the huge number of high-quality three-dimensional (3D) films and 3D TV in large-scale application. Meanwhile, mobile phones are also expected to be the largest 3D display application in the near future. There is no denial that the rapid development of stereoscopic technology has significantly enriched the way people perceive the world psychologically. Nevertheless, during capturing, coding, transmitting, processing and displaying of stereoscopic images, the distortion and interference introduced from outside are inevitable and unneglectable, which lead to the decrease of image quality and great visual discomfort. Thus, it is highly necessary to design effective methods to evaluate the perceptual quality of images, which is of vital importance for the performance optimization of image processing systems.

However, due to the special characteristics of stereoscopic images that are distinguished from two-dimensional (2D) images, for example, complex and non-intuitive interactions between multiple 3D visual cues, including depth perception, visual comfort and binocular characteristics such as binocular fusion and rivalry, automatically assessing the quality of stereoscopic images is still a challenging issue.

In recent years, although a large number of experimental studies on stereoscopic image quality assessment have been performed and various factors that affect stereoscopic perception have been investigated, it is still a puzzle in fully understanding the neural mechanism of visual cortex about how the human brain perceives and deals with stereoscopic natural image. As a result, studying upon stereoscopic image quality assessment is attracting more and more attention and a significant progress has been witnessed.

This book attempts to discuss the related topics about stereoscopic image quality assessment thoroughly and systematically. Firstly, the difference between 2D and stereoscopic image quality assessment is given. Secondly, the research of stereoscopic image quality assessment is discussed and analysed detailly, including a straightforward way based on existing 2D methods, a perceptual way based on human visual system properties and a new trend making use of deep learning models. Finally, some emerging challenges are described, and meanwhile a few new directions and trends are explored that are worth further investigations and research.

The authors would give particular thanks to Xiaoshu Xu who has made a significant contribution to the publication of this book. Moreover, the authors would express deep appreciation to the other students of Prof. Ding who have contributed to the research works presented in this book. For example, Ruizhe Deng, Yang Zhao, Xiaogang Xu, Zijin Gu, et al. have made their great efforts in researching on stereoscopic image quality assessment for the last several years.

Besides, the authors have received generous assistance and support from many of our colleagues including valuable information and materials used in this book, discussions, feedback, comments on and proofreading of various parts of the book, recommendations and suggestions that shaped the book as it is.

Due to our limited knowledge and energy, there inevitably exist some ambiguous interpretations and even mistakes in this book, which we welcome the readers and colleagues to point out.

Hangzhou, China Yong Ding
June 2020

Contents

Chapter 1
Introduction

Abstract Nowadays, objective image quality assessment (IQA) plays an important role for performance evaluation of image/video processing systems. Over the past few years, a variety of IQA methods have been introduced and they can be divided into three categories: full-reference IQA, reduced-reference IQA and no-reference IQA. All of these methods are clarified in detail in this book. In this chapter, the overall structure of the book is explained briefly and a summary of each of the following chapters is also provided.

Keywords Image quality assessment · Performance evaluation · Stereoscopic image

The idiom saying "A picture us worth a thousand words" has illustrated the importance of visual information perceived by human beings from nature images. With the rapid development of image store and display technologies, the image applications are widely used in human daily life, including entertainment, communications, security, monitoring and medical treatment fields (Wang and Bovik 2009; Karam et al. 2009). Perceptual quality of images plays an essential role in human perceiving visual information by human brain. Unfortunately, various distortions could be introduced during image transmission, compression, encoding and decoding, leading to images suffering from potentially substantial loss (Larson and Chandler 2010). Therefore, it is urgent demand to monitor and evaluate quality of images in real-time, which is called image quality assessment (IQA). In the past few years, there are a number of studies on IQA research community, and several of them have achieved promising results. Recently, three-dimensional (3D) media applications like virtual reality (VR) and 3D television (3D-TV) have been invented to improve the quality of human life, which derives consumers' interests for high-quality image contents of stereo images instead of plane images (Shen et al. 2018). How to explore the perceptual quality

Y. Ding and G. Sun, *Stereoscopic Image Quality Assessment*,
Advanced Topics in Science and Technology in China 60,
https://doi.org/10.1007/978-981-15-7764-2_1

of stereo images is becoming the focus of research in recent years (Moorthy et al. 2013).

Since nature images are perceived by human eyes and processed in human brain, the quality score of distorted images can be obtained in subjective test accurately. However, it is time-consuming and laborious, and cannot realized in real time in uncontrolled environments (Wang et al. 2004). It triggers the urgent demand for developing real-time reliable objective IQA methods to explore the quality of images, whose predicted quality scores are expected to be consistent with subjective perceptual scores by human eyes.

Further, objective IQA methods can be classified into three fields, full-reference (FR), reduced-reference (RR) and no-reference (NR), in which the classification criterion is based on the principle whether reference image information is involved in IQA research. When providing original reference images for comparison, viewers or objective algorithms can better explore the perceptual quality of images, which is called FR IQA. In contrast, NR IQA is defined by predicting quality scores of distorted images without any corresponding reference image information. As a trade-off between FR and NR IQA, RR IQA is conducted with the assistance of reference images partly. Due to the unavailability of reference images in practice applications, the research of NR IQA is the most valuable and challenging among the three.

According to the different inputs, IQA also can be divided into two research directions: 2D IQA for plane images and 3D IQA (also named SIQA) for stereopairs. In the past few years, significant progress has been achieved in IQA research of both 2D and 3D images. Since the main target of this book is to evaluate the quality of stereoscopic images, so most of chapters focus on 3D IQA research community. Of course, because 2D IQA is fundamental of exploring SIQA, we give briefly an overview about IQA research for plane images before introducing 3D IQA research, and interested readers can refer to relevant books and papers.

This book provides a comprehensive survey covering major developments in the field of SIQA. Firstly, the development of SIQA research community is introduced systematically. Then more detailed discussions and analysis in particular methods are presented in each chapter, respectively. Finally, some emerging challenges are described when conducting the research on 3D IQA, and meanwhile a few new directions and trends are explored and discussed that are worth further investigations and research in the future.

For each chapter, we assay progress focusing on a particular way to assess the quality of stereopairs. Each chapter begins with a brief introduction related to the chapter title, also includes an overview of recent significant progress in such aspect of SIQA research, hoping to help readers better understand. Besides, numerous references are given at the end of each chapter, which are from recent classical relative works.

Chapter 2 gives a comprehensive overview of IQA for plane images briefly. The research of 2D IQA is fundamental of exploring 3D images. It is necessary to understand relative knowledges about IQA before introducing SIQA research formally. Therefore, this chapter introduces some of the IQA information necessary for the transition to 3D IQA research community, including public subjective 2D IQA databases and well-known IQA methods (Wang et al. 2004; Winkler 2012). Of course, not all aspects of such a large subject about 2D IQA can be covered only in one single chapter. Unfortunately, limited of space, more detailed information in 2D IQA refer to relative books and papers, where we recommend reading the book titled with "Visual quality assessment for natural and medical image" (Ding 2018).

Chapter 3 formally introduces the conception of SIQA, including subjective ratings and public SIQA databases. In addition, the different visual information with the plane image, called binocular vision (Fahle 1987; Howard and Rogers 1995), is also discussed in this chapter briefly, which can give a illustrate understanding for readers.

Chapter 4 provides the general framework of SIQA based on 2D IQA models weighting. Directly extending some well-known 2D IQA models into SIQA tasks is a preliminary exploration for assessing the quality of stereopairs. In later research, as one of most important 3D visual factors, disparity map is also considered to assist quality prediction for stereoscopic images. Perhaps such a framework is not perfect from today's perspective, but it is necessary for us to learn and understand the contents of this chapter as a pioneer of SIQA research.

Chapter 5 illustrates the binocular vision caused by stereo images in detail. When the two views of stereopairs are combined into stereopsis, binocular visual properties (e.g., depth perception, binocular rivalry and visual discomfort) will occur, especially for asymmetrically distorted stereopairs. Deep analysis about binocular vison is given in this chapter briefly, and some relative works on SIQA fields considering binocular properties are introduced and discussed subsequently.

Chapter 6 focus on the importance of human visual system (HVS) on image quality evaluation. Since the HVS is the receiver of visual information for nature images, simulating the properties of HVS is a meaningful and valid for prediction performance improvements (Krüger et al. 2013). There are many visual properties in HVS that have been explored in previous research, from which visual saliency (Zhang et al. 2014) and just noticeable difference (Legras et al. 2004) are selected to discuss in detail in this chapter, respectively.

Chapter 7 gives a new trend for SIQA research community. Recently, deep learning has applied into many image processing tasks and achieved promising results than before. Many researchers begin to attempt to employ convolutional neural networks (CNNs) into SIQA fields, expecting CNN can automatically learning visual representations related with image quality rather than using hand-crafted visual features. However, the biggest obstacle, inadequate training data, need to be addressed before designing more complex deep learning models. There are many strategies applied in CNN-based SIQA models for alleviating the problem, including patch-wise (Zhang et al. 2016; Oh et al. 2017), transfer learning (Ding et al. 2018; Xu et al. 2019) and extending datasets (Liu et al. 2017; Dendi et al. 2019), which derive a series of SIQA models using CNN architectures.

Chapter 8 gives a summary of SIQA research community described in previous chapters, and meanwhile discusses the challenge issues and new trends of stereoscopic image quality assessment in the future.

The book is intended for researchers, engineers as well as graduate students working on related fields including imaging, displaying and image processing, especially for those who are interested in the research of SIQA. It is believed that the review and presentation of the latest advancements, challenges, and new trends in the stereoscopic image quality assessment will be helpful to the researchers and readers of this book.

References

Dendi SVR, Dev C, Kothari N, Channappayya SS (2019) Generating image distortion maps using convolutional autoencoders with application to no reference image quality assessment. IEEE Signal Process Lett 26(1):89–93

Ding Y (2018) Visual quality assessment for natural and medical image. Springer. https://doi.org/10.1007/978-3-662-56497-4

Ding Y, Deng R, Xie X, Xu X, Chen X et al (2018) No-reference stereoscopic image quality assessment using convolutional neural network for adaptive feature extraction. IEEE Access 6:37595–37603

Fahle M (1987) Two eyes, for what? Naturwissenchaften 74(8):383–385

Howard IP, Roger BJ (1995) Binocular vision and stereopsis. Oxford University Press, New York

Karam LJ, Ebrahimi T, Hemami SS, Pappas TN (2009) Introduction to the issue on visual media quality assessment. IEEE J Sel Topics Signal Process 3(2):189–190

Krüger N, Janssen P, Kalkan S, Lappe M, Leonardis A et al (2013) Deep hierarchies in the primate visual cortex: what can we learn for computer vision? IEEE Trans Pattern Anal Mach Intell 35(8):1847–1871

Larson EC, Chandler DM (2010) Most apparent distortion: full-reference image quality assessment and the role of strategy. J Electron Imaging 19(1):1–21

Legras R, Chanteau N, Charman WN (2004) Assessment of just-noticeable differences for refractive errors and spherical aberration using visual simulation. Optom Vis Sci 81(9):718–728

Liu X, van de Weijer J, Bagdanov AD (2017) RankIQA: learning from rankings for no-reference image quality assessment. International Conference on Computer Vision, Venice, Italy, pp 1040–1049

Moorthy AK, Su CC, Mittal A, Bovik AC (2013) Subjective evaluation of stereoscopic image quality. Sig Process Image Commun 28(8):870–883

Oh H, Ahn S, Kim J, Lee S (2017) Blind deep S3D image quality evaluation via local to global feature aggregation. IEEE Trans Image Process 26(10):4923–4935

Shen L, Li K, Feng G, An P, Liu Z (2018) Efficient intra mode selection for depth-map coding utilizing spatiotemporal, inter-view correlations in 3D-HEVC. IEEE Trans Image Process 27(9):4195–4206

Wang Z, Bovik AC (2009) Mean squared error: love it or leave it?—A new look at signal fidelity measures. IEEE Signal Process Mag 1:98–117

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

Winkler S (2012) Analysis of public image and video databases for quality assessment. IEEE J Sel Topics Signal Process 6(6):616–625

Xu X, Shi B, Gu Z, Deng R, Chen X et al (2019) 3D no-reference image quality assessment via transfer learning and saliency-guided feature consolidation. IEEE Access 7:85286–85297

Zhang L, Shen Y, Li H (2014) VSI: a visual saliency-induced index for perceptual image quality assessment. IEEE Trans Image Process 23(10):4270–4281

Zhang W, Qu C, Ma L, Guan J, Huang R (2016) Learning structure of stereoscopic image for no-reference quality assessment with convolutional neural network. Pattern Recognit 59:176–187

Chapter 2
Brief Introduction on 2D Image Quality Assessment

Abstract In this chapter, a brief introduction about 2D image quality assessment is given. Firstly, some public image quality databases are introduced which provide ground-truth information for training, testing and benchmarking. Secondly, IQA performance metrics including SROCC, KROCC, PLCC and RMSE to compare the accuracy of different IQA methods are provided. Finally, the general frameworks of 2D IQA methods containing full-reference (FR), reduced-reference (RR) and no-reference (NR) are illustrated based on specific algorithms.

Keywords 2D Image Quality Assessment · Databases · Correlation coefficient

2.1 Introduction

Visual perception information is an indispensable part of our daily life. With the development of multimedia display and transmission technology, people can obtain a lot of high-definition pictures through mobile phones, laptops, tablet computers and other electrical devices at any time. However, as an important medium of carrying information, image is inevitably polluted in the process of acquisition, reproduction, compression, storage, transmission or restoration, which finally could result in quality degradation. Therefore, objective image quality assessment (IQA) has become one of the focuses of people's research.

To begin with, researchers recognized the fact that the most reliable IQA method is human subjective judgment. Since human beings are the ultimate receivers of the visual information, the results of subjective judgement are considered to be the most accurate and reliable for perceiving the quality of images. However, directly utilizing observers to make subjective judgments on image quality is time-consuming and laborious, which is difficult to apply in real-time image processing systems. Therefore, designing an objective quality assessment algorithm to correlate with the results of subjective judgement is the mainstream in the research field of IQA. Recently, towards advancing progress on objective IQA research, a large number of classic and state-of-the-art IQA algorithms have been invented. In order to evaluate the accuracy of an IQA algorithm, ground-truth information should be obtained for training,

Y. Ding and G. Sun, *Stereoscopic Image Quality Assessment*,
Advanced Topics in Science and Technology in China 60,
https://doi.org/10.1007/978-981-15-7764-2_2

testing and benchmarking. Ground-truth can also be recognized as image quality database in the field of IQA. A classic IQA database consists of a set of reference images and its corresponding distorted images. In addition, the most important part of IQA database is the subjective quality ratings of distorted images obtained by subjective judgement. The concept of IQA database is initiated by Video Quality Experts Group to evaluate the performance of the IQA metrics. Therefore, in this chapter, we will first briefly introduce the main publicly available image quality databases in the recent years.

Generally, the performance of an IQA algorithm is analyzed and evaluated by statistical methods. Similarity analysis often refers to solving the correlation coefficient between each objective algorithmic score and subjective (differential) mean opinion score (DMOS/MOS), so as to compare the prediction accuracy of each objective IQA algorithm. The higher the correlation, the better the performance, and the lower the correlation, the worse the performance. The most commonly used correlation coefficients include Spearman Rank-Order Correlation Coefficient (SROCC), Kendall Rank-Order Correlation Coefficient (KROCC), Pearson Product-Moment Correlation Coefficient (PLCC) and Root Mean Squared Error (RMSE). In general, PLCC and RMSE are typically used to represent the prediction accuracy and consistency, while SROCC and KROCC can be regarded as a measure of the prediction monotonicity. Higher values of PLCC, SROCC and KROCC and lower values of RMSE correspond to higher performance. The more information of the four correlation coefficients will be given in the next section.

In the next part of this chapter, we will give a brief introduction about the general framework of modern IQA methods. Generally speaking, researchers always divide IQA methods into three categories, full-reference (FR), reduced-reference (RR), and no-reference (NR). As the term suggests, this classification is based on the participation of reference during IQA operation (Wang and Bovik 2006). We define an image of undistorted version as its reference image (or original image). FR IQA methods mean that viewers can obtain more information from reference images and the quality of distorted image can be calculated by comparing the local similarity between the reference and distorted images. In contrast, NR IQA methods predict the perceptual quality of a distorted image without any information assistance of original images. The implementation of RR IQA method requires the assistance of some information from the reference image. For IQA methods belonging to different classifications, their frameworks will be different to some extent. We will spend the rest of this chapter explaining each of them. On the other hand, their basic principles are exactly the same. This means that no matter which type of algorithms we are proposing, it is inevitably required to extract quality-aware features from images, then quantify the features and map the results to the final IQA scores. For three different frameworks, we will each propose a classical algorithm to illustrate.

2.1.1 Public Image Quality Databases

Image quality database is particularly significant for objective IQA methods. The development of the former plays an important role in promoting the latter. Previous IQA methods usually performs well on a specific IQA database, but achieves poor performance on other databases. However, a qualified IQA method should not be limited to a specific IQA database. In another words, it needs to have good generalization ability. Therefore, different IQA databases are urgently needed for the research of IQA. This section summarizes several well-known public databases, such as LIVE, TID2008, TID2013, MICT, IVC, A56, LAR, WIQ, DRIQ and so on. The following gives a brief introduction to these mentioned IQA databases.

The entire Laboratory for Image and Video Engineering (LIVE) database developed by the University of Texas at Austin. LIVE includes twenty-nine high-resolution and quality color reference images, which were collected from the Internet and photographic CD-ROMS and their distorted images. The distortion types include JPEG, JEPG2000, Gaussian white noise (WN), Gaussian blur (GB) and fast fading (FF), and each distortion type is accompanied by different degrees of distortion. The total number of distorted images is 779, including 175 JPEG distorted images, 169 JPEG2000 distorted images, 145 WN distorted images, 145 Blur distorted images, 145 FF distorted images. LIVE database also provides DMOS value for each image as subjective quality score, where the higher value demonstrates the worse quality of the image.

Categorical Subjective image quality database includes 30 reference images and 866 corresponding distorted images. The distortion types include JPEG, JP2K, Additive Gaussian white noise, Additive pink Gaussian noise, Gaussian blur and Global contrast decrements (Larson and Chandler 2010).

Tampere Image Quality (TID 2008) database is created by 25 reference images from the Tampere University of Technology, Finland. The 25 reference images are obtained from the Kodark Lossless True Color Image Suite, among which the first 24 are natural images, and the 25th is computer-generated images. By introducing 17 distorted types, 1700 test images are generated, and its 17 distortion types are AGN (Additive Gaussian noise), ANC (Additive Gaussian noise in color components), SCN (Spatially correlated noise), MN (Masked noise), HFN(High frequency noise), IN (Impulse noise), QN (Quantization noise), GB (Gaussian blur), DEN (Denoising), JPEG (JPEG compression), JP2K (JP2K compression), JGTE (JPEG transmission errors), J2TE (JP2K transmission errors), NEPN (Non eccentricity pattern noise), Block (block-wise distortions of different intensity), MS (Mean shift) and CTC (Contrast change).

TID 2013 database (Ponomarenko et al. 2013) includes 3000 distorted images, and adds 7 types of distortion based on the TID 2008, which are CCS (Change of color saturation), MGN (Multiplicative Gaussian noise), CN (Comfort noise), LCNI (Lossy compression of noisy images), ICQD (Image color quantization with dither), CHA (Chromatic aberrations) and SSR (Sparse sampling and reconstruction). The