Electronic Theses and Dissertations

2015

Investigating Potential Combinations of Visual Features towards Improvement of Full-Reference and No-Reference Image Quality Assessment

Ashirbani Saha
University of Windsor

INVESTIGATING POTENTIAL COMBINATIONS OF VISUAL FEATURES
TOWARDS IMPROVEMENT OF FULL-REFERENCE AND NO-REFERENCE
IMAGE QUALITY ASSESSMENT

by

Ashirbani Saha

A Dissertation
Submitted to the Faculty of Graduate Studies
through the Department Electrical and Computer Engineering
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
at the University of Windsor

Windsor, Ontario, Canada

2015

Investigating Potential Combinations of Visual Features towards Improvement of
Full-Reference and No-Reference Image Quality Assessment

by
Ashirbani Saha

APPROVED BY:

Dr. T. D. Bui,
Department of Computer Science and Software Engineering,
Concordia University

Dr. I. Ahmad,
School of Computer Science

Dr. M. Ahmadi,
Electrical and Computer Engineering

Dr. N. C. Kar,
Electrical and Computer Engineering

Dr. Q. M. Jonathan Wu, Advisor,
Electrical and Computer Engineering

February 9, 2015

Declaration of Co-Authorship/Previous Publications

I. Co-Authorship Declaration

I hereby declare that this dissertation incorporates material that is result of joint research, as follows:

This dissertation also incorporates the outcome of research undertaken under the supervision of and in collaboration with professor Dr. Q. M. Jonathan Wu. The collaboration is covered through Chapters 3, 4, 5 and 6 of the dissertation. In all cases, the key ideas, primary contributions, experimental designs, data analysis and interpretation, were performed by the author, and the contribution of the collaborator was primarily through the provision of valuable suggestions for the representation of ideas, for the analysis of the results for the experiments carried out and editorial activities throughout the process of dissemination of the work.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my dissertation, and have obtained written permission from the only co-author to include the above material in my dissertation.

I certify that, with the above qualification, this dissertation, and the research to which it refers, is the product of my own work.

II. Declaration of Previous Publication

This dissertation includes 5 original papers that have been previously published/submitted for publication in peer reviewed conference proceedings and journals, as follows:

Chapter	Publication title/full citation	Publication status
Chapter 3	A. Saha and Q. M. Jonathan Wu, "Perceptual Image Quality Assessment Using Phase Deviation Sensitive Energy Features Elsevier Signal Processing, 93(11), 3182-3191, 2013.	Published
	A. Saha and Q. M. Jonathan Wu, "A Study on Using Spectral Saliency Detection Approaches for Image Quality Assessment", in proceedings of International Conference on Acoustics, Speech and Signal Processing(ICASSP), 2013, pp.1889-1893, Vancouver, Canada.	Published
Chapter 4	A. Saha and Q. M. Jonathan Wu, "Full-Reference Image Quality Assessment by Combining Global and Local Distortion Measures", 2013	Submitted
Chapter 5	A. Saha and Q. M. Jonathan Wu, "High Frequency Content based Stimulus for Perceptual Sharpness Assessment in Natural Images", 2014	Submitted
Chapter 6	A. Saha and Q. M. Jonathan Wu, "Utilizing image scales towards totally training free blind image quality assessment", IEEE Trans. on Image Processing, 2015	Accepted with minor revisions

I certify that I have obtained a written permission from the copyright owner(s) to include the above published material(s) in my dissertation. I certify that the above

material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my dissertation does not infringe upon anyones copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my dissertation, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my dissertation.

I declare that this is a true copy of my dissertation, including any final revisions, as approved by my dissertation committee and the Graduate Studies office, and that this dissertation has not been submitted for a higher degree to any other University or Institution.

Abstract

Objective assessment of image quality is the process of automatic assignment of a scalar score to an image such that the rating or score corresponds to the score provided by the Human Visual System (HVS). Despite extensive studies since the last two decades, it remains a challenging problem in image processing due to the presence of different types of distortions and limited knowledge of the HVS. Existing approaches for assessing the perceptual quality of images have relied on a number of methodologies that directly apply known properties of the HVS, construct hypotheses considering the HVS as a blackbox and use hybrid approaches that apply both of the techniques. All of these methodologies have relied on different types of visual features for Image Quality Assessment (IQA).

In this dissertation, we have studied the problem of different types of IQA from the feature extraction point of view and showed that effective combinations of simple visual features can be used to develop IQA approaches having competitive performance with the state-of-the-art. Our work is divided into four parts each having the final goal to bring about performance improvement in the areas of Full-Reference (FR) and No-Reference (NR)-IQA. We have gradually moved from FR to NR-IQA in the works presented in this dissertation. First, we propose improvements in two existing FR-IQA techniques by introducing changes in the features used. Next, we propose a new FR-IQA technique by extracting image saliency as global features and combining them with the local features of gradient and variance to improve the performance. For NR-IQA, we propose a novel technique for sharpness detection in natural images using simple features. The performance of this method provides improvement over the existing methods. After working with the specific purpose NR-IQA, we propose a general purpose technique using suitable features such that no training with pristine or distorted images or subjective quality scores is required. This technique, despite

having no reliance on training, provides competitive performance with the state-of-the-art techniques. The main contribution of the dissertation lies in identification and analysis of effective features and their combinations for improving three different sub-areas of IQA.

Dedication

to

the loving memory,

of

my aunt Jyotsna Gayen

and

my grandparents

Sneha Lata and Pasupati Nath Saha

and

Rekha Rani and Sital Ranjan Gayen

Acknowledgements

I would like to thank my supervisor Dr. Q. M. Jonathan Wu for all his support during the entire program, patience and belief in my abilities. I thank my committee members Dr. M. Ahmadi, Dr. N. C. Kar and Dr. I. Ahmad for their valuable suggestions and insightfulness. I thank the external examiner Dr. T. D. Bui for all his suggestions and help in improving the quality of the dissertation. I would extend my gratitude towards Dr. B. Shahrrava for his guidance during my coursework with him. I would like to take this opportunity to express my gratitude towards my mentors whom I met during my high school and undergraduate studies for motivating me to carry on the pursuit of knowledge. As learning is a never-ending process that permeates throughout the life, I humbly offer my gratitude to mother nature for making me a part of it and for helping me to realize it.

I would like to thank our departmental graduate secretary, Ms. Andria Ballo for all her help during my studies at the University of Windsor. I thank the Student Health Services at the University of Windsor for helping me to learn a better lifestyle during these strenuous years. My colleagues and friends in the CVSS Laboratory have been helpful, supportive and source of several good memories I have gathered during my studies. I thank the city of Windsor for the friendly environment it offered and for being my home away from home.

Back in India, my family has been with me always with unconditional love and support. I express my love and respect for my father who always taught me to aim higher. My mother has been my support throughout my life in every aspect and put stupendous effort into my upbringing. I thank God for blessing me with such wonderful parents. I thank my loving elder sister and her family for being there beside my parents and taking care of them. My uncle and aunts have also motivated me with their determination and successful careers. My father-in-law has been an

embodiment of diligence and inspired me to question my efforts. At the same time, his and my mother-in-law's optimism and faith in my abilities have been important sources of my motivation.

My husband, Dibyendu, has stood by me through thick and thin, and has emboldened my spirits from time to time towards chasing my dreams. Besides writing his own dissertation, he made enormous efforts to tie all the loose ends of our lives and to maintain a positive attitude towards it.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADM	Additive and Detailed Loss based Measure
apgn	additive pink Gaussian noise
awgn	additive white Gaussian noise
awgn-pink	more intense additive Gaussian noise in color channels compared to luminance channels
BIQA	Blind Image Quality Assessment
BIQES	Blind Image Quality Evaluator based on Scales
BIQI	Blind Image Quality Index
blockwise-distortion	local block-wise distortions of different intensities
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator
CBIQ	Codebook based Image Quality
CORNIA	Codebook Representation for No-Reference Image Assessment
CSF	Contrast Sensitivity Function
CSIQ	Categorical Subject Image Quality
DCT	Discrete Cosine Transform
DMOS	Differential Mean Opinion Score
DSCQS	Double Stimulus Continuous Quality Scale

DSIS	Double Stimulus Impairment Scale
DWT	Discrete Wavelet Transform
Eigen SR	Eigen Spectral Residual
Eigen PQFT	Eigen Phase of Quaternion Fourier Transform
FR	Full-Reference
FT	Fourier Transform
FSIM$_c$	Feature-Similarity Index for Color Images
FSIM	Feature-Similarity Index
gblur	Gaussian blur
global-contrast	global contrast decrement
GSM	Gaussian Scale Mixture
GSSIM	Gradient based Structural Similarity
HFT	Hypercomplex Fourier Transform
HPF	High Pass Filter
HPFSM	High Pass Filter based Sharpness Measure
HVS	Human Visual System
IFC	Information Fidelity Criterion
IGM	Internal Generative Mechanism
IQA	Image Quality Assessment
ITU	International Telecommunication Union

IVC	IRCCyN/IVC Image Quality Database
IWSSIM	Information-Content Weighted Structural Similarity Index
JNB	Just Noticeable Blur
jpeg	JPEG compression
jpeg2k	JPEG2000 compression
jpeg2k-trans-error	JPEG2000 transmission errors
jpeg-trans-error	JPEG transmission errors
KROCC	Kendall's Rank-Order Correlation Coefficient
LIVE	Laboratory for Image and Video Engineering
LPC-SI	Local Phase Coherence Sharpness Index
LQF	Latent Quality Factors
MAD	Most Apparent Distortion
MAE	Mean Absolute Error
MICT	Multimedia Information and Communication Technology
MOS	Mean Opinion Score
MSSSIM	Multiscale Structural Similarity Index
MVG	Multivariate Gaussian Model
NIQE	Natural Image Quality Evaluator
NR	No-Reference
NRQI	No-Reference Quality Index

NSS	Natural Scene Statistics
pattern-noise	non eccentricity pattern noise
PC	Phase Congruency
PDSESIM	Phase Deviation Sensitive Energy Features based Similarity
PFT	Phase Fourier Transform
PLCC	Pearson's Linear Correlation Coefficient
PQFT	Phase of Quaternion Fourier Transform
PS	Programme Segment
PSF	Point Spread Function
PSNR	Peak-Signal-to-Noise Ratio
QAC	Quality Aware Clustering
QDCT	Quaternion Discrete Fourier Transform
QFT	Quaternion Fourier Transform
RMS	Root Mean Squared
RMSE	Root Mean Squared Error
RR	Reduced-Reference
spatial-corr-noise	spatially correlated noise
SR	Spectral Residual
SRNSS	Sparse Representation of Natural Scene Statistics
SROCC	Spearman's Rank-Order Correlation Coefficient

SR-SIM	Spectral Residual based Similarity
SSCQE	Single Stimulus Continuous Quality Evaluation
SSIM	Structural Similarity Index
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVR	Support Vector Regressor
TC	Test Condition
TID2008	Tampere Image Quality Database 2008
TS	Test Session
UWT	Undecimated Wavelet Transform
UWTSM	Undecimated Wavelet Transform based Sharpness Measure
VIF	Visual Information Fidelity
VQEG	Video Quality Experts Group
VSNR	Visual-Signal-to-Noise Ratio
wn	white noise

Chapter 1

Introduction

The work presented in this dissertation is related to the investigation of improvements in full-reference and no-reference image quality assessment using feature combinations. Before delving into the details of the work proposed here, we discuss some preliminaries related to the work along with the motivation and objective of the proposed research.

1.1 What is Image Quality Assessment?

The set of words 'Image Quality Assessment', when taken literally, amounts to the evaluation of certain attributes of images. However, the phrase has a restricted and specific meaning when studied under the context of modern image processing. Image Quality Assessment (IQA) refers to the automatic evaluation of the quality of digital images in keeping with the judgment of humans. Quality of any image refers to its collective attributes that determine its appearance and ability to convey information to the human observer. Hence, this quality is purely subjective in nature. Thus, IQA techniques aim to automatically predict the image quality by using mathematical and computational models in keeping with the human observers. Many a times, the phrase 'objective image quality assessment' is used specifically to differentiate the process of automatic evaluation of perceptual image quality from the subjective evaluation of image quality. The subjective evaluation of image quality is necessary for several image processing applications as human beings are the ultimate users and hence judges of the attributes of an image. We present a discussion on the subjective

and objective methods of image quality assessment to understand their respective needs towards improving image processing methodologies. In the remaining part of this dissertation, the image assessed for quality evaluation is termed as a test or query or distorted image. The pristine quality counterpart of the test image, if exists, is called the reference image. The same applies when image sequences are discussed instead of images.

1.2 Subjective Methods of Image Quality Assessment

Subjective evaluation is considered to be the most reliable way of assessing the image quality [111]. However, any human observer has his own perception of quality which may differ from that of other observers. Hence, subjective methods employ the opinion of many viewers to evaluate the quality of any image. The observers need to rate any image under identical viewing conditions and also need to have a pre-defined rating scale to maintain uniformity throughout the evaluation process. Hence, International Telecommunication Union (ITU) has recommended some standards for carrying out subjective tests for quality assessment of images. The basic set of recommendations is provided in ITU-R Rec.BT.500 [26] through the work of Video Quality Experts Group (VQEG). Though the recommendations are presented for videos, they are equally applicable to images as videos consist of sequence of images. Depending on the type or variations in the stimulus presented to the observers and scale of ratings provided, the three commonly used test methodologies are as follows:

1. **Double Stimulus Continuous Quality Scale (DSCQS)**: In this procedure, multiple sequence pairs comprising of reference and test sequences are displayed to the human observer twice. The duration of the sequences is restricted to the typical value of 10 seconds. However, the observer is not informed about the

order of appearance of the reference and the test images as the order is decided randomly in each trial. An example of such a trial is presented in Fig. 1.1(a). Here, S1 and S2 are two sequences forming a reference and test sequence pair, not necessarily in the same order as mentioned. They are displayed twice after which the observer provides ratings for each of S1 and S2 in a numerical continuous scale ranging from 0 to 100 as shown in Fig. 1.1(d). The difference of these two ratings is considered as the rating of the test sequence, as it is reasoned to remove the subjective bias caused by the observer experience and the sequence content.

2. **Double Stimulus Impairment Scale (DSIS)**: Though this is also a double stimulus procedure, the presentation of stimulus varies in two different ways from that in DSCQS. Firstly, the reference sequence is always presented before the test sequence. Secondly, there is no repetition of the same sequence pair. Each trial of this standard is shown in Fig. 1.1(b). The user rating is provided for the test sequence in a discrete five-level scale which is shown in Fig. 1.1(e). The scale assesses the visibility of degradation present in test images as compared with the reference image and classifies from 'imperceptible' to 'very annoying' as presented in the figure.

3. **Single Stimulus Continuous Quality Evaluation (SSCQE)**: As the name itself says single stimulus, only the test sequence can be displayed to the observer. However, the sequences are much longer in duration and this method enables the collection of observer ratings dynamically with the help of a slider. The ratings are provided in continuous scale. The stage 1 test protocol for this method comprises of set Test Session (TS) (without separation) which is at least 30 minutes in duration. Each TS is segregated to a Programme Segment (PS) (duration of at least 5 minutes) signifying a specific type of video content (for e.g., news, drama, sports etc.) and processed according to one quality param-
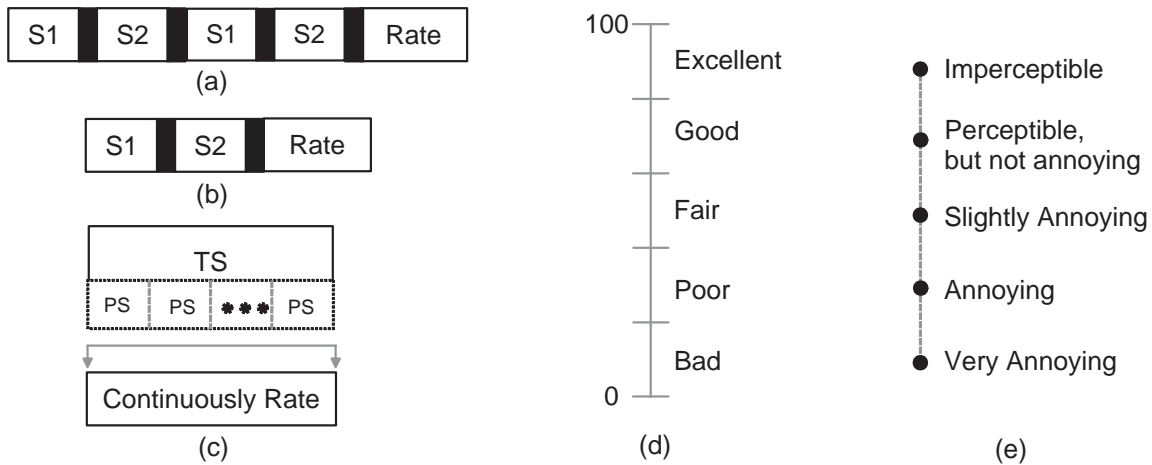
Figure 1.1: Trials in (a)DSCQS (b)DSIS and (c)SSCQE representing the variations in the stimuli. Continuous quality scale is presented in (d) and discrete scale is shown in (e).

eter or Test Condition (TC). Rating is provided by the user at a frequency of two ratings per second. The relevant diagram for this method is shown in Fig. 1.1(c).

Each of these procedures has different types of utilities that are suited to varying applications and suffers from limitations [120]. Since DSCQS techniques allow rapid evaluation of a test sequence with the help of reference after double viewing, the procedure works well to identify small differences in quality between the reference and test images. On the other hand, clearly visible degradations are better evaluated by DSIS. However, these methods rely on human memory and suffer from recency effect implying that more weight will be placed on whatever is retained by the memory. In this case, usually more weight is placed on the degradation seen in last 5-10 seconds [12]. Also, repetition and use of the reference image make the user familiar with the image content. Based on these drawbacks, SSCQE is developed and it finds application both in time-varying quality evaluation and aids instantaneous evalua-

tion. However, it has video content bias and is affected by the reaction times of the observers.

All of the test procedures discussed above need a minimum number of human observers to perform the related tasks. All the human observer ratings collected are then used to form the Mean Opinion Score (MOS) for each test sequence. However, in the updated report [107] by VQEG, Differential Mean Opinion Score (DMOS) scores, calculated using the difference between the rating for test and reference images, are used as subjective scores. Thus, the subjective methods require the support of human observers making it less economical and slow to be suited for real-world applications in image processing [111]. Therefore, objective methods for assessing the perceptual quality of images are needed for many real-world applications. However, for the performance evaluation of objective IQA methods, subjective scores in the form of MOS or DMOS are needed and subjective methods are applied to generate these scores in the existing databases.

1.3 Objective Methods of Image Quality Assessment

Objective methods of assessing image quality come up with a scalar score generated by computational models. The scalar score represents the quality of the test image. If the generation of the score requires the help of full information from the corresponding reference image, the method is called Full-Reference (FR). If no information is used from the corresponding reference image, the objective IQA technique is called No-Reference (NR). NR-IQA is also called Blind Image Quality Assessment (BIQA). Throughout this dissertation, these two terms will be used interchangeably. If partial information from the corresponding reference image is used to produce the objective quality score from the distorted image, the method is called Reduced-Reference (RR).

By partial information, we refer to some features (for example, mean, variance, edge information etc.) available from the reference image, instead of the availability of the reference image itself. Further details about this classification can be found in [111]. As the works proposed in this dissertation are related to FR and NR-IQA, we have devoted the Chapter 2 of our dissertation for the earlier works related to these. However, the evaluation methodology and databases used for FR and NR-IQA bear a lot of similarity. Hence, we discuss about the databases and evaluation methodologies in the current chapter. Also, the applications of objective IQA techniques are discussed in details.

1.3.1 Significance of IQA Databases

In general, the databases used for IQA contain pristine quality images which can be used as reference images along with the distorted images and corresponding MOS/DMOS values. They may also contain information about the type and intensity of degradation applied to the pristine image to generate the distorted image, the number of human observers employed, the raw scores provided by the observers and details about the subjective tests conducted to generate the MOS/DMOS values. Throughout our work, we have used all or relevant ones from a set of six benchmark databases which are most commonly used for the evaluation of objective IQA techniques. The details of these six databases are as follows:

1. **Laboratory for Image and Video Engineering (LIVE) Database** [94]:
   The database is developed by LIVE at the University of Texas at Austin, USA. It consists of 29 reference and 779 distorted color images rated by 29 subjects. Images degraded by five types of distortions namely, JPEG compression (jpeg), JPEG2000 compression (jpeg2k), white noise (wn), Gaussian blur (gblur), and JPEG2000 bit error generated by transmission through a simulated Rayleigh fading channel (also referred as fastfading) are present in this database. The

numbers of images with each of these degradations are 169, 175, 145, 145 and 145 respectively. These distorted images were generated with 5-6 levels of the corresponding distortion. The subjective scores are given in the form of DMOS and lie in the range 0-112.

2. **Categorical Subject Image Quality (CSIQ) Database** [42]: The database is developed by the Oklahoma State University, USA with the help of 35 subjects. It consists of 30 reference and 866 distorted color images. Images having six types of distortions namely, jpeg, jpeg2k, additive white Gaussian noise (awgn), gblur, additive pink Gaussian noise (apgn) (also called fnoise), and global contrast decrement (global-contrast) are present in the database. Every distortion type except the global contrast decrement (116 images) has 150 images each. 4-6 distortion levels were applied on the pristine images to generate them. The subjective scores are given in the form of DMOS and lie in the range between 0-1.

3. **Tampere Image Quality Database 2008 (TID2008)** [79]: The database is developed by Tampere University of Technology, Finland with the help of 838 human observers. It has 1700 distorted images generated from a set of 25 pristine color images. 17 types of distortions are applied with 4 different levels on each pristine image. The distortions present in the database are awgn, more intense additive Gaussian noise in color channels compared to luminance channels (awgn-pink), spatially correlated noise (spatial-corr-noise), masked noise, high frequency noise, impulse noise, quantization noise, gblur, denoising, jpeg, jpeg2k, JPEG transmission errors (jpeg-trans-error), JPEG2000 transmission errors (jpeg2k-trans-error), non eccentricity pattern noise (pattern-noise), local block-wise distortions of different intensities (blockwise-distortion), and contrast change. The subjective scores are given in the form of MOS and lie in the range between 0-8.

4. **A57 Database** [8]: This database is developed at the Cornell University, USA with the help of 15 human observers. It consists of 3 reference grayscale images and 54 distorted images generated by 6 distortions in three different levels. The distortions are jpeg, jpeg2k, awgn, jpeg2k using the dynamic contrast based quantization, gblur, and uniform quantization applied on the LH subband of a 5-level Wavelet decomposition of the pristine image. The subjective scores lie in the range between 0-1.

5. **Multimedia Information and Communication Technology (MICT) Database** [30]: The database is developed by MICT Laboratory at the University of Toyama, Japan with the help of 16 human observers. There are 14 reference color images and only two types of distortions jpeg and jpeg2k are applied at 6 different levels to each of these reference images to generate 168 distorted images (84 for each distortion). The subjective scores are given in the form of MOS and lie in the range 1-5.

6. **IRCCyN/IVC Image Quality Database** [4]: The database is developed by Institut de Recherche en Communications et Cybernétique de Nantes (IR-CCyN), France with the help of 16 human observers. It has 10 pristine quality images and 4 different types of distortions are applied at several levels to different images. jpeg, jpeg2k, LAR coding, and gblur are the applied distortions. In some images, jpeg is applied to all channels and to the luminance channel only for generating separate images. The total number of distorted images present in the database is 185. The subjective scores are given in the form of MOS and lie in the range between 1-5.

1.3.2   Techniques to Evaluate an IQA Method

To evaluate the performance of an objective IQA method, five evaluation measures for the quantitative analysis of the proposed method are used. These evaluation measures

are Spearman's Rank-Order Correlation Coefficient (SROCC), Kendall's Rank-Order Correlation Coefficient (KROCC), Pearson's Linear Correlation Coefficient (PLCC), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). SROCC and KROCC are non-parametric rank based correlation measures and useful for judging the *prediction monotonicity* of any IQA method. Higher values of SROCC and KROCC indicate that the objective assessment scores are more consistent with the corresponding MOS/DMOS values and therefore the assessment method is better. On the other hand, PLCC is a measure of *prediction accuracy* whereas MAE and RMSE are the measures of *prediction error* [107]. Therefore, higher values of PLCC and lower values of MAE and RMSE are desirable. For PLCC, MAE and RMSE, a 5-parameter logistic function [93] is used for mapping between the subjective and objective scores. The logistic function is given by

$$Predicted\ Quality(x) = \beta_1 \left[0.5 - \frac{1}{1 + \exp(\beta_2(x - \beta_3))} \right] + \beta_4 x + \beta_5, \qquad (1.1)$$

where $(\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ are the 5 parameters. As per [107], the logistic transform is used to bring the objective scores in a common ground with the DMOS/MOS by providing a non-linear mapping between them. However, this logistic function provides monotonic non-linear mapping. The measures SROCC and KROCC are based on magnitude-wise ranking of data values and hence do not depend on the mapping. SROCC has been widely used due to its simplicity. Hence, we have used SROCC for parameter selection and distortion wise performance. For analysis of the overall performance of the proposed method(s), SROCC, KROCC, and PLCC are used. This is because their absolute values lie between 0 and 1, irrespective of the ranges of subjective scores in different databases. In contrast, MAE and RMSE values are dependent on the ranges of the subjective scores given in the respective databases.

1.3.3 Applications

There are three major applications of IQA as mentioned in [111]. They include the monitoring of image quality in image processing applications, benchmarking of image processing systems and algorithms, and optimizing image processing systems and their parameter settings. Perceptual models of IQA and properties of Human Visual System (HVS) have often inspired the areas of digital watermarking [33, 39] and steganalysis [1] for better schemes. Visual quality assessment methods have found applicability in evaluation of image fusion [87], image and video coding [115, 123], evaluation of image denoising techniques [127] and perceptual feature based automatic parameter selection [67] for image denoising. The recent applications of IQA include evaluation of image deblurring [55] and image super resolution [126] as well. Stereoscopic image quality [24, 50] and medical image quality evaluation [5, 66] have also emerged as two application based research directions for IQA.

1.3.4 Challenges

When automatic assessment of image quality emerged as a research area in image processing, it was meant to attain an objective that was self-explanatory with a fixed set of problems. The objective is to perceptually evaluate different types of distortions as per the HVS. We have discussed about distortions in the databases already. Evaluating different types of distortions with the limited knowledge of HVS makes the job of image quality assessment very difficult. However, along with the research and development taking place in this field, we have more challenges to combat with in order to achieve a complete and practical solution. Thus, the earlier set of challenges are prevailing and getting augmented with the newly identified significant problems as various directions for potential solutions are revealed through different research works. The definition of quality of an image is itself dubious as the observer's cognitive ideas and the task assigned alter the results. Presence of prior information

regarding the pristine image and earlier ideas about the image content affect the user's decision [99, 111]. However, due to the similarity in visual processing strategies in humans, they can consistently judge, if a digital manipulation has enhanced or degraded the quality of any image. The models of HVS developed mainly incorporate the properties of V1 cortex of human brain. However, much of information from other parts of the brain has not been incorporated and is not available for the purpose of incorporation [7]. Also, the models are found simpler when compared to the complexity of HVS and natural images [111]. The existing state-of-the-arts IQA methods evaluate images degraded by a single type of distortion only. Also, there exists no FR-IQA technique that works better than all others methods across all distortions present in all databases. Hence, substantial improvement is needed to achieve this. Researchers on NR-IQA techniques are also trying to improve their performances amidst all the challenges prevailing.

1.4 Motivation

As discussed earlier, various application areas of IQA along with its challenging problems make it an interesting area to work with. This has invariably left a positive influence as well as the present surge in the development of different types of IQA methods and has created a diversity towards achieving the goal. In spite of the limited knowledge of HVS, the various methodologies that showed effectiveness towards the partial solution of the problem are indeed fascinating. All of these methodologies depend on different types of features and their combinations to evaluate perceptual quality images. However, there still exist several limitations and hence the need to improve. This has inspired us to explore, investigate and make contributions towards the field of IQA in general. The motivation to work on each selected sub-area of IQA is mentioned along with the discussion of related works presented in Chapter 2.

1.5 Objective and Scope

The effort culminated in this dissertation has been targeted towards FR-IQA and NR-IQA. The main objective of the dissertation is to know and explore the state-of-the-art techniques in the related areas and be able to contribute towards some of the existing problems by making improvements. By investigating on effective combination of features, we have tried to contribute towards three diversified channels of IQA. To begin with, during the initial years of the work we tried to make changes in some of the existing approaches. Next, we have worked towards developing a novel FR-IQA method based on diverse feature combinations and our understanding and interpretation of perceptual quality assessment. From there on, we have tried to work on NR-IQA using effective features. First, we focused on a distortion specific NR-IQA. The specific distortion chosen is absence of sharpness or presence of blur because it is one of the commonly occurring distortions. Then, we have moved on to the general purpose NR-IQA. Through this dissertation, we have made simultaneous efforts to gather knowledge and make contributions in the field of IQA. Our focus in this dissertation has remained confined within the limits of various types of IQA problems and different solution approaches to these problems using visual feature combinations.

1.6 Main Attributes

The major attributes of the dissertation are enlisted as follows:

- As an initial step in the field of IQA, we have studied and tried to improve the existing techniques Feature-Similarity Index (FSIM) [129] and Spectral Residual based Similarity (SR-SIM) [128] by using modification of various features. With the proposed modifications, the performance of the techniques has been enhanced.

- A methodology is proposed for FR-IQA using the combination of global and local features of distortions. This approach achieves improved performance in two databases as compared to the state-of-the-art techniques. The overall performance of the proposed method is also better than the state-of-the-art techniques.

- A methodology is proposed for blind evaluation of image sharpness using simple combination of features indicating high frequency content in images. The proposed technique improves over the performance of the state-of-the-art techniques in several databases.

- A general purpose technique for blind evaluation of image quality is proposed. The methodology is unique as it does not use training with any of the three data, namely, pristine images, distorted images and subjective score of quality and relies only on features. In spite of this, the approach achieves competitive performance with the state-of-the-art techniques in several databases.

1.7   Organization

The work presented in the dissertation has been segmented into several chapters in accordance with their main features to facilitate continuity and improve readability. The dissertation is organized as follows:

- Chapter 2 presents the earlier works related to FR-IQA and NR-IQA.

- Chapter 3 presents the basic ideas for improvements in two existing techniques, and analyzes the results obtained by these improvements.

- Chapter 4 proposes a new FR-IQA technique based on the combination of local and global features of the reference and distorted images.

- Chapter 5 proposes a specific purpose NR-IQA technique to evaluate perceptual sharpness using high frequency content based features in images.

- Chapter 6 proposes a general purpose NR-IQA technique using features from different scales of the image to successfully evaluate the quality of a distorted image without the need for any kind of training for its execution.

- Chapter 7 concludes the dissertation by discussing its contributions, limitations and the scopes for future work.

Chapter 2

Related Works

In this chapter, we discuss the existing works related to the research work presented in this dissertation. The presentation of related works is in the same order as the rest of the chapters are arranged. First, the literature related to the FR-IQA is discussed. Next, the existing techniques corresponding to the special purpose NR-IQA are discussed. Finally, the literature related to existing general purpose NR-IQA is discussed. This allows a modular representation of the related research work which are connected with the common thread of IQA and yet, diverse enough in their applicability and procedures.

2.1 FR-IQA Techniques

In the broad area of signal processing, Peak-Signal-to-Noise Ratio (PSNR) is significantly used as an effective measure for comparison of signals and the measurement of signal fidelity. However, in the role of a measure for perceptual image quality, PSNR is not consistent enough [112]. As determined by experimental results presented in the works [114, 129], PSNR is good enough for the perceptual quality evaluation of noisy images, but it is not effective as an IQA method for other kinds of distortions. Interestingly, calculation of PSNR does not involve the spatial relationships with the neighboring pixels and it is one of the best perceptual evaluators for mainly content-independent noise [114, 129]. Still, the general applicability of PSNR is very low when considered with other types of distortions or across several types of distortions. This is where the development and enhancement of FR-IQA techniques became nec-
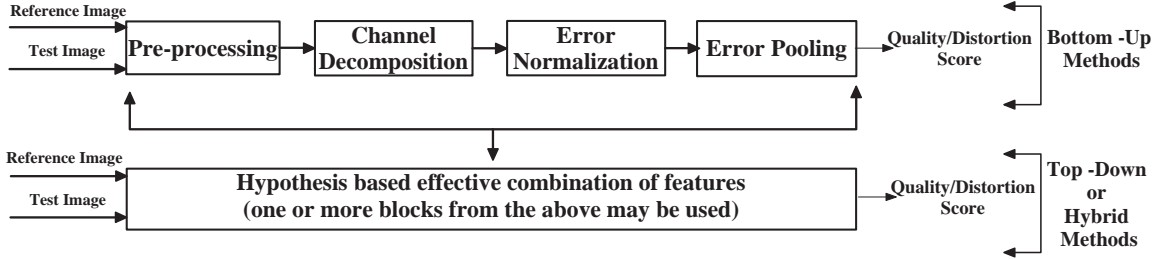
Figure 2.1: Comparison of FR-IQA techniques based on bottom-up, top-down and hybrid approaches.

essary [111]. Since, FR-IQA techniques already use full information of the reference image, it is very important for them to have applicability across several distortions. Therefore, from the last decade, different techniques have been developed for FR-IQA to increase its general applicability. Discrete Cosine Transform (DCT) domain based PSNR has been shown to have much better correlation with HVS in [80] and the method is called PSNR-HVS-M. It is shown in [114] that under effective pooling strategy, PSNR can act as a measure of perceptual IQA. There are FR-IQA techniques that employ the properties of HVS directly and these methods are called bottom-up methods for IQA. There are others, which have hypotheses supporting some properties of HVS or an assumption about the overall functionality of HVS. These methods are called top-down techniques or methods relying on overarching principles. Some existing methods of FR-IQA are combined approaches involving both of the aforesaid techniques and hence may be termed as hybrid techniques [111]. A discussion on several popular techniques of FR-IQA exhibiting their effective performance on different kinds of distortions follows below.

The bottom-up methods in IQA relied on HVS based modeling in IQA. They usually follow the four steps of pre-processing, channel decomposition, error normalization and error pooling as shown in Fig. 2.1. We will describe each of these steps briefly as they have relevance with several existing works related to IQA in general

16

and the research proposed in this dissertation. As we find, some of these blocks are also used to extract information in many existing IQA techniques.

1. **Pre-processing:** Pre-processing in an IQA technique may be an operation or a set of auxiliary operations that are required to be performed before the actual procedures related to IQA start. They may include image registration as the reference and distorted images need proper alignment with respect to each other. Most of the IQA techniques need perfect registration of the reference and test images to perform well. The other operations that may be involved in the phase of pre-processing include color space conversion, color to grayscale conversion, image size adjustment, conversion of pixel values to luminance values, application of Point Spread Function (PSF) simulating the optics of eye and Contrast Sensitivity Function (CSF) following the dependence of eye on the frequency to perceive a change in contrast. Often, CSF is modeled at the later stage of channel decomposition instead of the pre-processing stage so that channels of different frequencies may be obtained by a particular decomposition.

2. **Channel Decomposition:** Channel decomposition is inspired by the fact that a large number of neurons in the primary visual cortex of the human brain are efficient in processing visual stimuli of specific frequency, orientation and location. This inspired the use of Fourier decomposition, Gabor decomposition, DCT transforms, steerable pyramids, cortical transforms and log-Gabor transforms. Often these transforms are used decompose the signals into different frequency and orientations for further processing.

3. **Error Normalization:** The reference and test images are subjected to the channel decompositions and the difference of their coefficients is called error. This error is often processed by CSF to make it perceptually reasonable. Sometimes, contrast masking is also applied for normalization in terms of the error in the neighborhood coefficients. More details about normalization can be found

in [111].

4. **Error Pooling:** After the normalization, the job that remains is to derive a scalar score signifying the distortion or quality. This process of deriving a scalar score from the set of all errors is called pooling. Often these pooling procedures are derived from the properties of HVS. The mean of the error often serve as the score which is called mean pooling. Similarly, we have max pooling, percentile pooling, Minkowski pooling, information-weighted pooling and saliency-weighted pooling as popular pooling strategies in the existing literature.

Some of the models that follow these four steps are: Daly model [14], Safranek Johnson model [103], Lubin model [59, 60], Teo-Heeger model [104] and Watson's models [117, 118]. However, these models significantly differ in their ways of processing information in each stage. In spite of the success of these methods, limited knowledge of HVS and the specific problems [111] with these approaches triggered research in IQA in a different direction which assumes the HVS as a blackbox and hypothesize about its functionality and hence are overarching or top-down in nature. In this regard, the most popular method for FR-IQA is Structural Similarity Index (SSIM) [113]. The hypothesis in this method is about the sensitivity of the HVS in the loss of structural information in images. This approach involves computation of perceptual image quality using the luminance, contrast and structure comparison images obtained from the reference and test images to form a quality map. The final score is derived from the quality map as its mean value. In the past few years, several modifications of SSIM have been proposed. These include Gradient based Structural Similarity (GSSIM) [9] which calculates SSIM on gradient images and Multiscale Structural Similarity Index (MSSSIM) [116] which iteratively downsamples the reference and test images to embed multiscale information in the IQA method to arrive at the value of image quality. Content-partitioned structural similarity [44] classifies

the images into four regions depending on the smoothness and edge properties. The work presented in [96], analyzes the most effective values for the parameters used in SSIM. These methods are based on high level properties of vision. Recently, Gradient Similarity [52] combines low level properties of vision like masking effect and visibility threshold along with structure, luminance and contrast comparison to formulate the quality measure. Other approaches that use low and mid-level properties of vision are Visual-Signal-to-Noise Ratio (VSNR) [8] and Most Apparent Distortion (MAD) [43]. VSNR identifies visual distortions and finds perceived contrast and global precedence properties of the image's edges while calculating the perceptual quality score. On the other hand, MAD evaluates, if the amount of distortion present is near-threshold or supra-threshold and applies different strategies in either case to evaluate perceptual quality.

Information theoretic approach towards FR-IQA has been followed to formulate Information Fidelity Criterion (IFC) [92] and Visual Information Fidelity (VIF) [91]. Information-Content Weighted Structural Similarity Index (IWSSIM) [114] uses local information weighted pooling with MSSSIM to improve the correlation between subjective scores and objective values. Therefore, IWSSIM is a bridge between information theoretic approaches and structural similarity. These approaches rely on the mutual information between two images and the formulation is based on Gaussian Scale Mixture (GSM). Multiscale transforms for IQA have been explored in the works [15, 41, 76]. Using the Haar Wavelet and its space-frequency localization properties, local contrast at various resolutions is used to find image quality in [41]. In [15, 76], coefficients from Discrete Wavelet Transform (DWT) have been used to analyze image quality. Image saliency based approach [62], blockwise Singular Value Decomposition (SVD) of images [95] are also shown to correlate well with HVS. The aforementioned FR-IQA methods differ in principles as well as their execution times and in their correlation with the subjective assessment of quality.

Recently, a phase congruency [38] based technique called FSIM/Feature-Similarity

Index for Color Images (FSIM$_c$) [129] has been proposed. This technique relies on calculating phase congruency and gradient information from the reference image and the test images. The score is based on the similarity of the gradient information and phase congruency values at the corresponding locations. In a similar framework as FSIM, a faster technique called SR-SIM [128] has been developed and shown to perform well in three databases. SR-SIM uses saliency maps derived from spectral residual [32] technique to find perceptual similarity between pixels as well as to pool the quality map obtained from the reference and distorted images. As already discussed, MAD [43] segregates the distortions present in an image as near threshold or supra-threshold and uses dual strategies for evaluating those distortions. In the same context, another approach is based on defining the distortions as additive or detail loss based and combining their measurements. The method is called Additive and Detailed Loss based Measure (ADM) [48]. A recent approach [121] uses the principal of Internal Generative Mechanism (IGM) of human brain and dissolves the given images into predicted and disorderly parts. Then, the distortions on these two parts are separately evaluated by different techniques and are combined to form the objective score. MAD, ADM and VSNR employ the direct applications of the HVS properties like contrast sensitivity and visual masking. Most of the methods mentioned in the previous discussions form the state-of-the-art techniques of FR-IQA as none of the methods discussed can single-handedly outperform others across all databases and distortions therein. Hence, our target in this domain has been to improve the performance of the existing techniques and develop competitive techniques with the existing ones.

2.2 NR-IQA Techniques for Perceived Sharpness Evaluation

As we already discussed in the previous chapter, the existing works on NR-IQA or BIQA can be broadly separated into two genres namely: specific purpose and general purpose IQA. After working with FR-IQA, we selected the problem of NR-IQA to work with. Since removal of reference image created problems for a smooth transition from FR to NR-IQA, we choose to work in specific purpose IQA first. We fixed the specific purpose as perceptual sharpness. The antonym of sharpness is blur. The absence of sharpness does not necessarily mean the presence of blur [108]. However, increase of blurriness hints about loss of sharpness. Hence, several works in existing literature are found to treat both of the problems simultaneously. Since, the dissertation concentrates on investigation of potential features in IQA, our literature review takes a look at the different types features that have been used in this regard.

The existing techniques in the literature involve several visual cues to evaluate perceptual sharpness or blur without the aid of any reference image. The set of variety of visual cues used for blur or sharpness is quite large. Depending on the type of features used to evaluate blur or sharpness, the techniques can be classified as spatial domain, transform domain and hybrid techniques. Though all of the approaches mentioned below do not fall exactly into the category of quantifying perceptual distortions formed due to loss of sharpness, they give information about different types of visual cues that aided in detecting sharpness or blur and the variety of techniques that are developed over several decades to evaluate sharpness or blur levels.

2.2.1 Spatial Domain based Techniques

Spatial domain blur detection methods are those which rely solely on features extracted or pixel properties from the image in the space domain only. Image variance

from entire image is used as a measure in [16] as variance is likely to decrease in an image with the presence of blur. A globally sharp image has lower correlation between adjacent pixels. Thus, when the previous pixel is predicted from the current one, the prediction error is likely to be higher. The variance of this error is used determine a globally blurred or sharp image [106]. Since, neighboring pixels are more correlated in the blurred image compared to its sharper version, an autocorrelation function based method is proposed in [2]. Image intensity histograms [36, 40] based methods are also used to determine blur in auto-focussing applications, though they are not likely to evaluate the perceptual degradation caused by it. By discriminating between different levels of perceptible blur on the same image, a no-reference blur method is proposed by Crete [13]. On the other hand, the image edge or gradient based features happen to be widely used for determination of perceptual blur. The edge feature based methods only rely on the edge width or the spread of edges. The work leading to this approach is [65]; the width of an edge (vertical or horizontal) determined by separation of the extrema positions closest to any edge. This width is considered to be the local edge value and the average of such values over an image served to be the measure of the blurriness of the image. The width of edge using the local gradient direction is used the to measure blurriness in [78]. Degradations caused by the blur uniformly over an whole image are also evaluated using Renyi entropy based anisotropy method [22]. However, depending on the image content, the perceived blurriness of an image varies and this is incorporated through the concept of Just Noticeable Blur (JNB) in the edge width[18] to determine the probability of blur detection for each block region in an image. Finally, a Minkowski metric based pooling is used to determine overall perceived blur. A modification to the same approach using cumulative probability of blur detection is proposed in [75]. Combination of edge based features and variance are used to obtain a sharpness method in [11]. Based on the argument that singular value decomposition of local gradient values represent the components along the direction of significant change of gradient direction, a sharpness measure is proposed

22

in [132].

2.2.2   Transform Domain based Techniques

Transform domain approaches began with failure of blurred image spectra to follow
the power law as done by natural images [19]. The kurtosis obtained from the Fourier
spectra is used to measure the image sharpness in [130]. Some blurriness evaluation
methods using high and low frequency power measures are presented in [21]. A
DCT based approach is proposed in [64]. 2D DCT coefficient based kurtosis is used
to calculate the image sharpness in [6]. Wavelet transform based blur detection
approaches are presented in [105, 109]. The techniques discussed so far did not involve
any kind of prior training. A Wavelet feature based on Support Vector Machine
(SVM) classification and confidence is used to design the quality score of degraded
images in [10]. A complex Wavelet based computation of local phase coherence [28],
called Local Phase Coherence Sharpness Index (LPC-SI), is shown to perform well
in four blur databases among the recent approaches developed for perceptual blur
detection.

2.2.3   Hybrid Techniques

Hybrid techniques are those which rely on both spatial domain and transform domain
methods to evaluate blur. The work proposed by Sadaka et.al [85] uses JNB and
incorporates foveal pooling into it using by Gabor filtering based visual attention map.
A combination of features obtained from local power spectrum, gradient histogram
span and color based maximum saturation is used to detect blur regions in a partially
blurred image and then used to classify these blur regions into motion and out-of-
focus blurred regions [54]. A recent approach S3 [108], uses both spatial and spectral
information to perceptually evaluate blur. In addition, they also come up with a
blur map to identify the sharp and blurred regions in an image. Apart from these
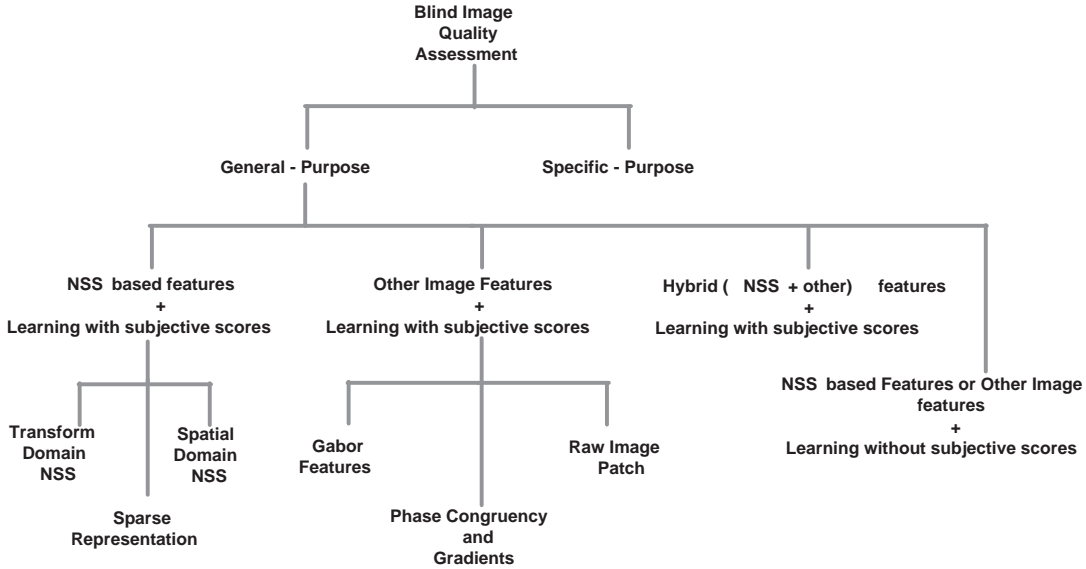
Figure 2.2: Different types of BIQA methods.

techniques, the general purpose image quality assessment techniques also evaluate distortions caused by blur. However, recent perceptual blur or sharpness evaluators are shown to perform better than the general purpose techniques [28]. Our approach in this area has been to combine the high frequency content and contrast based features from natural images to compute the perceived sharpness in a natural image.

2.3 General Purpose NR-IQA Techniques

Recent years have witnessed a substantial proliferation of the general purpose BIQA methods. In this section, our discussion focusses on general purpose BIQA as it is related more to the proposed work. A summary of different types of existing methods is shown in Fig. 2.2.

Most of the general purpose approaches are based on feature extraction and training based on the given human scores in the databases. Since different kinds of degradations affect an image's properties and structures in different manner, these

methods rely on features to train the way the degradations affect the human visual systems. One of the first methods for general purpose IQA is Blind Image Quality Index (BIQI) [71] which relies on identifying the type of degradation on the image and then predict its perceptual quality value. This method can also be called as the bridge between specific purpose and general purpose techniques. The accurate identification of the degradation type prior to the quality assessment is very difficult and most of the other approaches (discussed in the following paragraphs) omitted this step and predicted the image quality directly.

The learning based methods, at first, mostly used properties of Natural Scene Statistics (NSS) as features. Wavelet coefficient statistics are parameterized by generalized Gaussian distribution and these parameters over various scales and orientations are used as features to train SVM and Support Vector Regressor (SVR) in [71]. For testing, similar features extracted from the query image are used to obtain probability of the image belonging to different distortions (using SVM) and these probabilities are combined with the distortion specific support vectors regressors' outputs to predict the quality. DCT coefficient statistics [84] and contourlet domain statistics [58] have been used to train SVR and non-linear functions respectively. Another approach [102] combined complex Wavelet domain statistics with distortion texture statistics and blur/noise statistics to form features for training. The approach in [72] followed a similar route to BIQI but with an enhanced set of NSS features and improved the performance. A spatial domain approach that uses NSS properties is Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [68]. Several properties of mean subtracted contrast normalized luminance coefficients are used as quality features in this approach. The approach is much faster compared to previous NSS based approaches. A method based on sparse representation of NSS features, called Sparse Representation of Natural Scene Statistics (SRNSS), is proposed in [29] and shown to be performing well in several databases.

In another type of approach, Gabor Wavelet features from raw image patches are

used to form visual codebooks and support vector regression is used to predict the quality scores. This method requires images of same or similar degradations for training and is called Codebook based Image Quality (CBIQ) [124]. The number of feature vectors are very high in this method. Another codebook representation based method is Codebook Representation for No-Reference Image Assessment (CORNIA) [125] which learns using the raw image patches but the learning procedure is unsupervised. Entropy and gradients of the distorted image and its phase congruency images are used as features to train a generalized regression neural network in [45]. The method provided competitive performance with BIQI in LIVE database. The aforementioned methods deviated from the NSS features but delivered competitive performance with the NSS based methods.

All of the methods discussed earlier used training with features and DMOS/MOS values provided in the databases. A recent approach called using a technique called Quality Aware Clustering (QAC) uses learning but without the scores [122]. For training, a dataset consisting of reference images and corresponding distorted images of four common distortions is developed. Overlapping patches from a distorted image are selected and evaluated against the reference patches by an FR-IQA method. The similarity score obtained by the evaluation is normalized using the sum of the quality scores of the 10% lowest quality patches. This normalized score for each patch becomes its indicator of quality instead of any DMOS/MOS score. The method achieves competitive performance with the state-of-the-art techniques. Another technique [69] also trains without human scores and uses the NSS features introduced in [68] to decide the Latent Quality Factors (LQF). The method achieves competitive performance with some existing methods.

Some recent approaches [46, 70] work without training on human scores as well on distorted images. The method in [70] generates a parameterized Multivariate Gaussian Model (MVG) based on NSS from high quality images by drawing quality aware features from it. For any distorted image, the distance between the NSS feature

model and the MVG model fitted to the distorted image becomes the indicator of quality. This method,named as Natural Image Quality Evaluator (NIQE), delivers competitive performance with the state-of-the-art techniques in the LIVE database. Another approach that uses no modeling is [46]. They use simple relationship with the features of mean phase congruency, entropy and mean gradient magnitude to accomplish their tasks and the method is called No-Reference Quality Index (NRQI). However, prior to this, they use the first step of BIQI [71] to classify the distortions. Hence, the method implicitly involves training on distorted images. Though the performance of these techniques attempting to minimize training is not the best, they provide a methodology to continue research in the direction of no-training IQA. Any kind of training makes the IQA measure dependent on the training ensemble and requires effort and time to generate the trained model. Motivated by these techniques, we attempt to use the multiscale information from the image to establish an NR-IQA technique, absolutely free from any training with human scores, pristine images or distorted images.

2.4   Discussion on the Existing Approaches

The works related to the three sub-areas of IQA are presented in this chapter. In the FR-IQA techniques, a generalized model for the bottom-up methods has been outlined. However, the model's use has been restricted by the advent of overarching methods that search for features based on the hypotheses, though they may also use the overarching blocks. The variety of features used to develop FR-IQA techniques are visibly huge. As one progresses with the specific purpose NR-IQA, features affected by a specific type of distortion are extracted and used for computing the score. Hence, in some general purpose NR-IQA techniques, detection of the distortion type is followed by the computation of quality score. In another direction in general purpose BIQA, deviation from NSS models are captured. In a third direction, effective features

are clustered by supervised or unsupervised learning. Efforts have also been spent to reduce the elements used in the learning process for NR-IQA. The existing approaches in three types of IQA discussed, insinuates at the effectiveness of features and their combinations to arrive at the IQA score for an image. Also, the aim of several techniques has been to provide feasible solution to the problem of IQA from an engineering point of view rather than delving into the actual functionalities of HVS and its mathematical modeling. Hence, we concentrate on searching effective features and their combinations to improve the performance level in the three sub-areas.

Chapter 3

Improvements in Existing Methods

This chapter presents our initial efforts in the field of IQA. We initiated with FR-IQA and have tried to work with existing techniques in this field. We have changed the features used in two existing techniques to improve the performance and have reported our findings in this chapter. The common thing which binds these two techniques is the framework which is used by both of these techniques. We discuss about the framework and then propose changes in the features used in this framework. The remainder of the chapter is arranged as follows. The framework is discussed in details in Section 3.1. The modification made to the first technique related to this framework and experimental analysis are discussed in Section 3.2. Section 3.3 discusses the changes made to the second technique and related experiments. Finally, the chapter is concluded in Section 3.4.

3.1 Introduction

A recent work by Zhang *et al.* develops Feature Similarity Index (FSIM) using phase congruency [38] for FR-IQA. In this section, we discuss the framework in details. The block diagram of the framework used in their technique is shown in Fig. 3.1. This framework is also inspired by the similarity calculation technique used in SSIM [113]. For the grayscale reference and test images I_R and I_T, the method applies the appropriate scale selection procedure mentioned in [110] to the images to obtain the images I_1 and I_2 respectively. From the resulting images, the values of the phase congruency maps PC_1 and PC_2 are calculated. Since, the phase congruency is contrast invariant,
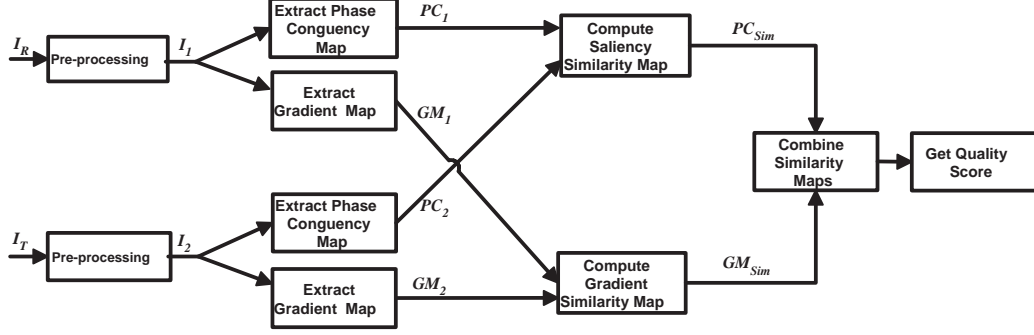
Figure 3.1: Block diagram of the framework used in FSIM.

the gradient magnitude features of the images are used to account for the sensitivity of HVS on the local contrast. The gradient magnitude images GM_1 and GM_2 are calculated using the Scharr operator [35]. Then, FSIM map is calculated for a pixel x in I_1 as

$$FSIM_{map}(x) = PC_{sim}(x)GM_{sim}(x) \max\left(PC_1(x), PC_2(x)\right). \qquad (3.1)$$

where PC_{sim} and GM_{sim} are calculated as

$$PC_{sim}(x) = \left(\frac{2.PC_1(x).PC_2(x) + \epsilon_1}{PC_1{}^2(x) + PC_2{}^2(x) + \epsilon_1}\right)^a \qquad (3.2)$$

and

$$GM_{sim}(x) = \left(\frac{2.GM_1(x).GM_2(x) + \epsilon_2}{GM_1{}^2(x) + GM_2{}^2(x) + \epsilon_2}\right)^b. \qquad (3.3)$$

Here, a and b are used to adjust the relative contribution of phase congruency similarity and gradient magnitude similarity. For simplicity, the values of a, b are chosen to be 1. ϵ_1 and ϵ_2 are constants in the equations required to avoid instability arising due very low denominator values. $\max\left(PC_1(x), PC_2(x)\right)$ acts as the weight for pooling and hence the FSIM measure is defined as

$$FSIM = \frac{\sum\limits_{x \in \Psi} FSIM_{map}(x)}{\sum\limits_{x \in \Psi} \max\left(PC_1(x), PC_2(x)\right)}. \qquad (3.4)$$

Here, Ψ stands for the indices of the elements or pixels of I_1. The processing changes a little for color images. Color images are converted to YIQ space with 'Y' component

representing the luminance, 'I' and 'Q' components representing the chrominance components respectively. From the luminance components of the test and reference images, $FSIM_{map}$ is calculated first. The chrominance similarities are calculated as

$$IP_{sim}(x) = \frac{2.IP_1(x).IP_2(x) + \epsilon_3}{IP_1{}^2(x) + IP_2{}^2(x) + \epsilon_3} \qquad (3.5)$$

and

$$QP_{sim}(x) = \frac{2.QP_1(x).QP_2(x) + \epsilon_4}{QP_1{}^2(x) + QP_2{}^2(x) + \epsilon_4} \qquad (3.6)$$

where IP_1 and QP_1 are chrominance components for the image I_1. Here, ϵ_3 and ϵ_4 are again constants in the equations required to avoid instability arising due very low denominator values. Using the luminance and chrominance similarity, feature similarity color map $FSIM_{cmap}$ is formed by

$$FSIM_{cmap}(x) = (IP_{sim}(x)QP_{sim}(x))^\lambda FSIM_{map}(x), \qquad (3.7)$$

where λ is used for assigning importance to the chrominance similarities. Then the feature similarity index for color images (FSIM$_c$) is calculated as

$$\text{FSIM}_c = \frac{\sum\limits_{x \in \Psi} FSIM_{cmap}(x)}{\sum\limits_{x \in \Psi} \max\left(PC_1(x), PC_2(x)\right)}. \qquad (3.8)$$

The values of FSIM and FSIM$_c$ overlap for grayscale images. P. Kovesi's phase congruency algorithm [38] is used in the implementation of FSIM and FSIM$_c$. FSIM and FSIM$_c$ demonstrated competitive performance with the state-of-the-art algorithms on six publicly available databases. Using the aforementioned framework, we propose two works in this dissertation depending on the following facts:

1. The prediction accuracy for noisy images remained low in FSIM. The design of phase congruency is meant to derive dimensionless features (which are not influenced by noise) from noisy images as well and therefore as we will elaborately discuss in the coming section, the sensitivity of phase congruency towards noise is deliberately reduced under the assumption that the noise is additive with a

31

constant power spectrum. However for the task of FR-IQA, the way of noise reduction and subsequent normalization of energy by the sum of amplitude might suppress some features which are actually necessary for HVS. This hypothesis is, therefore, the motivation for our proposed work which is explained in the following section. The new method proposed using this technique is called Phase Deviation Sensitive Energy Features based Similarity (PDSESIM).

2. In another new FR-IQA method proposed by Zhang *et al.* [128], the framework used in FSIM is used. However, the phase congruency features are replaced by Spectral Residual (SR) based saliency map [32]. Following SR, several variants of spectral measures for visual saliency have come up. These new variants differ in their computational times as well as in performances and have established themselves better than or competitive with SR as measures of visual saliency. The effectiveness of these saliency measures in FR-IQA is still an open question. In our work, a study to evaluate the performance of the recent spectral approaches for visual saliency (hence spectral saliency) for FR-IQA is presented.

3.2   From FSIM to PDSESIM

First, we present a discussion on phase congruency and related schemes. Then, we discuss the modifications, followed by our experiments and analysis.

3.2.1   Phase Congruency

Phase congruency is a method for feature detection from signals using the phase information of signals. The method is proposed by Morrone et. al. [73, 74]. The main idea of the method is : features of signal also correspond to those points where phase congruency or similarity is maximum. This is shown by the Fig. 3.2 where a 1-D triangular wave and its first few harmonics obtained by the Fourier series
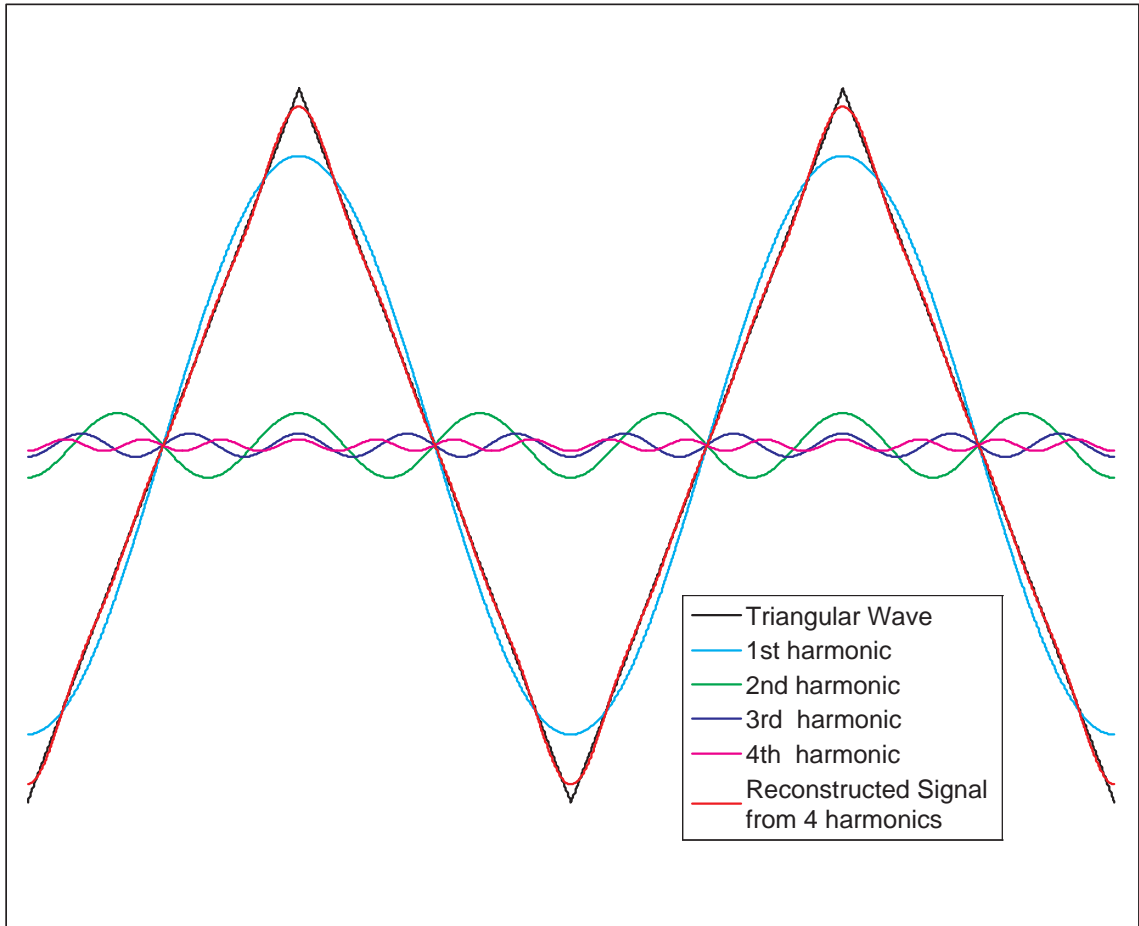
Figure 3.2: Demonstration of phase congruency among the harmonics of a triangular wave and reconstructed signal using four harmonics.

are plotted. At the crests and troughs of the triangular wave, one can find the all harmonics to be in phase together. It is shown in the work of Morrone, maximal in-phase Fourier components correspond to points that belong to image features like step edges, deltas, lines or combinations as well. Their model of feature perception is called local energy model which proves that local energy of a signal is directly proportional to the phase congruency of the signal and perceived features exist at the points of maximum phase similarity of a signal. Let $g(x)$ be a square integrable

periodic (f_0) waveform, the Fourier series of $g(x)$ is given by

$$g(x) = \sum_n a_n \sin(2\pi n f_0 x + \theta_n) \quad n \geq 0, a_n \geq 0. \tag{3.9}$$

Here a_n and θ_n denote the amplitude and phase of the nth harmonic of the expansion. The Hilbert Transform of $g(x)$ is defined as

$$g_h(x) = -\sum_n a_n \cos(2\pi n f_0 x + \theta_n) \quad n > 0. \tag{3.10}$$

By the definition itself, $g_h(x)$ does not have a DC component. Since $g(x)$ and $g_h(x)$ form a quadrature pair, the local energy is defined as

$$E(x) = \sqrt{\left(\sum_n a_n \sin\left(2\pi n f_0 x + \theta_n\right) \right)^2 + \left(-\sum_n a_n \cos\left(2\pi n f_0 x + \theta_n\right) \right)^2}. \tag{3.11}$$

Each component $a_n \sin(2\pi n f_0 x + \theta_n)$ and $b_n \cos(2\pi n f_0 x + \theta_n)$ are the projections of a function with amplitude a_n and argument $\text{modulo}(2\pi n f_0 x + \theta_n, 2\pi)$ in the space spanned by the orthogonal signals g and g_h. Thus, $E(x)$ is the magnitude of the resultant signal formed by the sum of all such functions. If the argument of this resultant signal is $\overline{\phi(x)}$ which is also the amplitude weighted local mean phase angle of all the Fourier terms at the point x, $E(x)$ can be formulated as $E(x) = \sum_n a_n \cos(\phi_n(x) - \overline{\phi(x)})$. This signifies, as the deviations of the phase of the Fourier components from $\overline{\phi(x)}$ decrease, the value of local energy will increase, implying higher similarity among the phases of the signal components. Now in Eqn. 3.11, if the instantaneous phases of all harmonics are same then $(2\pi n f_0 x + \theta_n) = \phi(x)$ and

$$E(x) = \sqrt{\left(\sum_n a_n \sin\left(\phi(x)\right) \right)^2 + \left(\sum_n a_n \cos\left(\phi(x)\right) \right)^2}$$
$$= \sum_n a_n \tag{3.12}$$

The phase congruency is therefore defined as

$$PC(x) = \frac{E(x)}{\sum_n a_n}. \tag{3.13}$$

34

$PC(x)$ is a dimensionless, positive quantity and bounded by the upper limit 1 as, $E_{\max}(x) = \sum_n a_n$. Kovesi [38], developed a convenient measure to calculate local energy using odd and even log Gabor Wavelets. As depicted in Eqn. 3.11, both of the quadrature components are free from DC. Log Gabor Wavelets offer large bandwidth with zero DC and they have Gaussian transfer function on logarithmic scale as defined by Fields [20]. If $ge_n(x)$ and $go_n(x)$ are the responses of the even and odd log Gabor filter respectively at scale n (signifying different center frequencies and geometric scaling of a reference filter), the amplitude $A_n(x)$ is calculated as

$$A_n(x) = \sqrt{ge_n{}^2(x) + go_n{}^2(x)}. \tag{3.14}$$

The quadrature signals g and g_h that are required for the calculation of local energy can be estimated as

$$g(x) \approx \sum_n ge_n(x) \tag{3.15}$$

and

$$g_h(x) \approx \sum_n go_n(x). \tag{3.16}$$

Hence, direction of the unit vector representing the local energy is found as

$$(u(x), u_h(x)) = \left(\frac{g(x)}{\sqrt{g^2(x) + g_h{}^2(x)}}, \frac{g_h(x)}{\sqrt{g^2(x) + g_h{}^2(x)}} \right). \tag{3.17}$$

Then local energy can be found using the dot product of the unit vector and filter responses as

$$
\begin{aligned}
E(x) &= \sum_n A_n cos(\phi_n(x) - \overline{\phi(x)}) \\
&= \sum_n \left[(ge_n(x), go_n(x)) \, (u(x), u_h(x))^T \right] \\
&= \sum_n \left[ge_n(x)u(x) + go_n(x)u_h(x) \right].
\end{aligned}
\tag{3.18}
$$

As noticed by Kovesi, the calculation of phase congruency becomes ill conditioned if the values of the Fourier components become small, or if there is a single frequency in

the signal. Hence, a small positive constant δ is added in the denominator to account for this. Two additional changes are brought into the calculation of energy $E(x)$. A term more sensitive to the change of phase deviation is introduced to calculate the energy of the signal. Increased value of phase congruency means the cosine of the phase deviation $(\phi_n(x) - \overline{\phi(x)})$ at the corresponding point is high and phase deviation is low. Therefore, the sine of the phase deviation will be low in magnitude. Thus, the energy contributed by each component is calculated as

$$E_n(x) = A_n(x) \left[\cos(\phi_n(x) - \overline{\phi(x)}) - \left| \sin(\phi_n(x) - \overline{\phi(x)}) \right| \right] \tag{3.19}$$

where $\left| \sin(\phi_n(x) - \overline{\phi(x)}) \right|$ is obtained from the cross product as

$$\left| \sin(\phi_n(x) - \overline{\phi(x)}) \right| = |ge_n(x)u_h(x) - go_n(x)u(x)| \tag{3.20}$$

This energy is proved to be more phase deviation sensitive measure for the phase congruency. To obtain these filter banks in 2D, a Gaussian spreading function is multiplied with a radially varying log Gabor function in the radial frequency-angle (ω, θ) plane. Thus the 2D log Gabor function is obtained as

$$G(\omega, \theta) = \exp\left(\frac{\log(\omega/\omega_0)}{2\log(k/\omega_0)} \right)^2 . \exp\left(-\frac{(\theta - \theta_0)^2}{2\sigma_\theta^2} \right) \tag{3.21}$$

Here, ω_0 is the filter's center frequency which controls the scale n, the term (k/ω_0) is responsible for the bandwidth and shape of the filter. θ_0 is responsible for the orientation of the 2D log Gabor filter. The design of the log Gabor Wavelet filter bank has been made such that a uniform coverage of the spectrum in all orientations is obtained and so the ratio between the angular spacing between the filters and σ_θ is made constant. To reduce the sensitivity to noise, a noise circle is calculated and values of energy greater than this noise N_{θ_0} (which is the radius of the circle) are considered in the calculation of phase congruency. For each orientation θ_0, the smallest scale filter has the largest bandwidth and hence the noise is calculated using

36

that for θ_0. The defnition of phase congruency becomes

$$PC(x) = \frac{\sum\limits_{n} \lfloor E_n(x) - N_{\theta_0} \rfloor}{\delta + \sum\limits_{n} A_n(x)}. \tag{3.22}$$

3.2.2 Proposed Modification

The proposed method uses phase deviation sensitive energy features instead of phase congruency to form energy maps. The energy is phase deviation sensitive as it is calculated in the way Kovesi used it to enhance the sensitivity towards the phase similarity as shown in Eqn. 3.19. However, in the proposed method we refrain from the calculation of N_{θ_0} which removes much of the noise actually present in the images. As for the visual clarification, we show in Fig. 3.3, the phase congruency maps calculated from a reference and distorted image (distorted with additive white Gaussian noise) from LIVE database [94] along with the corresponding phase deviation sensitive energy images calculated with and without noise adjustment. The dynamic range of the energy images is adjusted to be within 0 to 1 for viewing. As revealed from the diagram, though the noise present in the image (Fig. 3.3(e)) is creating a visual disturbance, much of it is removed when the noise is adjusted (Fig. 3.3(h)). For the energy maps without any noise adjustment, the visual distortions are visible in the energy map (Fig. 3.3(g)). However, even after noise adjustment, the phase congruency map still shows some undesirable visual disturbance in the said areas. This is due to the fact, in the definition of phase congruency given in Eqn. 3.22, some amount of noise causing the high energy is removed. However, such noise is also responsible for the changes in the denominator as $\sum\limits_{n} A_n(x)$. But the denominator remains unchanged thereby leaving some traces of noise in the phase congruency maps. If the phase deviation sensitive energy is less than the corresponding noise value in the appropriate scale, the phase congruency value is zero otherwise the effect of noise is prevalent in the denominator and therefore in $PC(x)$. This feature is brought out by the images as well. Fig. 3.3(h) shows that noise adjusted energy has left very little amount of

noise or very low values of noise in the energy map. However, Fig. 3.3(f) is derived from Fig. 3.3(h) by dividing each point of it by $\sum_n A_n(x), (x \in \Psi)$. Hence, more traces of noise can be found in Fig. 3.3(f). Therefore, we use the phase deviation sensitive energy features instead of the phase congruency values. Higher values of energy features are required corresponding to noisy pixels as they are shown to influence the HVS. It is shown in different works [52, 129], that, PSNR which relies just on pixel differences without any knowledge of the their interdependence, is able to correlate better as a perceptual measure of noise. Therefore, the differences arising due to noise is important together with the interdependence of pixels. We calculate the phase deviation sensitive energy features from the luminance component of reference image as E_{s1} as

$$E_{s1}(x) = \sum_n E_{n1}(x). \tag{3.23}$$

where $E_{n1}(x)$ calculated for I_1 as per Eqn. 3.19. Similarly, we compute E_{s2} from the test image. We form $\hat{E}_{s1}$ and $\hat{E}_{s2}$ which have values less than or equal to unity by

$$\hat{E}_{s1}(x) = E_{s1}(x) / \max_{y \in \Psi}(E_{s1}(y)); \tag{3.24}$$

and

$$\hat{E}_{s2}(x) = E_{s2}(x) / \max_{y \in \Psi}(E_{s2}(y)); \tag{3.25}$$

In this way, we normalize the energy range before comparison with each other. However, this is totally different from the earlier normalization by $\sum_n A_n(y), (y \in \Psi)$ as it changed each energy value according to the energy response at each value of y. In our case, the normalization is just a uniform scaling of all phase deviation sensitive energy values. The energy similarity is calculated as

$$\hat{E}_{sim}(x) = \frac{2.\hat{E}_{s1}(x).\hat{E}_{s2}(x) + \epsilon_1}{\hat{E}_{s1}^2(x) + \hat{E}_{s2}^2(x) + \epsilon_1}. \tag{3.26}$$

We calculate gradient similarity in our method and hence the phase deviation sensitive energy based image similarity (PDSESIM) map is formed as

$$PDSESIM_{map}(x) = \hat{E}_{sim}(x)GM_{sim}(x) \max\left(\hat{E}_{s1}(x), \hat{E}_{s2}(x)\right). \tag{3.27}$$
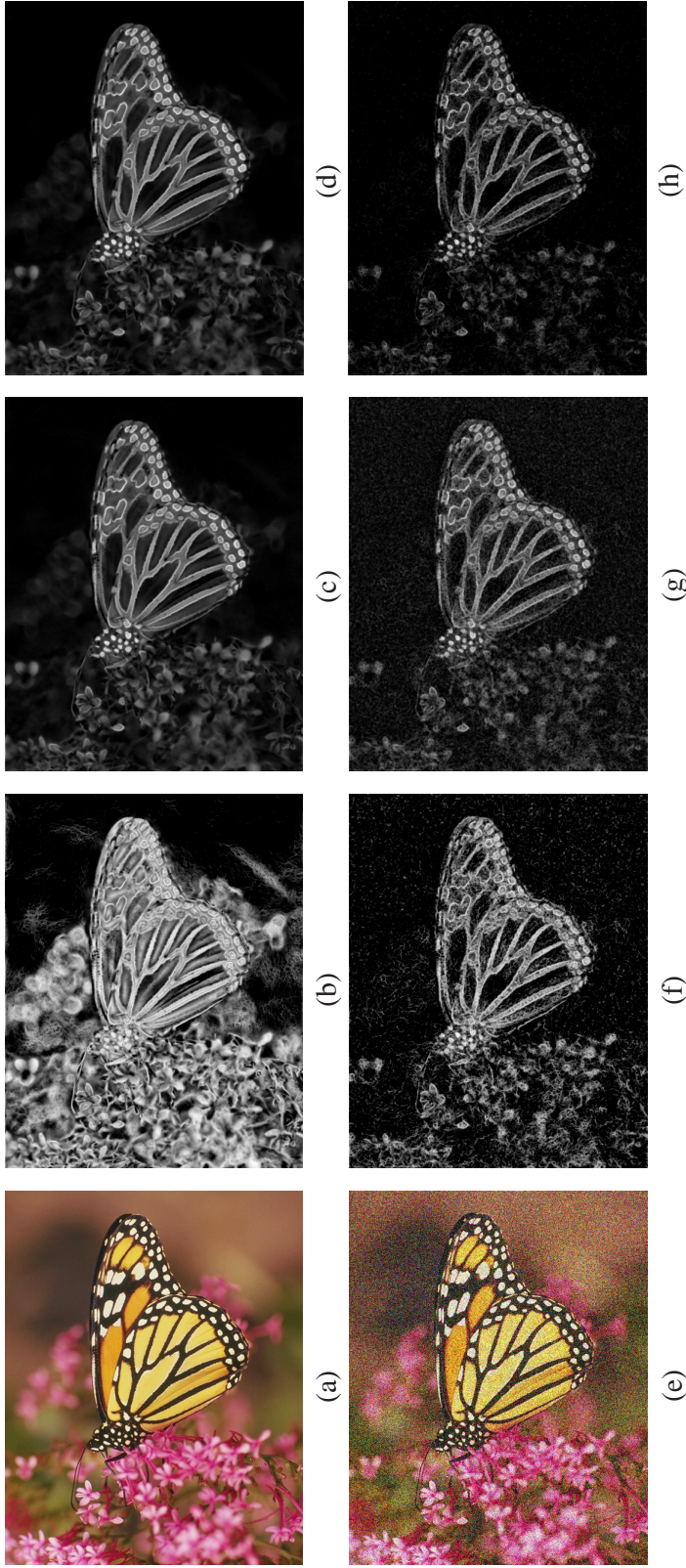
Figure 3.3: (a),(e) Reference 'Monarch' image and corresponding test image with adaptive white Gaussian noise [94]; (b),(f) Phase congruency maps corresponding to (a) and (e) resp; (c),(g) Energy maps without removing noise corresponding to (a) and (e) resp; (d),(h) Energy maps after removing noise corresponding to (a) and (e) resp.

39

This is extended to color images by

$$PDSESIM_{cmap}(x) = (IP_{sim}(x)QP_{sim}(x))^{\lambda} PDSESIM_{map}(x). \qquad (3.28)$$

For grayscale images, $PDSESIM_{cmap}(x) = PDSESIM_{map}(x)$. The proposed measure is calculated as

$$PDSESIM = \frac{\sum\limits_{x \in \Psi} PDSESIM_{cmap}(x)}{\sum\limits_{x \in \Psi} \max\left(\hat{E}_{s1}(x), \hat{E}_{s2}(x)\right)}, \qquad (3.29)$$

We have used the values of the parameter $\lambda = 0.3$ as used in the implementation in [129]. The values of the parameters for computing sensitive features are the same as used in the implementation by Kovesi. The values of the parameters used to find energy and gradient similarity maps are discussed in the following section.

3.2.3   Experimental Results

We discuss various experiments and their results to demonstrate the promise of the proposed method. This section is divided into three parts for discussion on parameter selection, distortion-wise performance on individual databases and overall performance in all the databases. The proposed method is evaluated in six databases already discussed in Chapter 1. For analysis of the overall performance of the proposed method, several evaluation measures have been used.

- **Parameter Selection**

  The parameters for computing phase congruency are kept same as in the implementation by Kovesi. The values of ϵ_3 and ϵ_4 used in chrominance similarity measurement are the same as mentioned in [129]. Here, we discuss the selection of parameters ϵ_1 and ϵ_2 used to find the phase deviation sensitive energy feature similarity (Eqn. 3.26) and gradient similarity (Eqn. 3.3) between the images. The values of $\hat{E}_{s1}$ and $\hat{E}_{s2}$ vary from 0 to 1. Therefore the value of ϵ_1 is varied

from 0.05 to 0.95 with an interval of 0.05. ϵ_2 is used for gradient images and its value is varied from 10 to 250 with an interval of 10. Therefore, for each value of ϵ_1, we have 25 values of ϵ_2. Together we have $19 \times 25 = 475$ different pairs for the values of (ϵ_1, ϵ_2) each of which is denoted by an index value (j) varying from 1 to 475. The SROCC values across the six databases obtained using each index value j for PDSESIM and FSIM$_c$ is shown in Fig. 3.4. Starting from the beginning, if a non-overlapping set of 25 indices are formed, each of those sets correspond to a fixed value of ϵ_1 and 25 different values of ϵ_2. We notice a repetitive pattern of the SROCC values for each of such non-overlapping sets. As revealed in the diagram, the highest SROCC value for PDSESIM has exceeded the highest SROCC value obtained by FSIM$_c$ in every database. However, for the following experiments we need to fix the parameter values. The maximum value SROCC can take is 1. For every parameter index, we calculate the sum of squares of deviation of corresponding SROCC value from 1. We call this quantity 'correlation error' CE_j which is calculated for the jth parameter index as

$$CE_j = \sum_{d=1}^{6}(1 - SROCC_{d,j})^2, \tag{3.30}$$

where d represents each of the six databases and $SROCC_{d,j}$ is the SROCC value in d^{th} database for j^{th} parameter set. The plot of correlation error for all the parameter indices is shown in Fig. 3.5. The minimum of the SROCC is obtained for index value $j = 10$ which corresponds to the parameter values $\epsilon_1 = 0.05$ and $\epsilon_2 = 100$. For the rest of the experiments, these values are used.

- **Distortion wise Performance in Different Databases**

Though the aim of FR-IQA is to find generic measures for perceptual assessment of quality for images, the distortion-wise performance of the methods is crucial enough. If the distortion type is known, this analysis will facilitate the selection of a distortion specific particular method. It may also happen, that
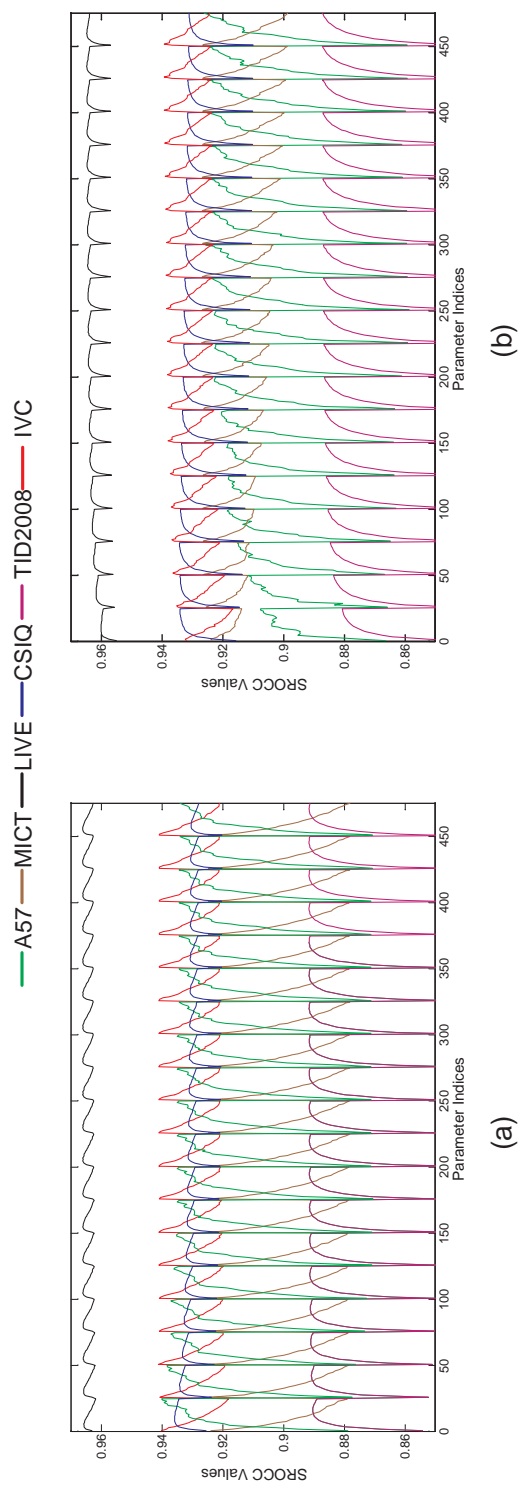
Figure 3.4: (a) Fluctuations in SROCC values using PDSESIM (b) Fluctuations in SROCC values using FSIM_c.
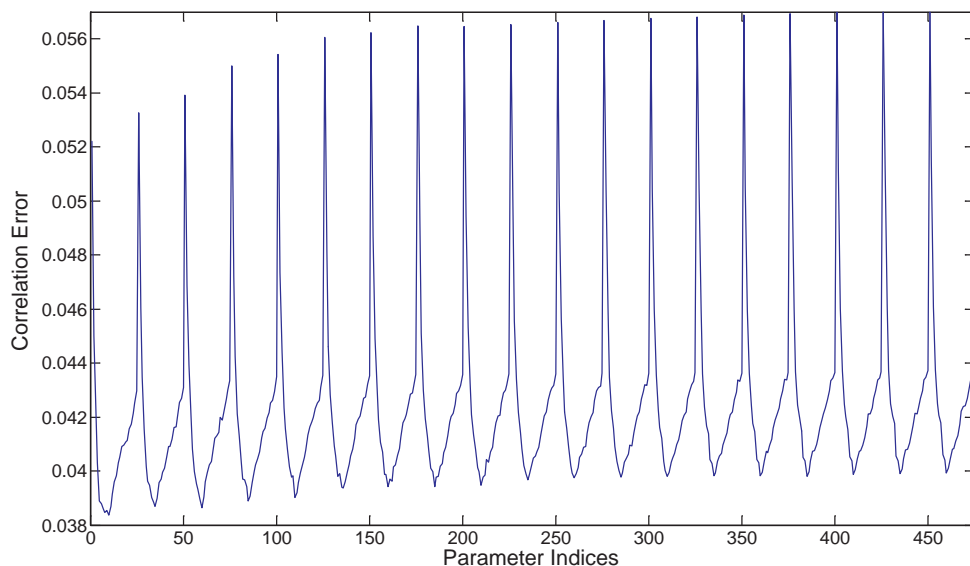
Figure 3.5: Variation of correlation error with parameter indices using PDSESIM.

a method has good overall performance for all databases; however, for one or more distortions it is a bad evaluator of perceptual quality.

Here, we look into the distortion wise performance of the proposed method. However, we have chosen only the largest three databases where number of distortions is 5 or more. Thus, LIVE, CSIQ and TID2008 databases are chosen. The results for the distortion-wise performance of the proposed method is presented in Table 3.1. In the table, the top two performances for each distortion type in each database are highlighted in bold. Next, the best performance among the phase congruency related methods : Measure [56], FSIM, $FSIM_c$ and PDSESIM is underlined. In the databases, distortion wise performance shows that there exists no single approach that is superior to others in all cases. However, the proposed approach lies among the top two approaches in 12 of the 28 cases discussed here. These 12 cases belong to various distortions in different databases. Hence, greater generalization ability is exhibited by PDSESIM than

the other methods. For 9 of these 28 cases, VIF lies within the top two methods. For most of the noises in the databases, the superiority of PSNR and PSNR-HVS-M [80] is observed. Though PSNR is not among the top two performers for wn in LIVE database, it is indeed very close to the second best. Also, we notice that IWSSIM performs well for blur images in all of the three databases and MSSSIM does the same in two databases. The comparison among the Measure [56], FSIM, FSIM_c and PDSESIM shows that PDSESIM has indeed improved the generalization ability of FSIM_c which is its closest contender. In 21 of the 28 cases, PDSESIM proves itself to be superior to the other three. Also in each database, PDSESIM has performed better than the other three for majority of the distortions. PDSESIM improves the performance for most of the noises compared to Measure [56], FSIM and FSIM_c.

- **Overall Performance in Different Databases**

Here, we evaluate the overall performance of PDSESIM across all databases. This evaluation is very important as it conveys, how well a method is suited to predict the perceptual quality caused by different types of distortions, taken all together. We have 4 types of experiments here. In the first one, we show graphically, how well the objective score is fitting with the subjective scores of the database. Secondly, we compare the performance of the proposed method with the state-of-the art methods quantitatively, using 5 evaluation measures in 6 databases. Thirdly, we compute the average performance of all the methods in 6 databases. The final experiment compares the execution times of all the methods.

In Fig. 3.7, the performance of the PDSESIM is presented for the 6 databases. The fitting gives visual idea about the performance of the PDSESIM. The quantitative evaluation is presented in Table 3.2. From the table, it is clear that PDSESIM demonstrates improved performance compared to Measure [56]

Table 3.1: Distortion-wise performance comparison using SROCC for different FR-IQA measures in LIVE, CSIQ and TID2008 databases.

Database	Distortion Type	PSNR	PSNR-HVS-M	SSIM	MSSSIM	IWSSIM	VIF	VSNR	MAD	Measure [56]	FSIM	FSIM$_c$	PDSESIM
LIVE	jpeg2k	0.8954	0.9540	0.9614	0.9628	0.9649	0.9696	0.9677	0.5501	0.8511	0.9716	**0.9723**	**0.9706**
	jpeg	0.8809	0.9650	0.9764	0.9814	0.9808	**0.9846**	0.6272	0.9763	0.9088	0.9834	0.9840	**0.9844**
	wn	0.9854	**0.9863**	0.9694	0.9733	0.9667	**0.9857**	0.8182	0.9844	0.9056	0.9652	0.9716	0.9802
	gblur	0.7823	0.9314	0.9517	0.9543	**0.9719**	0.9728	0.6403	0.9464	0.9388	0.9707	**0.9709**	0.9690
	fastfading	0.8907	0.9144	0.9556	0.9471	0.9442	**0.9650**	0.6813	**0.9569**	0.9362	0.9499	0.9520	0.9529
CSIQ	awgn	0.9363	**0.9584**	0.9256	0.9471	0.9380	**0.9575**	0.9254	0.9542	0.7308	0.9262	0.9360	**0.9543**
	gblur	0.9291	0.9711	0.9245	0.9711	**0.9781**	**0.9744**	0.9445	0.9681	0.9486	0.9728	0.9728	**0.9741**
	global-contrast	0.8623	0.8655	0.7399	**0.9528**	**0.9540**	0.9345	0.8714	0.9210	0.3997	0.9420	0.9439	0.9486
	fnoise	0.9338	0.9436	0.8925	0.9330	0.9057	0.9509	0.9082	**0.9568**	0.7713	0.9233	0.9369	**0.9526**
	jpeg2k	0.8879	0.9512	0.9219	0.9631	0.966	**0.9703**	0.9172	0.9614	0.9043	0.9652	0.9662	**0.9676**
	jpeg	0.9361	0.9701	0.9205	0.9682	0.9682	0.9671	0.9485	**0.9752**	0.8835	0.9684	0.9703	**0.9760**
TID2008	awgn	**0.9115**	**0.9110**	0.8310	0.8094	0.7869	0.8800	0.7728	0.8388	0.8565	0.8566	0.8758	0.8985
	awgn-pink	**0.9068**	0.8479	0.8134	0.8064	0.7920	0.8785	0.7793	0.8258	0.8538	0.8527	0.8931	**0.9022**
	spatial-corr-noise	**0.9229**	**0.9262**	0.8438	0.8195	0.7714	0.8703	0.7665	0.8678	0.8472	0.8483	0.8711	0.9039
	masked noise	**0.8487**	0.7555	0.7561	0.8155	0.8087	**0.8698**	0.7295	0.7336	0.8021	0.8021	0.8264	0.7659
	high frequency noise	**0.9323**	0.9136	0.8919	0.8685	0.8662	0.9075	0.8811	0.8864	0.9099	0.9093	0.9156	**0.9219**
	impulse noise	**0.9177**	**0.8339**	0.7072	0.6868	0.6465	0.8331	0.6471	0.0650	0.7473	0.7452	0.7719	0.7925
	quantization noise	0.8699	**0.9007**	**0.8745**	0.8537	0.8177	0.7956	0.8270	0.8160	0.8558	0.8564	0.8726	0.8622
	gblur	0.8682	0.9219	0.9596	**0.9607**	0.9636	0.9546	0.9330	0.9197	0.9461	0.9472	0.9472	0.9393
	denoising	0.9381	0.9572	0.9595	0.9571	0.9473	0.9189	0.9286	0.9434	0.9595	0.9603	0.9618	**0.9660**
	jpeg	0.9011	0.9338	0.9270	**0.9348**	0.9184	0.9170	0.9174	0.9275	0.9279	0.9279	0.9294	**0.9341**
	jpeg2k	0.8300	0.9762	0.9723	0.9736	0.9738	0.9713	0.9515	0.9707	0.9773	0.9773	**0.9780**	**0.9824**
	jpeg-trans-error	0.7665	0.8105	0.8668	0.8736	0.8588	0.8582	0.8055	0.8661	0.8704	0.8708	**0.8756**	**0.8851**
	jpeg2k-trans-error	0.7765	**0.9321**	0.8707	0.8525	0.8203	0.8510	0.7909	0.8394	0.8529	0.8544	0.8555	**0.8718**
	pattern-noise	0.5931	0.6812	0.7168	0.7336	**0.7724**	0.7608	0.5716	**0.8287**	0.7488	0.7491	0.7514	0.7627
	blockwise-distortion	0.5852	0.6238	**0.8529**	0.7617	0.7623	0.8320	0.1926	0.7970	0.8491	**0.8492**	0.8464	0.7321
	intensity-shift	0.6974	0.6932	**0.7575**	0.7374	0.7067	0.5132	0.3710	0.5161	0.6711	0.6720	0.6554	0.6012
	contrast change	0.6126	0.5815	0.6329	0.6400	0.6301	**0.8190**	0.4239	0.2723	**0.6542**	0.6481	0.6510	0.6310

45

and FSIM in all databases and in 4 out of 6 databases compared to FSIM$_c$. As compared to the state-of-the-art methods, PDSESIM's next strongest contender is MAD which lies within top two performers in 3 of the 6 databases. PDSESIM is the best performer in the TID2008 database which has maximum number of images as well as maximum types of distortions. Also, FSIM$_c$, IWSSIM, VIF and VSNR has been in the top two performers in at least one database. However, there exists no method which has performed equally well in all databases. To further evaluate the performance of the proposed method, we averaged the SROCC, KROCC and PLCC scores across all databases as shown in Table 3.3a. In the case of simple averaging, we find PDSESIM performs best among all the FR-IQA measures using all of SROCC, KROCC and PLCC. In weighted averaging, the number of distorted images in each database (see Chapter 1, Section 1.3.1), has been used as the weight for the databases. PDSESIM is the best performer for all of the three correlation scores here as well.

Finally, we come to the comparison of execution times of the aforesaid measures. The execution time has been found out by running the MATLAB codes for each method on a computer with a Pentium Core 2 duo processor with frequency of 2 GHz using a pair of reference and distorted images of size 768×512. It shows from the Table 3.3b, that the processing times of PSNR, SSIM and VSNR are very low. Compared to those, the processing times of IWSSIM, FSIM, FSIM$_c$, VIF, MAD and PDSESIM are much higher with PDSESIM having the lowest execution time among all of these. In the earlier experiments, PDSESIM showed improved performance distortion-wise and in overall performance evaluation. Also, similar to FSIM$_c$, PDSESIM works on color images, but its execution time is lower compared to that of FSIM which works on grayscale images only. This is due to the omittance of noise circle calculation from the measure of phase deviation sensitive energy.

Table 3.2: Overall performance comparison of different FR-IQA methods in six databases.

Database	Criterion	PSNR	PSNR-HVS-M	SSIM	MSSSIM	IWSSIM	VIF	VSNR	MAD	Measure [56]	FSIM	FSIM$_c$	PDSESIM
LIVE	SROCC	0.8756	0.9295	0.9479	0.9513	0.9566	0.9636	0.6481	**0.9669**	0.8848	0.9634	0.9645	**0.9652**
	KROCC	0.6865	0.7659	0.7963	0.8049	0.8178	0.8282	0.4879	**0.8421**	0.6975	0.8339	0.8365	**0.8378**
	PLCC	0.8722	0.9250	0.9449	0.9489	0.9519	0.9604	0.7018	**0.9675**	0.8851	0.9597	0.9613	**0.9616**
	MAE	10.4830	7.9368	6.9109	6.6701	6.3805	6.0952	14.6406	**5.2071**	10.0560	5.9301	**5.8236**	5.8339
	RMSE	13.3659	10.3792	8.9437	8.6143	8.3757	7.6089	19.5193	**6.9068**	12.7122	7.6705	7.5236	**7.4929**
CSIQ	SROCC	0.8057	0.8221	0.8368	0.9132	0.9212	0.9194	0.8132	**0.9466**	0.7543	0.9242	0.9309	**0.9359**
	KROCC	0.6078	0.6523	0.6325	0.7386	0.7522	0.7532	0.6279	**0.7963**	0.5728	0.7561	0.7684	**0.7786**
	PLCC	0.7999	0.8117	0.8144	0.8986	0.9142	**0.9277**	0.8016	**0.9505**	0.8120	0.9115	0.9186	0.9244
	MAE	0.1193	0.1056	0.1144	0.0857	0.0789	0.0742	0.1138	**0.0631**	0.1177	0.0783	0.0745	**0.0703**
	RMSE	0.1576	0.1655	0.1525	0.1168	0.1082	**0.0980**	0.1590	**0.0817**	0.1544	0.1098	0.1060	0.1042
TID2008	SROCC	0.5245	0.5612	0.8081	0.8528	0.8559	0.7496	0.7046	0.8340	0.8785	0.8805	**0.8840**	**0.8892**
	KROCC	0.3696	0.4509	0.6056	0.6543	0.6636	0.5863	0.5340	0.6445	0.6913	0.6946	**0.6991**	**0.7089**
	PLCC	0.5545	0.6029	0.8061	0.8419	0.8572	0.8075	0.6507	0.8311	0.8656	0.8733	**0.8758**	**0.8807**
	MAE	0.8918	0.7221	0.6139	0.5616	0.5245	0.5837	0.7085	0.5543	0.5002	0.4912	**0.4868**	**0.4716**
	RMSE	1.1168	1.0885	0.7948	0.7247	0.6915	0.8007	1.0274	0.7491	0.6731	0.6543	**0.6482**	**0.6360**
A57	SROCC	0.6189	0.8962	0.4059	0.8435	0.7750	0.6224	**0.9355**	0.9023	0.5098	0.9181	0.9181	**0.9299**
	KROCC	0.4309	0.7261	0.2780	0.6529	0.5880	0.4592	**0.8031**	0.7233	0.3722	0.7639	0.7639	**0.7793**
	PLCC	0.6347	0.8748	0.4149	0.8394	0.7652	0.6137	**0.9500**	0.9043	0.5785	0.9252	0.9252	**0.9341**
	MAE	0.1607	0.0932	0.1847	0.1119	0.1182	0.1417	**0.0574**	0.0856	0.1518	0.0794	0.0794	**0.0754**
	RMSE	0.1899	0.1211	0.2238	0.1337	0.1587	0.1957	**0.0770**	0.1051	0.2006	0.0933	0.0933	**0.0877**
MICT	SROCC	0.6132	0.8480	0.7870	0.8874	**0.9202**	0.9077	0.8608	**0.9362**	0.7219	0.9059	**0.9067**	0.8963
	KROCC	0.4443	0.6568	0.5922	0.7029	**0.7537**	0.7315	0.6745	**0.7823**	0.5385	0.7302	**0.7303**	0.7175
	PLCC	0.6498	0.8575	0.7996	0.8924	**0.9246**	0.9136	0.8698	**0.9367**	0.7378	0.9075	**0.9072**	0.8989
	MAE	0.7746	0.4984	0.5649	0.4287	**0.3653**	0.4012	0.4599	**0.3493**	0.6292	0.3980	**0.4015**	0.4201
	RMSE	0.9517	0.6448	0.7551	0.5650	**0.4775**	0.5096	0.6193	**0.4384**	0.8532	0.5270	**0.5275**	0.5491
IVC	SROCC	0.6884	0.8832	0.7789	0.8980	0.9125	0.8964	0.7993	0.9146	0.8053	0.9262	**0.9293**	0.9286
	KROCC	0.5218	0.6935	0.5939	0.7203	0.7339	0.7158	0.6053	0.7406	0.6093	0.7564	**0.7636**	0.7639
	PLCC	0.6983	0.8902	0.7915	0.9106	0.9228	0.9026	0.8026	0.9210	0.8111	0.9375	**0.9390**	0.9381
	MAE	0.6875	0.4349	0.5297	0.3701	0.3684	0.4050	0.5490	0.3673	0.5275	0.3384	**0.3277**	0.3228
	RMSE	0.8770	0.5563	0.7593	0.5103	0.4704	0.5276	0.7325	0.4753	0.7195	0.4246	**0.4203**	0.4246

3.2.4 Discussion

Compared to the existing state-of-the-art approaches, the proposed method exhibited improved performance in distortion wise perceptual analysis. The method also demonstrated better generalization performance than its nearest competitors FSIM, MAD and VIF in the different experiments conducted but has faster computation time compared to these. Using several evaluation criteria depicting prediction monotonicity, accuracy and error, the overall performance of the method is found to be robust and competitive with the state-of-the arts. The robustness of the method is also corroborated by the best overall average performance of PDSESIM across six benchmark FR-IQA databases.

3.3 Comparison of Spectral Saliency based Techniques

The framework discussed in the computation of FSIM is used to generate another technique for FR-IQA. This technique uses saliency maps obtained from the reference and distorted images using a technique called SR [32] for computing saliency instead of the Phase Congruency (PC) maps. SR is one of the state-of-the-art techniques in visual saliency detection. SR has been used earlier for FR-IQA but embedded in a different framework [61, 63]. Other saliency based FR-IQA techniques use visual saliency for the purpose of pooling in the later stages of the algorithm [114]. Since visual attention is one of the integral properties of HVS which the IQA methods aim to imitate, the justification of using image saliency for IQA is intuitive. Inspired by the potential and performance of SR for FR-IQA, this work attempts to explore and evaluate the performance of several recently developed spectral saliency detection methods in FR-IQA. The spectral methods selected are those which have gradually evolved being inspired from SR and showed better or competitive performance with

the same. The current work is an effort to study how these newly emerged spectral saliency methods support FR-IQA and evaluate their performance by comparison with the state-of-the-arts and SR.

3.3.1 Selected Spectral Approaches

During the evaluation process, we change the approach for saliency extraction method and generate the quality scores keeping the framework same. Here, we discuss the different spectral saliency computation methods used to compute saliency maps used during the evaluation.

1. **Spectral Residual (SR)**: Calculation of SR involves the residual log amplitude (computed as the difference of log amplitude and low-pass filtered log amplitude) and phase of the Fourier transform. The inverse Fourier Transform (FT) is applied on the residual and the related phase to generate the saliency maps. SR based saliency maps are used to computer the objective quality scores and hence this technique for score computation is evidently SR-SIM.

2. **Phase Fourier Transform (PFT)**: The second approach to calculate the saliency map is PFT [27] based saliency which uses the only the phase of Fourier transform to compute the saliency maps. As shown in [27], a faster and better computation of saliency map is obtained using PFT compared to SR.

3. **Phase of Quaternion Fourier Transform (PQFT)**: Further innovation led towards the use of PQFT [27] to form saliency maps. Instead of a two-dimensional FT, the technique for computing Quaternion Fourier Transform (QFT), represents each pixel of a color image as a quaternion comprising of color, intensity and motion features (if possible). For the application of Quaternion Fourier transform, the color channels are required and hence, this method works on color images only.

4. **Eigen Spectral Residual (Eigen SR)**: As any quaternion can be represented in a polar form with the help of its eigen axis and eigen angle which are used for the generalization of the imaginary part of the quaternion, the approach presented in [90] computes saliency map based on the eigen axes and eigen angles of the quaternion. One of the techniques proposed in the work [90] is called Eigen SR. The residual is calculated as the difference of log amplitude and low-pass filtered log amplitude of the QFT of the image. Then inverse QFT is applied on the residual and the related eigen axis and eigen angle to generate the saliency maps.

5. **Eigen Phase of Quaternion Fourier Transform (Eigen PQFT)**: In the same work [90], another technique called Eigen PQFT [90] is proposed for spectral saliency determination. However, it differs in principle used in computation of PQFT as the inverse QFT is calculated based on the log amplitude, eigen axis and eigen angles instead of using quaternion phase as used in PQFT.

6. **Hypercomplex Fourier Transform (HFT)**: Another saliency detection approach using HFT based saliency detection has been recently introduced in [47]. The scale at which the saliency map will be perfect is decided by the saliency entropy criterion.

7. **Discrete Cosine Transform (DCT)**: Apart from Fourier Transform, an approach has used discrete cosine transform (DCT) to generate features for image saliency called image signature [31]. This method can extract saliency from both color and grayscale images.

8. **Quaternion Discrete Fourier Transform (QDCT)**: The earlier work related to image signature using DCT is further extended to QDCT [89] for image saliency detection.

As seen from the aforesaid methods, the computations for saliency map is done in spectral domain using different transforms (Fourier, DCT). Also, the color channels of the images are embedded using the concept of quaternion in several of the cases. All of the aforementioned spectral saliency detection methods have been chosen for this study. Therefore, altogether eight saliency determination algorithms have been used to obtain the quality scores from the images.

3.3.2 Experimental Evaluation

In this section, we discuss on two topics. Since, the basic framework consists of the parameters ϵ_1, ϵ_2 and λ, the values of these parameters are required for the evaluation. The evaluation is carried out using six benchmark databases.

- **Parameter Selection**

  First, we fix the value of λ as it decides the contribution of gradient similarity maps towards forming the SIM_{map}. The value of λ is taken to be as 0.5 similar to that mentioned in [128]. Next, we followed the same procedure for the selection of the remaining parameters ϵ_1 and ϵ_2 used to find the corresponding similarity maps. Hence, the value of ϵ_1 is varied from 0.05 to 0.95 with an interval of 0.05 and hence 19 different values are obtained. ϵ_2 is used for gradient images and its value is varied from 10 to 250 with an interval of 10. Then we estimate the correlation error, but the equation is slightly modified to

  $$CE_j = \left(\frac{1}{N_D} \sum_{d=1}^{N_D} (1 - SROCC_{d,j})^2 \right)^{0.5}, \qquad (3.31)$$

  where d represents each of the N_D databases and $SROCC_{d,j}$ is the SROCC value in d^{th} database for j^{th} parameter set. This equation is basically an average over the relevant number (N_D) of databases followed by a square root of the earlier definition of correlation error mentioned in Eqn. 3.30 and hence does not change the order of the error values for the FR-IQA measure. The plot of correlation

51

error for all the parameter indices using different types of spectral saliency based FR-IQA methods is shown in Fig. 3.6. The ϵ_1 and ϵ_2 values are chosen corresponding to index (j) for which CE_j is minimum. This experiment reveals different sets of parameter values for different saliency based approaches as shown in Table 3.4. In this method of selecting parameters, equal importance has been given to all of the databases which vary a lot in size, subjective score range and in types of distortion present in images. Since, the FR-IQA methods need to be consistent across several types of distortions, the parameters chosen in this way can have better applicability with many distortions. The range of the correlation error plots is also significant. A small range assures that the variation in parameter values has less effect on the method whereas a bigger range shows that Q_s is more dependent on the parameters. Again, the lesser the higher limits of CE, the closer is the performance of the method to perfection. The least higher limit in CE is exhibited by PQFT. Though the selected framework is embedded in the FSIM and SR-SIM, it is quite dependent on the parameters.

- **Comparison Results**

We have compared the seven aforesaid saliency based FR-IQA methods with each other and also with state-of-the-art methods SR-SIM, MSSSIM, IWSSIM, VIF, MAD and FSIM$_c$. Since the quaternion based approaches process color images only, there are no results for these approaches in A57 database which contains grayscale images only. The results for FSIM$_c$ is equal to those obtained using FSIM for grayscale images in A57 database. For SR based IQA, we have used the parameters used in the implementation of SR-SIM [128]. The comparison among different IQA methods is summarized in Table 3.5. SR-SIM has the best results for TID2008 database. PFT performs better in TID2008 and IVC database. PQFT gives a consistent performance, though it has never been
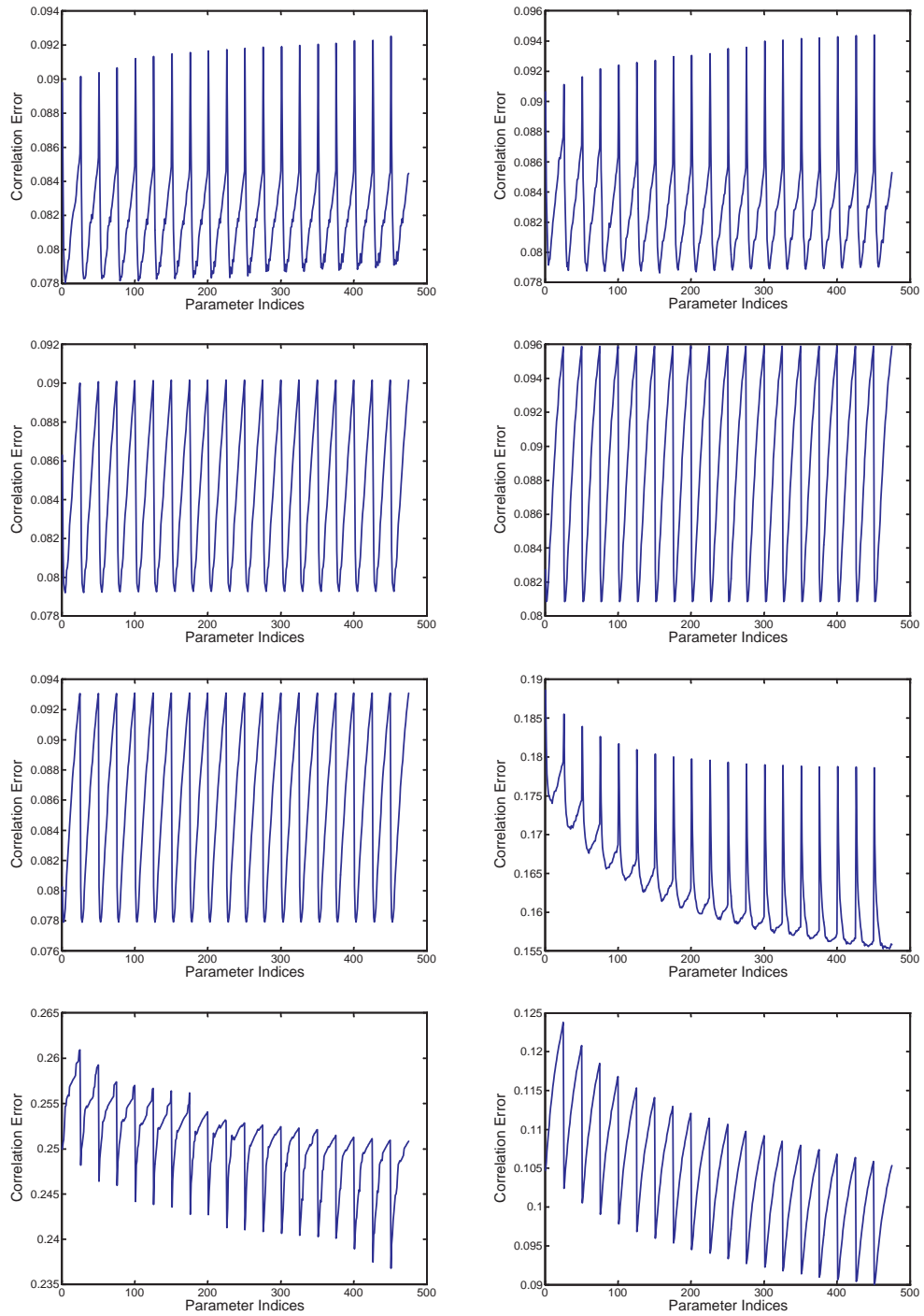
Figure 3.6: Correlation error across the databases for (a) SR, (b) PFT, (c) PQFT, (d) Eigen SR, (e) Eigen PQFT, (f) HFT, (g) DCT and (h) QDCT.

among the top two performers except for one database. It performs consistently well in all databases. Eigen SR performs better than SR-SIM in LIVE, IVC and MICT datasets. In TID2008, IVC and MICT database, Eigen PQFT performs better than PQFT. HFT and DCT based FR-IQA methods vary a lot from the best results of each database. However, the performance of QDCT based method is much better in all of the databases. In Table 3.6, we show the direct and weighted average SROCC values across all databases obtained using different FR-IQA methods. We find that saliency based methods, SR-SIM, PFT, PQFT and Eigen PQFT based methods have higher average SROCC. In [128], SR-SIM is shown to be faster than several FR-IQA techniques. PFT calculation is faster than SR [27]. PQFT, Eigen SR and Eigen PQFT involve quaternion fourier transform and hence are computationally expensive compared to PFT. Therefore, among the methods that have competitive performance with the state-of-the-arts, PFT is the fastest.

3.4 Chapter Summary

We have presented improvements in two existing techniques tied by a common framework. Firstly, we presented a phase deviation sensitive energy similarity index PDSESIM for FR-IQA using phase deviation sensitive energy features computed using log Gabor filters. The recent success of phase congruency based image features for FR-IQA is the motivation for this work. However, the phase congruency algorithm used for FR-IQA is less sensitive to noise depending on the noise removal technique used. The noise adjusted phase deviation sensitive energy features are used to compute phase congruency after normalization by filter amplitude response. Based on the sensitiveness of the HVS on noise, we have proposed the use of phase deviation sensitive energy features to evaluate perceptual quality of images. We have also demonstrated a technique to find the suitable parameters required to compute the phase deviation

sensitive energy similarity. Secondly, we have presented a comparative study of spectral saliency based approaches for the purpose of FR-IQA. This study enables us to see that several aspects of saliency based FR-IQA. Firstly, the raw values present in the saliency maps formed are very important and though they may form better saliency maps or help in better detection of salient areas, they may not help to design a better FR-IQA method. Secondly, we show the dependence of the framework (used in FSIM_c and SR-SIM) on the values of the parameters. Thirdly, this study enabled us to evaluate the relative capabilities of the spectral saliency based methods to form an FR-IQA method. The future work involves using suitable saliency measures for FR-IQA via a framework which has lesser dependence on parameters.
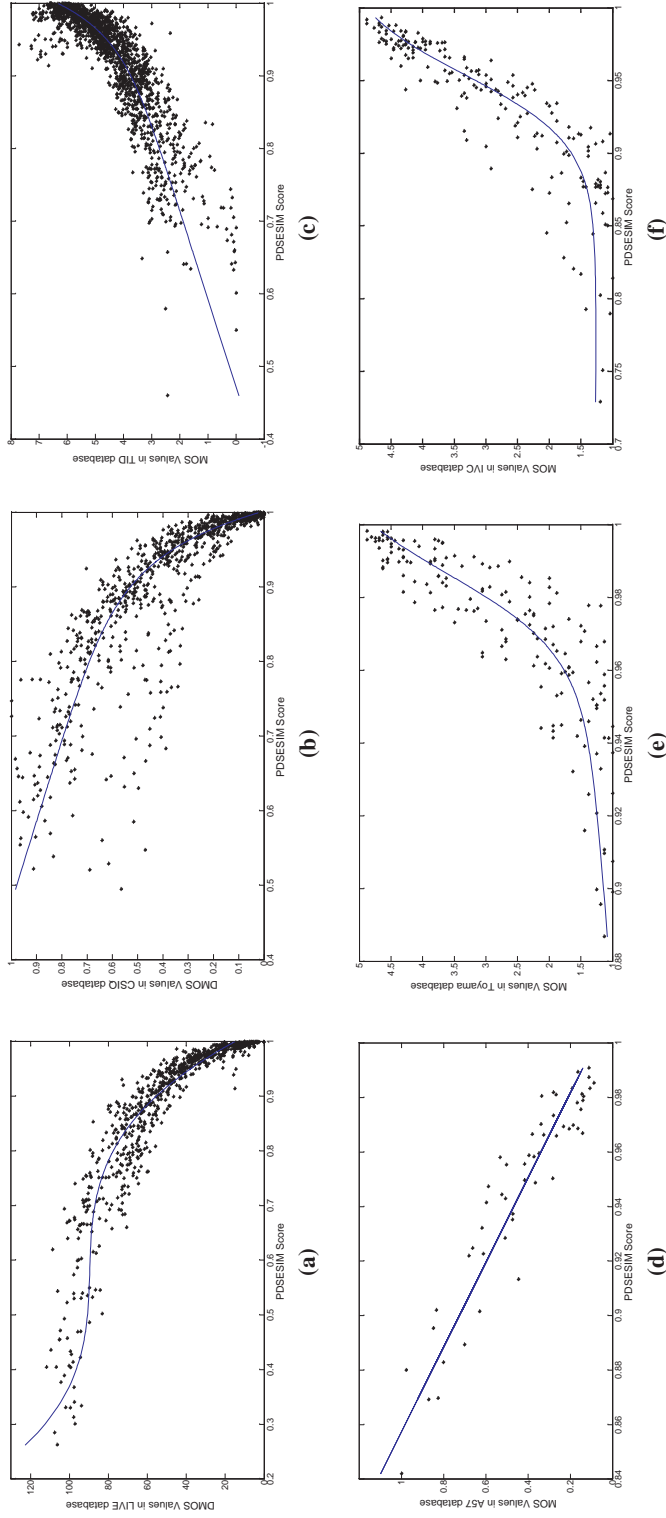
Figure 3.7: Scatter plots of the subjective evaluation scores(MOS/DMOS)versus PDSESIM scores in (a) LIVE (b) CSIQ (c) TID2008 (d) A57 (e) MICT and (f) IVC shown in black dots. The blue line/curve is the plot of the output of the predicted quality score by the logistic function versus MOS/DMOS.

Table 3.3a: Average performance comparison of different FR-IQA methods over six databases.

	Criterion	PSNR	PSNR-HVS-M	SSIM	MSSSIM	IWSSIM	VIF	VSNR	MAD	Measure [56]	FSIM	FSIM$_c$	PDSESIM
Direct Average	SROCC	0.6877	0.8234	0.7608	0.8910	0.8902	0.8432	0.7936	0.9168	0.7591	0.9197	**0.9223**	**0.9242**
	KROCC	0.5101	0.6576	0.5831	0.7123	0.7182	0.6790	0.6221	0.7549	0.5803	0.7559	**0.7603**	**0.7643**
	PLCC	0.7016	0.8270	0.7619	0.8886	0.8893	0.8543	0.7961	0.9185	0.7817	0.9191	**0.9212**	**0.9230**
Weighted Average	SROCC	0.6757	0.7314	0.8356	0.8908	0.8964	0.8457	0.7329	0.8971	0.8352	0.9117	**0.9153**	**0.9186**
	KROCC	0.5021	0.5879	0.6455	0.7104	0.7225	0.6861	0.5598	0.7326	0.6498	0.7434	**0.7491**	**0.7558**
	PLCC	0.6896	0.7474	0.8302	0.8828	0.8949	0.8736	0.7177	0.8972	0.8447	0.9055	**0.9087**	**0.9120**

Table 3.3b: Execution time (in seconds) comparison of PDSESIM with existing methods.

PSNR	PSNR-HVS-M	SSIM	MSSSIM	IWSSIM	VIF	VSNR	MAD	Measure [56]	FSIM	FSIM$_c$	PDSESIM
0.0073	4.8363	**0.1402**	0.7467	3.3204	7.4848	0.177	29.4643	15.9786	3.3695	3.6542	3.2551

57

Table 3.4: Different ϵ_1 and ϵ_2 values used with different spectral saliency based FR-IQA framework.

	PFT	PQFT	Eigen SR	Eigen PQFT	HFT	DCT	QDCT
ϵ_1	0.35	0.05	0.05	0.05	0.95	0.95	0.95
ϵ_2	70	50	20	30	220	10	10

Table 3.5: Performance of different spectral saliency approaches used in the FR-IQA framework in six databases.

Database	Criterion	SR	PFT	PQFT	Eigen PQFT	Eigen SR	HFT	DCT	QDCT	MSSSIM	IWSSIM	VIF	MAD	FSIM_c
LIVE	SROCC	0.9618	0.9647	**0.9659**	0.9629	0.9651	0.9403	0.9365	0.9467	0.9513	0.9566	0.9636	**0.9669**	0.9645
	KROCC	0.8299	0.8357	**0.8374**	0.8299	0.8354	0.7796	0.7697	0.8001	0.8049	0.8178	0.8282	**0.8421**	0.8365
	PLCC	0.9552	0.9615	0.9635	0.9619	**0.9638**	0.9389	0.9336	0.9429	0.9489	0.9519	0.9604	**0.9675**	0.9613
	MAE	6.3263	5.8746	5.6067	5.7533	**5.5954**	7.5383	8.0128	7.3918	6.6701	6.3805	6.0952	**5.2071**	5.8236
	RMSE	8.0792	7.5093	7.3135	7.4662	**7.2821**	9.4050	9.7837	9.0955	8.6143	8.3757	7.6089	**6.9068**	7.5236
CSIQ	SROCC	0.9319	0.9392	0.9344	0.9279	0.9319	0.9179	0.8570	**0.9454**	0.9132	0.9212	0.9194	**0.9466**	0.9309
	KROCC	0.7719	0.7827	0.7728	0.7631	0.7698	0.7401	0.6634	**0.7933**	0.7386	0.7522	0.7532	**0.7963**	0.7684
	PLCC	0.9241	0.9324	0.9227	0.9165	0.9202	0.9088	0.8854	**0.9526**	0.8986	0.9142	0.9277	**0.9505**	0.9186
	MAE	0.0734	0.0694	0.0726	0.0756	0.0735	0.0806	0.0974	**0.0607**	0.0857	0.0789	0.0742	**0.0631**	0.0745
	RMSE	0.1021	0.0960	0.1032	0.1075	0.1052	0.1098	0.1221	**0.0801**	0.1168	0.1082	0.098	**0.0817**	0.1060
TID2008	SROCC	**0.8913**	0.8870	0.8826	0.8797	0.8855	0.8279	0.7491	0.8544	0.8528	0.8559	0.7496	0.8340	0.8840
	KROCC	**0.7149**	0.7030	0.6952	0.6948	0.7013	0.6287	0.5580	0.6636	0.6543	0.6636	0.5863	0.6445	0.6991
	PLCC	**0.8854**	0.8827	0.8742	0.8777	0.8813	0.8154	0.7946	0.8660	0.8419	0.8572	0.8075	0.8311	0.8758
	MAE	**0.4543**	0.4782	0.4979	0.4872	0.4814	0.5905	0.6531	0.5153	0.5616	0.5245	0.5837	0.5543	0.4868
	RMSE	**0.6246**	0.6307	0.6522	0.6433	0.6344	0.7771	0.8149	0.6712	0.7247	0.6915	0.8007	0.7491	0.6482
A57	SROCC	**0.9295**	0.9170	-	-	-	-	0.5205	-	0.8435	0.7750	0.6224	0.9023	**0.9181**
	KROCC	**0.7779**	0.7583	-	-	-	-	0.3610	-	0.6529	0.5880	0.4592	0.7233	**0.7639**
	PLCC	**0.9247**	0.9177	-	-	-	-	0.5344	-	0.8394	0.7652	0.6137	0.9043	**0.9252**
	MAE	**0.0778**	0.0818	-	-	-	-	0.1592	-	0.1119	0.1182	0.1417	0.0856	**0.0794**
	RMSE	**0.0936**	0.0978	-	-	-	-	0.2079	-	0.1337	0.1587	0.1957	0.1051	**0.0933**
MICT	SROCC	0.8825	0.9054	0.9133	0.9131	0.9154	0.8624	0.9095	0.9145	0.8874	**0.9202**	0.9077	**0.9362**	0.9067
	KROCC	0.6975	0.7303	0.7421	0.7426	0.7462	0.6758	0.7381	0.7434	0.7029	**0.7537**	0.7315	**0.7823**	0.7303
	PLCC	0.8870	0.9056	0.9153	0.9164	0.9182	0.8599	0.9144	0.9198	0.8924	**0.9246**	0.9136	**0.9367**	0.9072
	MAE	0.4444	0.3950	0.3792	0.3755	0.3705	0.4835	0.3804	0.3671	0.4287	**0.3653**	0.4012	**0.3493**	0.4015
	RMSE	0.5782	0.5317	0.5053	0.5017	0.4968	0.6398	0.5071	0.4924	0.565	**0.4775**	0.5096	**0.4384**	0.5275
IVC	SROCC	0.9265	**0.9402**	0.9322	**0.9361**	0.9350	0.7515	0.8953	0.9212	0.8980	0.9125	0.8964	0.9146	0.9293
	KROCC	0.7560	**0.7809**	0.7671	**0.7739**	0.7717	0.5601	0.7083	0.7561	0.7203	0.7339	0.7158	0.7406	0.7636
	PLCC	0.9357	**0.9495**	0.9432	**0.9488**	0.9470	0.7511	0.9073	0.9301	0.9106	0.9228	0.9026	0.9210	0.9390
	MAE	0.3404	**0.3050**	0.3188	**0.3130**	0.3150	0.5877	0.4096	0.3267	0.3701	0.3684	0.4050	0.3673	0.3277
	RMSE	0.4306	**0.3826**	0.4052	**0.3850**	0.3918	0.8137	0.5126	0.4509	0.5103	0.4704	0.5276	0.4753	0.4203

Table 3.6: Average performance of different spectral saliency approaches used in the FR-IQA framework across six databases.

	Criterion	SR	PFT	PQFT	Eigen SR	Eigen PQFT	HFT	DCT	QDCT	MSSSIM	IWSSIM	VIF	MAD	FSIM$_c$
Direct	SROCC	0.9206	0.9256	**0.9257**	0.9239	**0.9266**	0.8600	0.8113	0.9164	0.8910	0.8902	0.8432	0.9168	0.9223
	KROCC	0.7580	**0.7652**	0.7629	0.7609	**0.7649**	0.6769	0.6331	0.7513	0.7123	0.7182	0.6790	0.7549	0.7603
Average	PLCC	0.9187	**0.9249**	0.9238	0.9243	**0.9261**	0.8548	0.8283	0.9223	0.8886	0.8893	0.8543	0.9185	0.9212
Weighted	SROCC	**0.9172**	**0.9191**	0.9162	0.9129	0.9170	0.8704	0.8240	0.9012	0.8908	0.8964	0.8457	0.8971	0.9153
	KROCC	**0.7541**	**0.7548**	0.7491	0.7454	0.7512	0.6853	0.6389	0.7310	0.7104	0.7225	0.6861	0.7326	0.7491
Average	PLCC	0.9119	**0.9154**	0.9097	0.9098	**0.9128**	0.8621	0.8516	0.9081	0.8828	0.8949	0.8736	0.8972	0.9087

Chapter 4

Full-reference image quality assessment by combining global and local distortion measures

Full-reference image quality assessment (FR-IQA) techniques compare a reference and a distorted/test image and predict the perceptual quality of the test image in terms of a scalar value representing an objective score. The evaluation of FR-IQA techniques is carried out by comparing the objective scores from the techniques with the subjective scores (obtained from human observers) provided in the image databases used for the IQA. Hence, we reasonably assume that the goal of a human observer is to rate the distortion present in the test image. The goal oriented tasks are processed by the human visual system (HVS) through top-down processing which actively searches for local distortions driven by the goal. Therefore local distortion measures in an image are important for the top-down processing. At the same time, bottom-up processing also takes place signifying spontaneous visual functions in the HVS. To account for this, global perceptual features can be used. Therefore, we hypothesize that the resulting objective score for an image can be derived from the combination of local and global distortion measures calculated from the reference and test images. We calculate the local distortion by measuring the local correlation differences from the gradient and contrast information. For global distortion, dissimilarity of the saliency maps computed from a bottom-up model of saliency is used. The remaining parts of the chapter are arranged in the following manner. The basic idea behind the proposed approach is discussed in Section 4.1. The motivating factors for the

proposed approach are presented in Section 4.2. The details of the proposed method
are presented in Section 4.3. The performance of the proposed method is analyzed in
Section 4.4. The concluding remarks are presented in Section 4.5.

4.1   Basic Idea

As already discussed in Chapter 2, Section 2.1, MAD, ADM and IGM apply different
strategies to separately evaluate different aspects of distortion. Inspired by all of
these, the aspect we focus on is the automatic evaluation of image quality based on
global and local perceptual visual cues and their combination. As explained in [98], vi-
sual processing is a simultaneous combination of bottom-up and top-down processes.
Bottom-up processing focuses on highlighting the relatively important regions in an
image [34]. On the other hand, top-down processing caters to goals and it searches
actively for local features [77] based on contextual information. In HVS, these pro-
cesses occur simultaneously and quickly. Whenever, a human subject is given the
task of evaluating perceptual quality, he has the final goal to rate the distortions
present. Based on this, we hypothesize that three main events taking place during
this evaluation: 1. Initially, the global image content determines the varying attention
(regionwise) of the evaluator; 2. existing regional distortions are used to assign local
quality evaluations; 3. finally, based on the attention, a refinement on the previous
evaluation is carried out and final rating is provided. To serve the first event, we
resort to the global perceptual features of the image. For the second step, local at-
tributes between the reference and test/distorted images are compared. Finally, third
step involves a pooling strategy taking place based on the global perceptual features.
In the present work, the perceptual quality of an image is expressed in terms of image
saliency maps (as global perceptual features) obtained from bottom-up saliency mod-
els, gradient information and local standard deviation (local features) also termed as
Root Mean Squared (RMS) contrast. The saliency map of an image represents global

information about how often a particular pixel is gazed at. On the other hand, local standard deviation and gradient information signify the local and contextual information of any pixel. The local correlation between the global information obtained from the reference and distorted images is computed. The local correlation between the gradient information obtained from the reference and distorted images is also calculated. These local correlations are combined with the local RMS contrast between the images. From the experimental results, we find that the integration of simple visual details of global perceptual difference information and local information may result in an effective FR-IQA technique. It differs from its predecessors in terms of treatment of local and global features by using the regional correlation. It has been shown in [52], gradient is structure-variant as well as contrast-variant. Thus similar variations in gradient magnitude values and standard deviation are expected for a pixel. However, change in standard deviation may not be caused by change in gradient magnitude only. Gradient orientation is also affected by presence of distortion. Thus, the proposed approach applies all of these visual details to arrive at the quality score. The performance analysis of the technique in six benchmark databases shows the promise of the proposed method as a competitive technique in FR-IQA. Also, we carry out analysis on distortion wise performance of the FR-IQA techniques using color based representation. This representation of the results clearly shows that with certain distortions, the FR-IQA techniques fail to perform well. The representation also depicts the competitive performance of the proposed method as also analyzed in Section 4.4.3.

4.2 Motivating Factors

The main motivation behind the proposed approach is to successfully combine simple visual cues, representing the global and the local information present in an image, to formulate a competitive FR-IQA technique. We have selected saliency maps as

global perceptual features for an image. For local features, block based gradients and standard deviations are used. The proposed technique obtains global distortion information by comparing global features obtained from the reference and distorted images using local correlation. The local distortion information is obtained by comparing local features using local correlation and local difference. In the following sections, we describe the importance of each of the global features and the roles they are expected to play as parts of an FR-IQA technique. Based on their properties, we hypothesize that the effective combination of these simple features can result in a competitive FR-IQA technique.

4.2.1 Global Features

The saliency map of an image is chosen as a representation of the global perceptual features of an image. Since saliency maps have perceptual information contained in them, they have been used as features and for pooling purposes [63, 128]. Given saliency maps, it becomes easier to point out the pixels which are perceptually more important than the others as they are more significantly gazed at. Many saliency methods rely on bottom-up (task-independent and data driven) processing only. Some techniques also apply top-down (task-driven, prior knowledge) approaches to enhance saliency [17]. A comparison of several types of spectral saliency based FR-IQA is presented in Chapter 3, Section 3.3. We chose the bottom-up technique for saliency calculation called spectral residual (SR) [32] and phase Fourier transform (PFT) [27] to be used in our method owing to their simplicity, fast computation, high average performance and applicability with both grayscale and color images (shown in Chapter 3, Section 3.3). Thus, we present two versions of the proposed method using SR and PFT separately. These versions are named as global local distortion with SR (GLD-SR) and global local distortion with PFT (GLD-PFT).

The saliency maps shown in the corresponding images as global features during our discussion of the motivation, are computed using SR only. As shown in Figs. 4.1(d–

f), saliency maps of different images (see Figs. 4.1(a–c)) having different distortions look similar. Though the points of gazes remain similarly distributed visually, subtle changes in values occur as shown in the Figs. 4.1(g–h)). Hence, the saliency maps do not vary in similar way for different distortions. Though, the face of the girl remains among the most salient parts in all the images, the changes in saliency values at various pixels are different. Thus, global perceptual information varies for different distortions and this change is conveyed by the saliency maps. This property of the saliency maps is used in the proposed technique. The local correlations of the saliency maps are expected to be higher if the variations of the saliency values in the block-wise neighborhood of the same pixel are similar. However, this similarity indicates similar variation in the global perceptual information for the neighborhood but same values of saliency may not be present in both of the corresponding blocks of the images. Also, local distortions may be ignored by the global representation. Hence, the use of local visual cues remains important. Local standard deviation and gradient information are used as local visual information and these are elaborated in the following section.

4.2.2   Local Features

As experimentally validated in the work of Bex and Makous [3], RMS contrast can be used to detect contrast changes. Therefore, RMS contrast has been chosen to extract local features for the proposed work. As per the definition of RMS contrast, it is similar to the standard deviation of luminance values. In Fig. 4.2, we show the RMS contrast maps for the reference and distorted images. Some structure differences are present between the images in Figs. 4.2(d) and (e) as revealed in their difference image Fig. 4.2(m). On the other hand, the dynamic range of RMS contrast is lower in Fig. 4.2(f) compared to that in Fig. 4.2(e) and hence their difference shown in Fig. 4.2(n) bears much similarity with Fig. 4.2(e). The gradient magnitude image in Fig. 4.2(g) shows the outlines of the blocks of the image shown in Fig. 4.2(a). The difference of gradient magnitude of the jpeg image and that of the

Figure 4.1: (a) and (c) are the jpeg and global contrast decremented images obtained from the reference image (b) [42]. The saliency maps of (a), (b) and (c) are shown in (d), (e) and (f) respectively. The difference of the saliency maps shown in (e) and (d) is (g). The difference of the saliency maps shown in (e) and (f) is (h).

reference image is shown in Fig. 4.2(o). This difference reveals smaller structural changes compared to Fig. 4.2(m). For the global contrast decremented image, the difference in gradient magnitude image Fig. 4.2(p) reveals magnitude changes (lower dynamic range) but this image is much sharper than Fig. 4.2(n). The gradient orientation image in Fig. 4.2(j) has changed significantly from Fig. 4.2(k). For jpeg image, significant change in gradient orientation can be seen through the difference image shown in Fig. 4.2(q). However, for contrast decremented image, the change in gradient orientation is very less as shown in Fig. 4.2(r). Thus, the RMS contrast, gradient magnitude and gradient orientation maps have different visual depictions for different distortions as shown in the third row of the Fig. 4.2 and each of these local features has its own relevance.

4.2.3   Correlation between the Feature Maps

We have used cross-correlation between the saliency maps, x-partial derivative maps and y-partial derivative maps to find the relative local variations of the selected features. The cross-correlation conveys whether the local and global information we employ in our technique, cause uniform or varying changes throughout different neighborhoods of the pixels. A block-wise correlation between the saliency maps is sure to indicate how similarly the saliency varies in the reference and the distorted images in the same neighborhood/region. Thus, from global perspective, we have the idea about the variation of saliency in the respective local neighborhood of a pixel. Similarly, for the x and y-partial derivatives of the local data, we get the varying trends in neighborhoods along the x and y directions. Any edge present in an image is sure to have a vertical and a horizontal component. These components are the x and y-partial derivatives. Thus, the correlation between x-partial derivatives of the reference and distorted images will depict the variation in neighborhood due to changes in edge strengths in vertical direction. Similarly, in horizontal direction, the correlation between y-partial derivatives is important. In Fig. 4.3, the importance of correlation

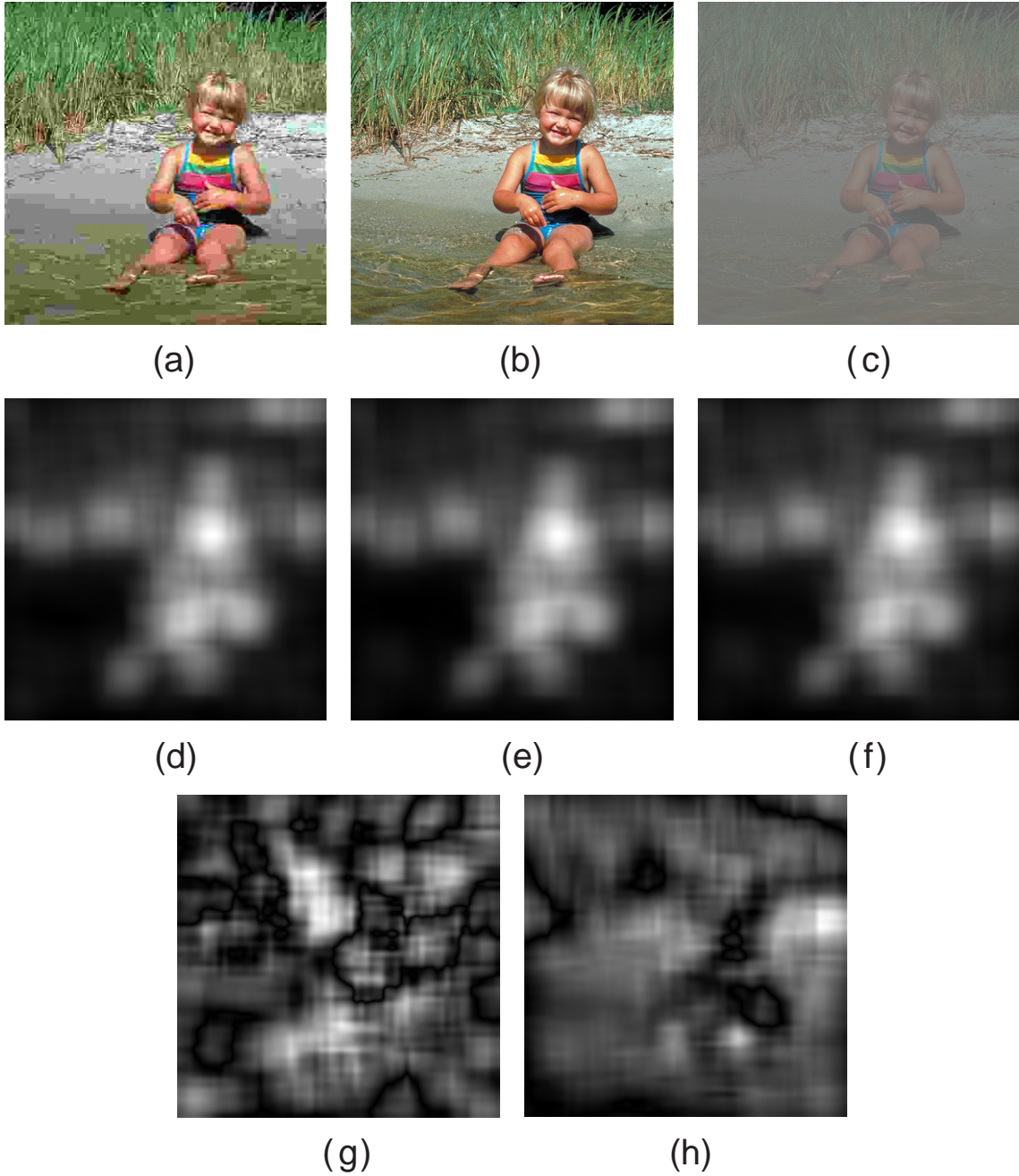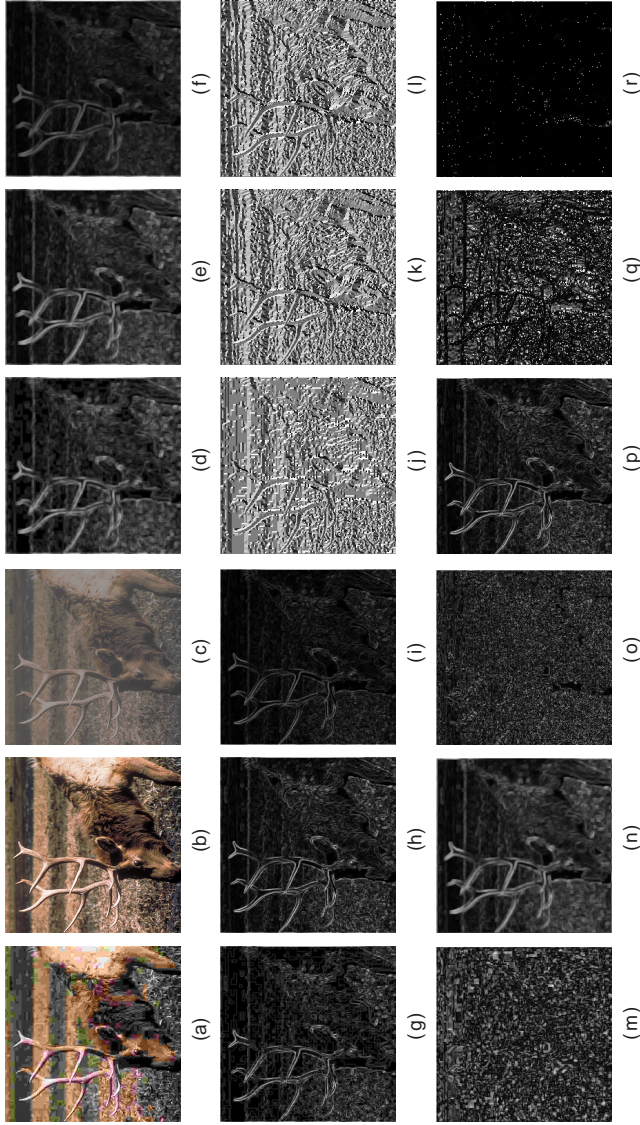Figure 4.2: In the first row, (a) and (c) are the jpeg and global contrast decremented images obtained from the reference image (b) [42]. The RMS contrast images corresponding to (a), (b), (c) are presented in (d), (e) and (f) respectively. The second row represents the corresponding gradient magnitude images (g, h, i) and the gradient orientation images (j, k, l). In the third row (m), (o), (q) represent the difference images calculated between the RMS contrast, gradient magnitude and orientation images of the reference and jpeg images, respectively. (n), (p), (r) represent the difference images calculated between the RMS contrast, gradient magnitude and orientation images of the reference and global contrast decremented images, respectively.

maps are shown for the distortion jpeg. The block-wise correlation between the x-partial derivatives (Figs. 4.3(e) and (g)) is found out for the reference and distorted images (Figs. 4.3(a) and 4.3(b), respectively). Similarly, the block-wise correlation between y-partial derivatives (Figs. 4.3(f) and (h)) is calculated. Therefore, two correlation maps are obtained. The maximum and minimum of these correlation maps are used to form the images shown in Figs. 4.3(j) and (k), respectively. The difference between these two images is shown in Fig. 4.3(l) and it shows to what extent the local data based maximum and minimum correlation varies. This figure depicts all the edges of blocks formed in the distorted image and hence is indicative of some structural changes.

4.3   Proposed Method

We describe the proposed method in details in this section. We sub-divide the section into three parts. The first part describes the computation of global and local feature information from the distorted images. The second part describes the use of these information to form distortion map and computation of the quality score while the third part analyzes the method using an example. The graphical depiction of the proposed method is presented in Fig. 4.4.

4.3.1   Global and Local Feature Extraction

The reference and distorted/test images I_R and I_T are pre-processed to obtain the images PI_R and PI_T respectively. The pre-processing steps are: (1) any color image is converted to grayscale; (2) the grayscale image is subjected to the process of automatic scale selection as mentioned in [110] and (3) the range of image intensity values is restricted between 0 and 1. In the next step, we calculate the saliency maps, gradient maps and local contrast maps. The saliency maps (S_R and S_T) of PI_R and PI_T are
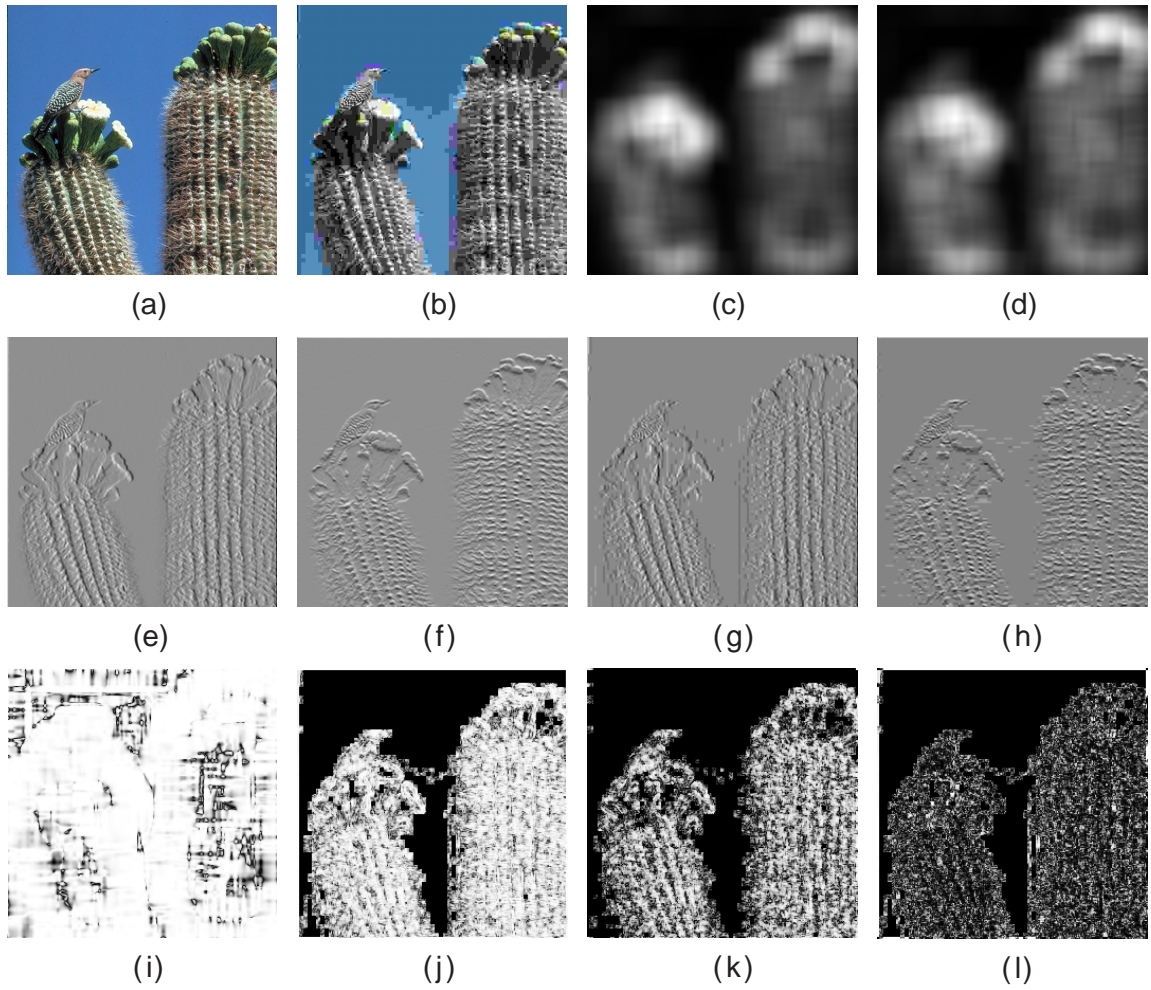
Figure 4.3: In the first row, (a) and (b) are the source and the jpeg images [42]; (c) and (d) are the respective saliency maps; (e) and (g) are the respective x - gradient maps while (f) and (h), the respective y-gradient maps. (i) is the block-wise correlation between (c) and (d); (j) and (k) are formed using the maximum and minimum of the block-wise correlation between (e) and (g) and that between (f) and (h). (l) represents the pixelwise differences between (j) and (k).
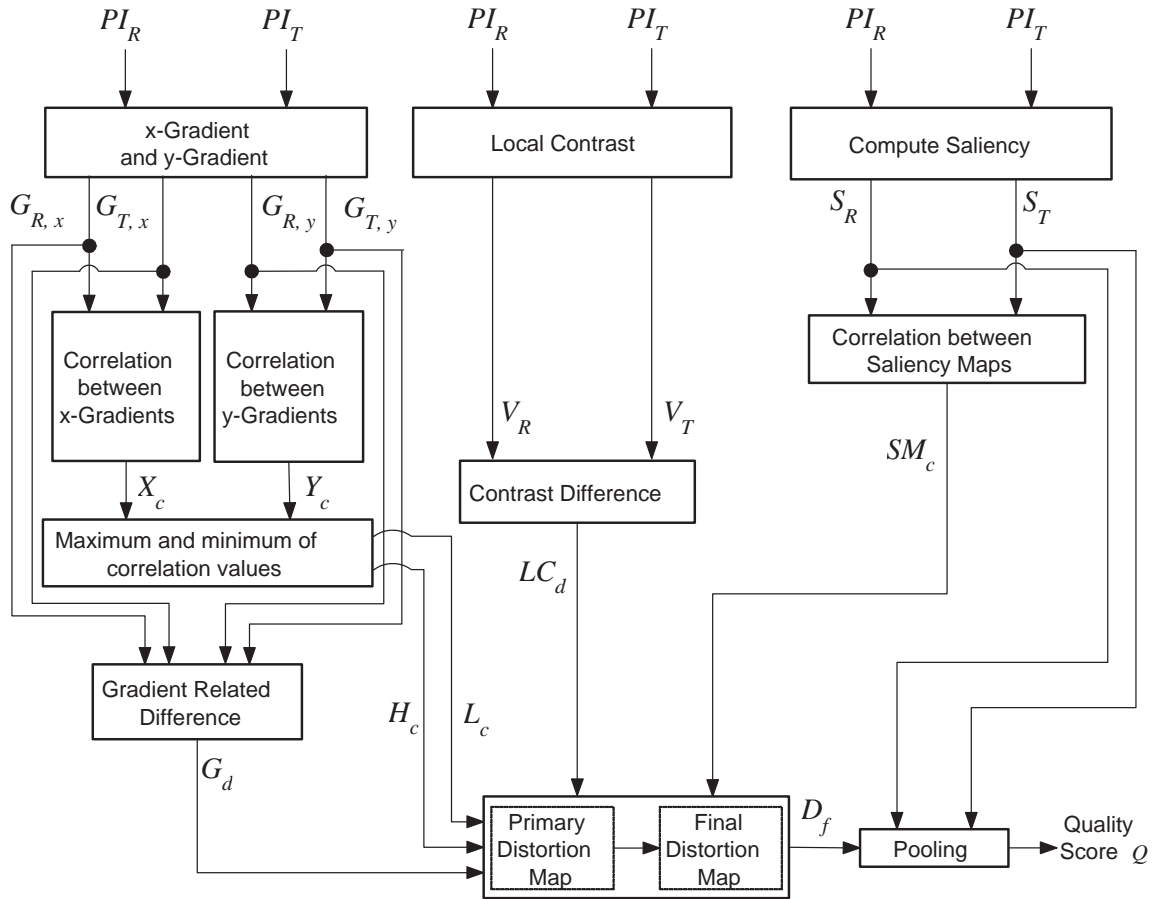
Figure 4.4: A schematic diagram of the proposed framework used in GLD-SR and GLD-PFT.

found by spectral residual method [32]. The local contrast is calculated as

$$V_R(p) = \left[\frac{1}{MN}\sum_{\nu\in\omega}(PI_{R,\nu} - \mu_\omega)^2\right]^{(0.5)}, \tag{4.1}$$

where μ_ω is the mean of the block ω of size $M \times N$ surrounding the pixel p of image PI_R. Similarly, one can compute the local contrast image V_T from PI_T. Then, the local contrast difference is found out as

$$LC_d(p) = \left[\frac{(V_R(p) - V_T(p))}{2}\right]^2. \tag{4.2}$$

For the local contrast and gradient calculation, we use a 3×3 window. The x-gradient ($G_{R,x}$ and $G_{T,x}$) and y-gradient ($G_{R,y}$ and $G_{T,y}$) images are found by the Scharr gradient operator [35]. Next, we find the local correlation image SM_c between the global saliency maps. From the gradient images, the gradient magnitudes are found out as

$$G_{R,M}(p) = (G_{R,x}^2(p) + G_{R,y}^2(p))^{\frac{1}{2}}, \tag{4.3a}$$

$$G_{T,M}(p) = (G_{T,x}^2(p) + G_{T,y}^2(p))^{\frac{1}{2}}. \tag{4.3b}$$

The gradient orientations are calculated as

$$G_{R,O}(p) = \arctan(G_{R,y}(p)/G_{R,x}(p)), \tag{4.4a}$$

$$G_{T,O}(p) = \arctan(G_{T,y}(p)/G_{T,x}(p)), \tag{4.4b}$$

such that the angles lie within $[-\pi, \pi]$. We calculate the gradient related difference as

$$G_d(p) = \left[\frac{\max\left(\frac{|G_{R,M}(p)-G_{T,M}(p)|}{\sqrt{2}}, \frac{|G_{R,O}(p)-G_{T,O}(p)|}{2\pi}\right)}{2}\right]^2 \tag{4.5}$$

which considers the maximum of the differences of gradient magnitudes and gradient orientations by using the function max(.). Since, the intensity values in the images are restricted between 0 and 1, the gradient magnitudes can have a maximum value

of $\sqrt{2}$. Hence, these magnitudes are divided by $\sqrt{2}$ such that their values remain below 1. Again, the maximum difference between the orientation angles can be 2π. Therefore, the orientation difference is divided by 2π such that its maximum value is restricted to 1.

4.3.2   Formation of Distortion Map

In order to form the distortion map first, the three local cross-correlation maps are found out between the pair of saliency maps, x-partial derivative maps and y-partial derivative maps. Thus, we have SM_c, X_c and Y_c denoting the local correlation maps corresponding to saliency maps, x-gradient maps and y-gradient maps respectively. Now SM_c is the correlation between the global features, and high value of this correlation implies greater global similarity between the images. High values of correlation for the local information imply similar local variation between the images. The maximum and minimum values of correlation between the local features for any pixel p are computed as $H_c(p) = \max(X_c(p), Y_c(p))$ and $L_c(p) = \min(X_c(p), Y_c(p))$ respectively. Now, we calculate the primary distortion measure for each pixel p as

$$D_p(p) = \frac{\max(|H_c(p) - L_c(p)|, (1 - X_c(p)), (1 - Y_c(p)), (1 - SM_c(p)))}{2} \times T(p), \quad (4.6)$$

where

$$T(p) = \left[LC_d(p) \times \frac{(1 - SM_c(p))}{2} \times G_d(p) \right]^{\frac{1}{3}}. \quad (4.7)$$

Therefore, D_p indicates a measure of difference between two pixels considering their local neighborhood as well as global perceptual importance. The first difference term within the function max(.) in Eqn. 4.6 shows the difference in correlation caused by varying changes in the neighborhoods along the horizontal and vertical directions. The correlation differences between the local contrast and gradient difference are also used to account for the local contrast and gradient changes. The difference in local contrast and gradients are also considered to find their variation together. The correlation ranges from -1 to 1. Hence, when any correlation term is subtracted from

1, the maximum value it can take is 2. Hence, a division by 2 is carried out. Thus, the maximum deviation of any type of correlation is captured in this equation. Regarding the calculation of $T(p)$ which uses the change in local distortion obtained from the RMS contrast and gradient difference, if there exists high correlation within the saliency maps, the difference/distortion values get effectively lowered. Therefore, multiplication by $(1 - SM_c(p))$ has been used. Hence, the product of these three terms are taken in Eqn. 4.7. The values of $T(p)$ will be larger when aforesaid differences are higher along with smaller values of $SM_c(p)$ (smaller values of $SM_c(p)$ imply significantly different local variations in global information of the reference and distorted image). $T(p)$ is multiplied by the maximum change captured through correlations calculated from global and local features. In this primary map, the treatment with all pixels remains the same. However, this map can undermine local variations in images as it relies more on the global variations.

Therefore, one more map is used based on the relative values of the correlation between the local and global data. Higher correlation in the global features compared to the local features does not indicate that the distortion is low, but local variations in luminance/RMS contrast should be regulated or considered, if needed. Therefore, the following map is formed for some selected pixels which are more correlated globally than locally. We find the set η_G of those pixels p for which the correlation between the global features is greater than that between local features, meaning $SM_c(p) > L_c(p)$. The distortion measure for them is calculated in two ways: first, the changes in RMS contrast are regulated and next, only the local differences are considered. For the first part,

$$A(p) = \begin{cases} \left(LC_d(p) \times \frac{(1 - SM_c(p))}{2} \right)^{\frac{1}{2}}, & \text{if } p \in \eta_G \\ 0, & \text{otherwise} \end{cases}. \tag{4.8}$$

This equation says that if the correlation between global features is higher than the minimum local correlation obtained from gradients, the local contrast difference has to be given varied importance as decided by $(1 - SM_c(p))$.

On the contrary, local distortions present in some pixels in set η_G may be significant if the local changes are high . Hence, the product of RMS contrast difference and gradient difference are considered for the second part.

$$B(p) = \begin{cases} (LC_d(p) \times G_d(p))^{\frac{1}{2}} & \text{if } p \in \eta_G \\ 0, \text{ otherwise} \end{cases} . \tag{4.9}$$

The final distortion map is calculated as

$$D_f(p) = D_p(p) + A(p) + B(p). \tag{4.10}$$

Finally, the quality score Q is calculated as

$$Q = k \frac{\sum\limits_{p=1}^{N_I} D_f(p) \times \max(S_R(p), S_T(p))}{\sum\limits_{p=1}^{N_I} \max(S_R(p), S_T(p))}, \tag{4.11}$$

where N_I is the total number of pixels in PI_R and saliency map values are used for the pooling the map. k is a constant taken as 10000 in all our experiments. It scales the objective scores which are otherwise very low. Multiplication by this constant does not have impact on the quantitative results rather it makes the interpretation of the scores easier. In the next section, we analyze the proposed method using the Q values obtained on a suite of distorted images.

4.3.3 Analysis using an Example

In the diagram shown in Fig. 4.5, we find the distorted images with their corresponding primary distortion maps and saliency weighted final distortion maps computed using GLD-SR. The dynamic ranges of these maps have been adjusted to increase visual details. The brighter an area is in the distortion map, the higher is its contribution in the distortion map. The distortion map indicates the assessment of degradation of the image prior to the pooling process. Also, this figure shows the difference between

the primary and the final distortion maps. For example, the relative contrast of the values is lower in the final map than the primary map. This is as if taking 'one more look' at the local distortions before evaluating the image. Hence, the final distortion map is necessary to calculate. The averages of the saliency weighted final distortion maps (the Q values) in Figs. 4.5(m–r), have increased gradually demonstrating the increasing degradation of perceptual quality as we move from left to right. As the images are arranged in increasing order of their DMOS values (provided in the CSIQ database), the objective scores should be in the same order as well. Since the proposed technique is a way to quantify the perceptual degradation, values of Q increase as perceptual degradation increases. The values of Q are found to be increasing in the same order as the DMOS values. We find that the proposed technique is able to maintain the same order of perceptual quality as given by human observers for different types of degradations. We find that only parameter the proposed technique has, is the window size which is fixed for all our experiments. Increasing window size has no significant influence on the performance but it increases the computational time.

4.4 Experiments and Performance Analysis

We are presenting the experiments and results of our proposed method in this section. The experiments are carried out in six leading databases. The proposed method has been compared with 12 other methods: PSNR, SSIM, MS-SSIM, VIF, VSNR, MAD, IWSSIM, FSIM$_c$, ADM, SR-SIM and IGM and PDSESIM. We have gradually moved from global representation of experiments and results to a much detailed one. In the first experiment, the average performance of the techniques across all distortions and databases are considered. In the next experiment, we look at individual performance of the FR-IQA techniques in 6 databases considering all distortions in a database together, along with a statistical analysis using F-test. Finally, we look at

0.3706 (a)   0.4675 (b)   0.5652 (c)   0.5758 (d)   0.6871 (e)   0.7497 (f)

(g)   (h)   (i)   (j)   (k)   (l)

5.8864 (m)   10.7984 (n)   17.2687 (o)   17.6331 (p)   31.5596 (q)   102.1301 (r)
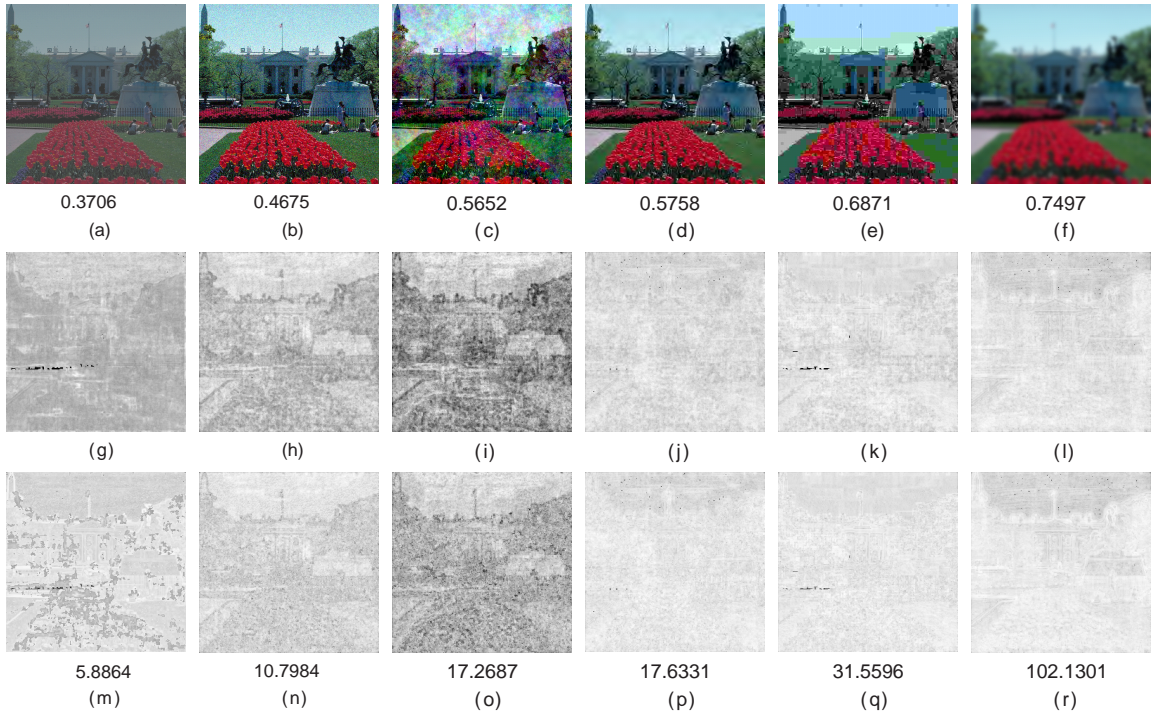
Figure 4.5: The objective values obtained by the proposed method GLD-SR on six images of CSIQ database [42]. The images (from left to right) in the first row have different types of distortions: contrast decrement, awgn, fnoise, jpeg2k, jpeg and gblur. The images are arranged in increasing order of their DMOS scores and hence perceptual quality decreases from (a) to (f). The respective DMOS values are shown below the figures in first row. The middle row shows the corresponding primary distortion maps. The final row shows the saliency weighted final distortion map which is used to calculate the objective score Q using GLD-SR. The respective Q values are shown below the figures in the last row.

the distortion wise analysis of the results in three largest databases. This three-fold experimentation helps to deeply analyze the proposed method by several aspects and shows its competitiveness to the state-of-the-art techniques.

4.4.1 Average Performance Comparison

The direct average for each evaluation measure shown in Table 4.1 is the arithmetic mean of the corresponding measure computed from six databases. The weighted average of an evaluation measure is the weighted mean of the measure values and the number of images in a database acts as its weight. The top performances for each evaluation measure in direct and weighted average are represented in bold font.We find that the proposed method GLD-SR has the highest average values in terms of all of the three measures for direct average. GLD-PFT has close performance with PDSESIM using SROCC and KROCC. The average PLCC score for GLD-PFT is better than all existing techniques and slightly below GLD-SR. For weighted average, GLD-PFT is the best followed by GLD-SR and PDSESIM using all of the three correlation scores. Thus, with the proposed methods GLD-SR and GLD-PFT, we have been able to achieve improved results compared to the state-of-the-art techniques across six databases considering all distortions together.

4.4.2 Database wise Performance Comparison

In this experiment, the overall performance of an FR-IQA technique in each of the six databases is evaluated and the results are shown in Table 4.2, using all five evaluation measures. The top three performances for each evaluation measure in each database are highlighted in bold. As we can find, none of the FR-IQA techniques are among the top three in all databases. VSNR, IWSSIM, ADM, SR-SIM and IGM are among the top performers in only one of the databases. MAD, PDSESIM and the proposed approaches GLD-PFT and GLD-SR are among the top performers in three to four

Table 4.1: Average performance comparison of GLD-SR and GLD-PFT over six databases.

Method	Measure	PSNR	SSIM	MS-SSIM	VIF	VSNR	MAD	IWSSIM	FSIM$_c$	ADM	SR-SIM	IGM	PDSESIM	GLD-SR	GLD-PFT
Direct	SROCC	0.6877	0.7608	0.8910	0.8432	0.7936	0.9168	0.8902	0.9223	0.9090	0.9206	0.9095	**0.9242**	**0.9271**	**0.9246**
Average	KROCC	0.5101	0.5831	0.7123	0.6790	0.6221	0.7549	0.7182	0.7603	0.7419	0.7580	0.7439	**0.7643**	**0.7689**	**0.7639**
	PLCC	0.7016	0.7619	0.8886	0.8543	0.7961	0.9185	0.8893	0.9212	0.9054	0.9187	0.9117	**0.9230**	**0.9278**	**0.9273**
Weighted	SROCC	0.6757	0.8356	0.8908	0.8457	0.7329	0.8971	0.8964	0.9153	0.9013	0.9172	0.9155	**0.9186**	**0.9191**	**0.9208**
Average	KROCC	0.5021	0.6455	0.7104	0.6861	0.5598	0.7326	0.7225	0.7491	0.7343	0.7541	0.7516	**0.7558**	**0.7575**	**0.7591**
	PLCC	0.6896	0.8302	0.8828	0.8736	0.7177	0.8972	0.8949	0.9087	0.9001	0.9119	0.9104	**0.9120**	**0.9193**	**0.9212**

databases; but the subset of these databases varies in each case. For the CSIQ and IVC databases, the proposed approach GLD-PFT gives best performance using all of the five evaluation measures. The performance of GLD-SR is better than the existing FR-IQA techniques in CSIQ, TID2008 and IVC databases using some of the evaluation measures. In TID2008 database, GLD-SR and GLD-PFT perform better than several state-of-the-art FR-IQA techniques using some of the evaluation measures.

To analyze the statistical effectiveness of GLD-SR and GLD-PFT, F-test has been used. The F-test is carried out using the residuals of the objective scores (after mapping) and the given subjective scores. The variance of the residuals for each database is calculated for all the FR-IQA techniques. Depending on the number of images in each database, the critical value of F-distribution F_c, is determined at 95% confidence level. For a given database, if the ratio of the larger variance to the smaller one is greater than F_c, the FR-IQA method with the smaller variance is significantly better than the one with larger variance. The result of the F-test is shown in Figs. 4.6 and 4.7. Fig. 4.6 shows the statistical comparison of GLD-SR with other thirteen techniques. Each cell in the figure with a numeric value has values containing '-1','0' or '1' only. '0' implies that GLD-SR is not statistically different from the corresponding FR-IQA technique given in the row heading for the database provided in the column heading. '1' implies that GLD-SR is better than the FR-IQA technique in the given databases indicated by the row and column headings respectively. Finally, '-1' implies significantly worse performance. We have also colored the cells depending on the its values (-1 : pink, 0 : grey, 1 : green) to make the comparison vivid. From the F-test results shown in Fig. 4.6, we understand that the more number of green cells signifies better performance of GLD-SR. Since, comparison of an FR-IQA technique with itself will return a '0', the row corresponding to GLD-SR is totally grey. Hence, the maximum number of green cells expected is 13*6 = 78. The number green cells obtained is 40. Hence, in 51.28% of the cases, GLD-SR has shown improvement.

Similarly, GLD-PFT has improved in 52.56% of the cases as shown in Fig. 4.7. While some performance decline is exhibited in the LIVE and MICT databases, in majority of the databases (CSIQ, TID2008, A57 and IVC) performance has improved.

4.4.3 Distortion wise Performance Comparison

In our last experiment, three databases (LIVE, CSIQ, TID2008) with the highest number of distorted images are chosen. In each of these databases, more than 4 distortions exist. Using SROCC, the performance of GLD-SR and GLD-PFT in comparison with several existing FR-IQA techniques for different distortions is demonstrated in Fig. 4.8. The top three performers for each distortion are highlighted in bold. An interesting point to be noted here: 13 methods (out of 14) have been among the top assessment methods for at least one distortion. Thus, this figure gives an insight about the suitable methods, if the distortion is fixed. VIF and GLD-SR are among the top methods for 9 distortions whereas, SR-SIM and IGM are among top methods for 10 and 14 distortions respectively. For several types of noises in the TID2008 database, PSNR performs better than the other techniques.

Apart from this, each cell with a numeric value is highlighted with a specific color depending on which group the value falls within. To elaborate, the background color in each cell depends on the value contained in the cell. As per the convention, absolute values of SROCC are shown and divided in four ranges. Hence, the range of absolute SROCC values [0,1] is divided in four non-overlapping groups: group 1 [0 - 0.7999] is represented by pink, group 2 : [0.8 - 0.8999] is presented by blue, group 3 [0.9 - 0.9499] is presented by orange and group 4 [0.95 - 1] is presented by green. The grouping is made non-uniform so as to segregate the strong performances in each distortion from the moderate and weak performances. First, we take a look at the results of TID2008 database. For the distortions like pattern noise, block-wise distortion, intensity shift and contrast change, the SROCC values for all of the techniques are very low. On the other hand, most of the techniques have higher SROCC scores for distortions

Table 4.2: Performance evaluation and comparison of GLD-SR and GLD-PFT in six databases.

Database	Measure	PSNR	SSIM	MS-SSIM	VIF	VSNR	MAD	IWSSIM	FSIM$_c$	ADM	SR-SIM	IGM	PDSESIM	GLD-SR	GLD-PFT
LIVE	SROCC	0.8756	0.9479	0.9513	0.9636	0.6481	**0.9669**	0.9566	**0.9645**	0.9460	0.9618	0.9581	**0.9652**	0.9624	0.9631
	KROCC	0.6865	0.7963	0.8049	0.8282	0.4879	**0.8421**	0.8178	**0.8365**	0.7976	0.8299	0.8250	**0.8378**	0.8290	0.8297
	PLCC	0.8722	0.9449	0.9489	0.9604	0.7018	**0.9675**	0.9519	**0.9613**	0.9359	0.9552	0.9567	**0.9616**	0.9506	0.9521
	MAE	10.483	6.9109	6.6701	6.0952	14.6406	**5.2071**	6.3805	**5.8236**	7.1836	6.3263	6.0420	**5.8339**	6.6008	6.4812
	RMSE	13.3659	8.9437	8.6143	7.6089	19.5193	**6.9068**	8.3757	**7.5236**	9.6222	8.0792	7.9557	**7.4929**	8.4907	8.3667
CSIQ	SROCC	0.8057	0.8368	0.9132	0.9194	0.8132	**0.9466**	0.9212	0.9309	0.9333	0.9319	0.9403	0.9359	**0.9539**	**0.9549**
	KROCC	0.6078	0.6325	0.7386	0.7532	0.6279	**0.7963**	0.7522	0.7684	0.7710	0.7719	0.7874	0.7786	**0.8091**	**0.8108**
	PLCC	0.7999	0.8144	0.8986	0.9277	0.8016	**0.9505**	0.9142	0.9186	0.9284	0.9241	0.9267	0.9244	**0.9506**	**0.9515**
	MAE	0.1193	0.1144	0.0857	0.0742	0.1138	**0.0631**	0.0789	0.0745	0.0758	0.0734	0.0698	0.0703	**0.0619**	**0.0612**
	RMSE	0.1576	0.1525	0.1168	0.0980	0.1590	**0.0817**	0.1082	0.1060	0.0977	0.1021	0.1020	0.1042	**0.0817**	**0.0810**
TID2008	SROCC	0.5245	0.8081	0.8528	0.7496	0.7046	0.8340	0.8559	0.8840	0.8617	**0.8913**	**0.8901**	**0.8892**	0.8817	0.8849
	KROCC	0.3696	0.6056	0.6543	0.5863	0.5340	0.6445	0.6636	0.6991	0.6842	**0.7149**	**0.7103**	**0.7089**	0.7001	0.7030
	PLCC	0.5545	0.8061	0.8419	0.8075	0.6507	0.8311	0.8572	0.8758	0.8683	**0.8854**	0.8842	0.8807	0.8879	**0.8909**
	MAE	0.8918	0.6139	0.5616	0.5837	0.7085	0.5543	0.5245	0.4868	0.4883	**0.4543**	0.4647	0.4716	0.4627	**0.4611**
	RMSE	1.1168	0.7948	0.7247	0.8007	1.0274	0.7491	0.6915	0.6482	0.6664	**0.6246**	0.6279	0.6360	0.6182	**0.6099**
A57	SROCC	0.6189	0.4059	0.8435	0.6224	**0.9355**	0.9023	0.7750	0.9181	0.8725	0.9295	0.9006	**0.9299**	0.9351	0.9100
	KROCC	0.4309	0.2780	0.6529	0.4592	**0.8031**	0.7233	0.5880	0.7639	0.6912	0.7779	0.7387	**0.7793**	0.7891	0.7457
	PLCC	0.6347	0.4149	0.8394	0.6137	**0.9500**	0.9043	0.7652	0.9252	0.8803	0.9247	0.9192	**0.9341**	0.9354	0.9217
	MAE	0.1607	0.1847	0.1119	0.1417	**0.0574**	0.0856	0.1182	0.0794	0.0914	0.0778	0.0749	0.0754	**0.0655**	0.0729
	RMSE	0.1899	0.2238	0.1337	0.1957	**0.0770**	0.1051	0.1587	0.0933	0.1167	0.0936	0.0968	0.0877	**0.0869**	0.0954
MICT	SROCC	0.6132	0.7870	0.8874	0.9077	0.8608	**0.9362**	0.9202	0.9067	**0.9378**	0.8825	0.8653	0.8963	0.8988	0.9010
	KROCC	0.4443	0.5922	0.7029	0.7315	0.6745	**0.7823**	0.7537	0.7303	**0.7820**	0.6975	0.6734	0.7175	0.7180	0.7210
	PLCC	0.6498	0.7996	0.8924	0.9136	0.8698	**0.9367**	0.9246	0.9072	**0.9415**	0.8870	0.8712	0.8989	0.9060	0.9084
	MAE	0.7746	0.5649	0.4287	0.4012	0.4599	**0.3493**	0.3653	0.4015	**0.3288**	0.4444	0.6149	0.4201	0.4017	0.3961
	RMSE	0.9517	0.7551	0.5650	0.5096	0.6193	**0.4384**	0.4775	0.5275	**0.4225**	0.5782	0.4811	0.5491	0.5302	0.5237
IVC	SROCC	0.6884	0.7789	0.8980	0.8964	0.7993	0.9146	0.9125	**0.9293**	0.9026	0.9265	0.9027	0.9286	0.9309	**0.9336**
	KROCC	0.5218	0.5939	0.7203	0.7158	0.6053	0.7406	0.7339	0.7636	0.7255	0.7560	0.7286	0.7639	0.7683	**0.7731**
	PLCC	0.6983	0.7915	0.9106	0.9026	0.8026	0.9210	0.9228	0.9390	0.8783	0.9357	0.9125	0.9381	0.9361	**0.9391**
	MAE	0.6875	0.5297	0.3701	0.4050	0.5490	0.3673	0.3684	**0.3277**	0.4471	0.3404	0.3702	0.3228	0.3332	**0.3242**
	RMSE	0.8770	0.7593	0.5103	0.5276	0.7325	0.4753	0.4704	**0.4203**	0.5885	0.4306	0.5036	0.4246	0.4297	**0.4198**

	LIVE	CSIQ	TID2008	A57	MICT	IVC
PSNR	1	1	1	1	1	1
SSIM	0	1	1	1	1	1
MS_SSIM	0	1	1	1	0	1
VIF	-1	1	1	1	0	1
VSNR	1	1	1	0	1	1
MAD	-1	0	1	0	-1	0
IWSSIM	0	1	1	1	0	0
FSIM$_c$	-1	1	1	0	0	0
ADM	1	1	1	1	-1	1
SR-SIM	0	1	0	0	0	0
IGM	-1	1	0	0	1	1
PDSESIM	-1	1	0	0	0	0
GLD-SR	0	0	0	0	0	0
GLD-PFT	0	0	0	0	0	0

Figure 4.6: F-ratio test for comparing GLD-SR with thirteen other FR-IQA techniques in six databases.

	LIVE	CSIQ	TID2008	A57	MICT	IVC
PSNR	1	1	1	1	1	1
SSIM	1	1	1	1	1	1
MS_SSIM	0	1	1	1	0	1
VIF	-1	1	1	1	0	1
VSNR	1	1	1	0	1	1
MAD	-1	0	1	0	-1	0
IWSSIM	0	1	1	1	0	0
FSIM$_c$	-1	1	1	0	0	0
ADM	1	1	1	0	-1	1
SR-SIM	0	1	0	0	0	0
IGM	0	1	0	0	1	1
PDSESIM	-1	1	1	0	0	0
GLD-SR	0	0	0	0	0	0
GLD-PFT	0	0	0	0	0	0

Figure 4.7: F-ratio test for comparing GLD-PFT with thirteen other FR-IQA techniques in six databases.

like blur, denoising, jpeg2k and jpeg. PSNR performs much better than most of the techniques in awgn, awgn-pink, spatial-corr-noise, masked noise, high frequency noise, impulse noise and quantization noise. For jpeg-trans-error and jpeg2k-trans-error, we have mixed results from the techniques but a strong performing technique is still missing. On the other hand, in CSIQ database, some techniques have lower SROCC values for contrast but other distortions are well dealt by most of them. Only SR-SIM is found to have strong performance in all distortions in CSIQ database. Most techniques have high SROCC scores for all distortions in LIVE database except for VSNR and PSNR. Also, SSIM, MS-SSIM, VIF, $FSIM_c$, PDSESIM and GLD-PFT have top performances for all distortions in LIVE database as indicated by the green color. Most techniques perform very well in LIVE database. Except for f-noise in the CSIQ database, all distortions present in the TID2008 database are present in LIVE and CSIQ databases and most techniques have high SROCC values for the common distortions present in LIVE, CSIQ and TID2008 databases. However, the percentage of green cells shows a huge decline when LIVE and CSIQ databases are compared with the TID2008 database.

From the above discussion, one can conclude that there is no assurance of better overall performance in a database, if the distortion wise performance is better. This is because, relative assessment of different distortions will be required together for overall performance. If the objective scores obtained from different distortions have different ranges, the overall performance in the database is likely to be poor. As we have seen, MAD has the best overall performance in LIVE database. Though MAD has strong performance in four out of five distortions in LIVE database while several others techniques have five strong performances, it still performs best when all of the distortions are considered together. A closer look at the results show that the proposed techniques follow a similar trend. GLD-SR and GLD-PFT have strong performance in CSIQ database for four out of six distortions. However, their overall performance is better than SR-SIM (6 strong performances), IGM (5 strong

performances) as well as MAD, VIF, IWSSIM and ADM in CSIQ database. Hence, both overall and distortion wise performance analysis are necessary to assess an FR-IQA technique.

4.4.4 Discussion on the Parameters

A look into the details of the methods reveals about the adjustable parameters present in the models. For PSNR, the definition is fixed and no adjustable parameters are involved. For SSIM, window size is one of the parameters and based on that local window size the structure, contrast and luminance similarity maps are generated. This way of generating similarity maps is followed in $FSIM_c$ and SR-SIM as well. It is shown in Chapter 3, Section 3.3 that this framework introduces parameters based on the number of similarity maps calculated and hence increases the effective number of adjustable parameters for the method. MS-SSIM has parameters similar to SSIM and the number of scales used also affects. However, number of scales and scale wise weights are fixed for MS-SSIM. For VIF, different types of modeling techniques are employed and in course of that, window size, number of Wavelet scales used and noise variance of modeling neural distortions are introduced as the adjustable parameters. MAD uses log-Gabor decompositions in which the scales and orientations need to be pre-defined while VSNR has the number of scales as the only adjustable parameter. ADM uses Wavelet decomposition to separate the image into two parts using 4 scales to compute the necessary parts. IGM uses auto-regression based models to generate the orderly and disorderly portions. The window size is a parameter as well as another parameter is introduced by the constant required in the luminance comparison maps. The saliency based method SR-SIM has window size and constants required by the similarity maps as the adjustable parameters. Similarly, for PDSESIM, tunable parameters are introduced by its similarity calculation framework.

Based on the discussion presented, we find that apart from PSNR, the parameter dependency is evident among the existing FR-IQA techniques. The proposed method

Database	Distortion	PSNR	SSIM	MS-SSIM	VIF	VSNR	MAD	IWSSIM	FSIM$_c$	ADM	SR-SIM	IGM	PDESIM	GLD-SR	GLD-PFT
LIVE	jpeg2k	0.8954	0.9614	0.9628	0.9696	0.5501	0.9677	0.9649	0.9723	0.9705	0.9701	0.9679	0.9706	0.9730	0.9720
	jpeg	0.8809	0.9764	0.9814	0.9846	0.6272	0.9763	0.9808	0.9840	0.9795	0.9823	0.9808	0.9844	0.9827	0.9815
	wn	0.9854	0.9694	0.9733	0.9857	0.8182	0.9844	0.9667	0.9716	0.9822	0.9810	0.9877	0.9802	0.9861	0.9852
	gblur	0.7823	0.9517	0.9543	0.9728	0.6403	0.9464	0.9719	0.9709	0.9652	0.9660	0.9529	0.9690	0.9499	0.9512
	fastfading	0.8907	0.9556	0.9471	0.9650	0.6813	0.9569	0.9442	0.9520	0.9177	0.9465	0.9194	0.9529	0.9541	0.9549
CSIQ	awgn	0.9363	0.9256	0.9471	0.9575	0.9254	0.9542	0.9380	0.9360	0.9583	0.9629	0.9631	0.9543	0.9512	0.9514
	gblur	0.9291	0.9245	0.9711	0.9744	0.9445	0.9681	0.9781	0.9728	0.9725	0.9766	0.9726	0.9741	0.9787	0.9781
	global-contrast	0.8623	0.7399	0.9528	0.9345	0.8714	0.9210	0.9540	0.9439	0.9508	0.9530	0.9548	0.9486	0.9300	0.9352
	fnoise	0.9338	0.8925	0.9330	0.9509	0.9082	0.9568	0.9057	0.9369	0.9487	0.9519	0.9427	0.9526	0.9423	0.9387
	jpeg2k	0.8879	0.9219	0.9631	0.9703	0.9172	0.9614	0.9660	0.9662	0.9666	0.9669	0.9683	0.9676	0.9654	0.9652
	jpeg	0.9361	0.9205	0.9682	0.9671	0.9485	0.9752	0.9682	0.9703	0.9748	0.9772	0.9782	0.9760	0.9776	0.9769
TID2008	awgn	0.9115	0.8310	0.8094	0.8800	0.7728	0.8388	0.7869	0.8758	0.8755	0.8990	0.9073	0.8985	0.8810	0.8792
	awgn-pink	0.9068	0.8134	0.8064	0.8785	0.7793	0.8258	0.7920	0.8931	0.8035	0.8954	0.8924	0.9022	0.8210	0.8245
	spatial-corr-noise	0.9229	0.8438	0.8195	0.8703	0.7665	0.8678	0.7714	0.8711	0.8668	0.9083	0.9121	0.9039	0.9083	0.9053
	masked noise	0.8487	0.7561	0.8155	0.8698	0.7295	0.7336	0.8087	0.8264	0.8007	0.7869	0.8016	0.7659	0.7768	0.7784
	high frequency noise	0.9323	0.8919	0.8685	0.9075	0.8811	0.8864	0.8662	0.9156	0.9213	0.9197	0.9214	0.9219	0.9044	0.9000
	impulse noise	0.9177	0.7072	0.6868	0.8331	0.6471	0.0650	0.6465	0.7719	0.8100	0.7667	0.8167	0.7925	0.7715	0.7715
	quantization noise	0.8699	0.8745	0.8537	0.7956	0.8270	0.8160	0.8177	0.8726	0.8799	0.8364	0.8787	0.8622	0.8608	0.8660
	gblur	0.8682	0.9596	0.9607	0.9546	0.9330	0.9197	0.9636	0.9472	0.9573	0.9549	0.9683	0.9393	0.9481	0.9491
	denoising	0.9381	0.9595	0.9571	0.9189	0.9286	0.9434	0.9473	0.9618	0.9379	0.9668	0.9704	0.9660	0.9748	0.9735
	jpeg	0.9011	0.9270	0.9348	0.9170	0.9174	0.9275	0.9184	0.9294	0.9317	0.9394	0.9472	0.9341	0.9437	0.9450
	jpeg2k	0.8300	0.9723	0.9736	0.9713	0.9515	0.9707	0.9738	0.9780	0.9669	0.9807	0.9844	0.9824	0.9790	0.9788
	jpeg-trans-error	0.7665	0.8668	0.8736	0.8582	0.8055	0.8661	0.8588	0.8756	0.8754	0.8881	0.8639	0.8851	0.8843	0.8842
	jpeg2k-trans-error	0.7765	0.8707	0.8525	0.8510	0.7909	0.8394	0.8203	0.8555	0.8613	0.8903	0.8877	0.8718	0.8949	0.8946
	pattern-noise	0.5931	0.7168	0.7336	0.7608	0.5716	0.8287	0.7724	0.7514	0.6743	0.7670	0.7282	0.7627	0.7478	0.7396
	blockwise-distortion	0.5852	0.8529	0.7617	0.8320	0.1926	0.7970	0.7623	0.8464	0.7534	0.7787	0.7908	0.7321	0.8089	0.8153
	intensity-shift	0.6974	0.7575	0.7374	0.5132	0.3710	0.5161	0.7067	0.6554	0.4694	0.5727	0.4829	0.6012	0.4759	0.4735
	contrast change	0.6126	0.6329	0.6400	0.8190	0.4239	0.2723	0.6301	0.6510	0.8513	0.6483	0.6412	0.6310	0.5713	0.5802

Figure 4.8: Distortion wise SROCC score comparison of GLD-SR and GLD-PFT in LIVE, CSIQ and TID2008 databases.

is dependent on the saliency map but it does not compute the similarity map unlike SR-SIM and PDSESIM. Hence, it has no tunable parameters apart from the window size. A window size is required to calculate the local correlation between the maps and it is same as the window size used to calculate the local standard deviation and gradient information.

4.5 Chapter Summary

In this chapter, we have tried to approach the problem of FR-IQA by hypothesizing about the simultaneous top-down and bottom-up processing that take place when a human observer rates the subjective quality of an image. Global perceptual information of the image stands for the bottom-up processing. On the other hand, local visual cues as gradient information and local standard deviation are used to imitate the top-down processing which is affected by the goal of the observer to rate the image quality. The regional correlation of the global information between the reference and test images and the same from the local information are used to compute the global and local distortion values. These values are combined to arrive at the final degradation score. Experiments carried out in six databases validate the promise of the proposed approach. The promising performance of the proposed method can be attributed to combination of local and global distortion measure which is a collection of simple visual cues proven to be effective from the existing IQA research. At the same time, this method also follows the drawback of the existing FR-IQA techniques to evaluate the certain types of distortions with lesser accuracy. Thus, the future work will be aimed at the distortion independent performance of the FR-IQA techniques. Also, in the proposed approach, the luminance component has only been used to formulate the technique. In future, we would study the effect of the chrominance related information on the quality score.

Chapter 5

No-reference Perceptual Sharpness Assessment using High Frequency Content

A blind approach to evaluate the perceptual sharpness present in a natural image is proposed. Though the existing literature demonstrates a set of variegated visual cues to detect or evaluate the absence or presence of sharpness, we emphasize in the current work that high frequency content and local standard deviation can form strong features to compute perceived sharpness in any natural image, and can be considered as an able alternative for the existing cues. Unsharp areas in a natural image happen to exhibit uniform intensity or lack of sharp changes between regions. Sharp region transitions in an image are caused by the presence of spatial high frequency content. Therefore, in the proposed approach, we hypothesize that using the high frequency content as the principal stimulus, the perceived sharpness can be quantified in an image. When an image is convolved with a high pass filter, higher values at any pixel location signify the presence of high frequency content at those locations. Considering these values as the stimulus, the exponent of the stimulus is weighted by local standard deviation to impart the contribution of the local contrast within the formation of the sharpness map. The sharpness map highlights the relatively sharper regions in the image and is used to calculate the perceived sharpness score of the image. The advantages of the proposed method lie in its use of simple visual cues of high frequency content and local contrast to arrive at the perceptual score, and requiring no training with the images. The remaining parts of the chapter are arranged as follows. The proposed method is discussed in details in Section 5.1. The related experiments and results are summarized in Section 5.2. The chapter is concluded by a summary
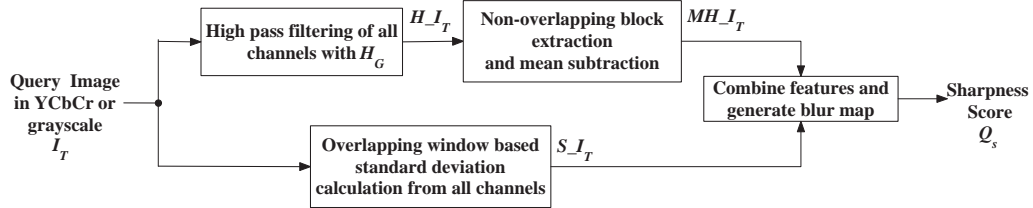
Figure 5.1: Block Diagram of the proposed method for evaluating perceptual sharpness.

presented in Section 5.3.

5.1 Proposed Method

In this section, the proposed method is discussed in details. The block diagram of the method is presented in Fig. 5.1. We describe the proposed method for a color image for generalization. In case of a grayscale image, the steps for the color image are performed for single channel only. Let I_c be the color image whose perceptual sharpness is to be determined. The pixel values of the image are scaled down between 0 and 1. At first, the image is transformed to YCbCr color space from RGB with the channels I_Y denoting luminance and I_{Cb} and I_{Cr} denoting the chrominance channels. This color space transformation is carried out as YCbCr space exhibits improved perceptual uniformity with respect to RGB [49]. Next, we have used either high pass filter (HPF) derived by subtracting Gaussian low pass filter (of width 3×3 and standard deviation 0.25) from all pass filter or the Undecimated Wavelet Transform (UWT) for extracting the high frequency content. These two filters are chosen as they form very basic and simple techniques for extracting high frequency content. The HPF based based technique will be referred as High Pass Filter based Sharpness Measure (HPFSM) and UWT based technique is named as Undecimated Wavelet Transform based Sharpness Measure (UWTSM). For the HPF, the filter is denoted

by H_G in the spatial domain and standard deviation of the low pass filter from which it is derived is indicative of its bandwidth. As calculated, the cut off frequency of 1D high pass filter derived from the Gaussian low pass filter of standard deviation 0.25 is $\frac{\pi}{4.934}$ rad/s. For HPF, the high frequency content is then obtained by

$$H\_I_Y = H_G \otimes I_Y, \tag{5.1}$$

$$H\_I_{Cb} = H_G \otimes I_{Cb}, \tag{5.2}$$

$$H\_I_{Cr} = H_G \otimes I_{Cr}, \tag{5.3}$$

where $\otimes$ denotes the two-dimensional convolution operation between two signals. When UWT is used for extracting high frequency content, the diagonal subbands from the single level UWT decomposition of all of the channels are chosen as H_I_Y, H_I_{Cb} and H_I_{Cr}. The diagonal subbands are chosen as they are result of processing by two high pass filters whereas each of other subbands are formed due to the application of at least one low pass filter [100]. The MATLAB implementation of two-dimensional UWT using 'db1' has been used for our purpose. Next, from each of these matrices locating and indicating high frequency content, the non-overlapping blockwise means of the high frequency content is subtracted and the absolute values of the differences are used to form matrices MH_I_Y, MH_I_{Cb} and MH_I_{Cr} from H_I_Y, H_I_{Cb} and H_I_{Cr} respectively. This process removes the blockwise high frequency content bias. The block size b_l used is fixed in all of our experiments. Now the local standard deviation with the same block size is calculated for all overlapping blocks from the image channels to obtain the matrices S_I_Y, S_I_{Cb} and S_I_{Cr} from I_Y, I_{Cb} and I_{Cr} respectively. The local standard deviation serves as the contrast measure in natural images [3]. The standard deviation is used to weigh the high frequency content to obtain the matrices T_Y, T_{Cb} and T_{Cr}. Each element (i,j) of the matrices is calculated

as follows:

$$T_Y(i,j) = \frac{[MH\_I_Y(i,j)]^\alpha \times S\_I_Y(i,j)}{\sum_k \sum_l (S\_I_Y(k,l))} \tag{5.4}$$

$$T_{Cb}(i,j) = \frac{[MH\_I_{Cb}(i,j)]^\alpha \times S\_I_{Cb}(i,j)}{\sum_k \sum_l (S\_I_{Cb}(k,l))} \tag{5.5}$$

$$T_{Cr}(i,j) = \frac{[MH\_I_{Cr}(i,j)]^\alpha \times S\_I_{Cr}(i,j)}{\sum_k \sum_l (S\_I_{Cr}(k,l))} \tag{5.6}$$

Here, α denotes the exponent for high frequency content obtained from each channel. We know from the aforementioned equations that, if the local contrast is high enough, the high frequency content is weighted more. Now the total stimulus due to the high frequency content weighted by local standard deviation is calculated as

$$TS(i,j) = \left(\frac{T_Y(i,j) + T_{Cb}(i,j) + T_{Cr}(i,j)}{3} \right)^{\frac{1}{\alpha}} \tag{5.7}$$

Now the raw sharpness map is generated by the following relation:

$$S_{map}(i,j) = \frac{\text{abs}\,(\log(\epsilon)) + \epsilon}{\text{abs}\,(\log(TS(i,j) + \epsilon)) + \epsilon}, \tag{5.8}$$

where ϵ is a small positive number used to avoid instability during taking logarithm and performing division. The values $T_Y(i,j)$, $T_{Cb}(i,j)$ and $T_{Cb}(i,j)$ are always less than 1 and greater than zero. Hence natural logarithm of their sum is negative. The closer the sum is to zero, the more is the absolute value of the logarithm. For sharper regions, higher values for contrast and high pass filtered content are expected. Hence, the sum is closer to 1 and $S_{map}(i,j)$ is high. For blurred images, the sum will be closer to zero and hence, the denominator will be high leading to lower values of $S_{map}(i,j)$. In order to generate the sharpness score, the border of the sharpness map is discarded depending on the block size, hence eliminating the border effect; the remaining map is called BS_{map}. The perceptual quality score is finally obtained as

$$Q_s = (\max(BS_{map})). \tag{5.9}$$

where max(.) calculates the maximum value of the argument. The max(.) operator here closely follows the approach of the HVS to combine visual responses in a non-linear fashion [25] and thus we apply maximum pooling to arrive at the sharpness

92

score. The BS_{map} lacks localization and hence we calculate the final localized map by median filtering. In order to enhance the visualization of the sharper regions along with localization, we do the following. First, we calculate γ as

$$\gamma = \frac{\max(BS_{map}) + \epsilon}{\text{mean}(BS_{map}) + \epsilon}. \tag{5.10}$$

The final localized blur map is calculated by

$$LBS_{map} = \exp(\gamma \times \text{block\_median}(BS_{map}, b_l + 2)) \tag{5.11}$$

where block_median(A, n) calculates overlapping blockwise median of the matrix A for a block size given by $n \times n$. Now we present two discussions on the intuitive explanation and the parameters of the proposed method.

Intuitive Explanation of the Proposed Method

The proposed method actually echoes the psychometric function presented in [81] which expresses the probability P_t of detection of a stimulus S_t by detector t is dependent on the contrast C_t related to the stimulus as

$$P_t = 1 - 2^{(-S_t C_t)^q}, \tag{5.12}$$

where q is the exponent. From this equation, it is revealed that the product $(S_t C_t)^q$ is of major importance in calculating the probability of detecting the stimulus. The stimulus S_t in our proposed method is the high frequency content derived after the high pass filtering of the signal. We do a basic thresholding in the stimulus by subtracting the blockwise mean as discussed earlier in this section. The C_t is obtained in the form of the local standard deviation. For each channel we calculate the product of the exponentiated high frequency content and local standard deviation to obtain T_Y, T_{Cb} and T_{Cr}. The exponent α can be compared to the exponent q in Eqn. 5.12. However, it differs from q is one aspect; α values are used to exponentiate the principal stimulus only. After that, a summation of these values takes place to generate a score.

93

The final form of the measure in Eqn. 5.8, is used to increase the range of the values obtained from proposed method. The most important part is therefore the argument of the logarithm present in the denominator for calculating raw sharpness map in this equation. As we designed the measure, the maximum value of the sharpness map BS_{map} is selected as representative perceived sharpness. Hence, it is theoretically ensured in the final form that the value of the measure increases with the perceived sharpness.

Discussion about the Parameters Involved

The block size b_l and map exponent α are the two main parameters of the implementation. The block size b_l is used in several computations and a fixed value of 7 is used for it in every experiment. Larger values of b_l increases the computation time. α is the parameters which control the contribution of stimulus to determine the final quality score. Hence, we have carried out experiments with the α values. The effect of changing these values is shown in Fig. 5.2. As we see, the performance improves remarkably for α value greater than 1. Since, the elements of the matrices MH_I_Y, MH_I_{Cb} and MH_I_{Cr} lie between 0 and 1, an exponent value less than 1 increases the values. Exponent values greater than 1 decrease their values. We find from the experiments that better performance are achieved, if the values of the stimulus remain same or decrease. In the experiments, $\alpha = 2$ is used as for α values greater than 1, higher and less fluctuating SROCC values are obtained for all databases.

5.2 Experiments and Results

In this section, we discuss several experiments conducted using the proposed method. First, we demonstrate the qualitative and quantitative results on perceived sharpness for different sharpness levels on the same image. Second, we discuss the performance of our methods on a set of different images with varying amount of blur. In the third
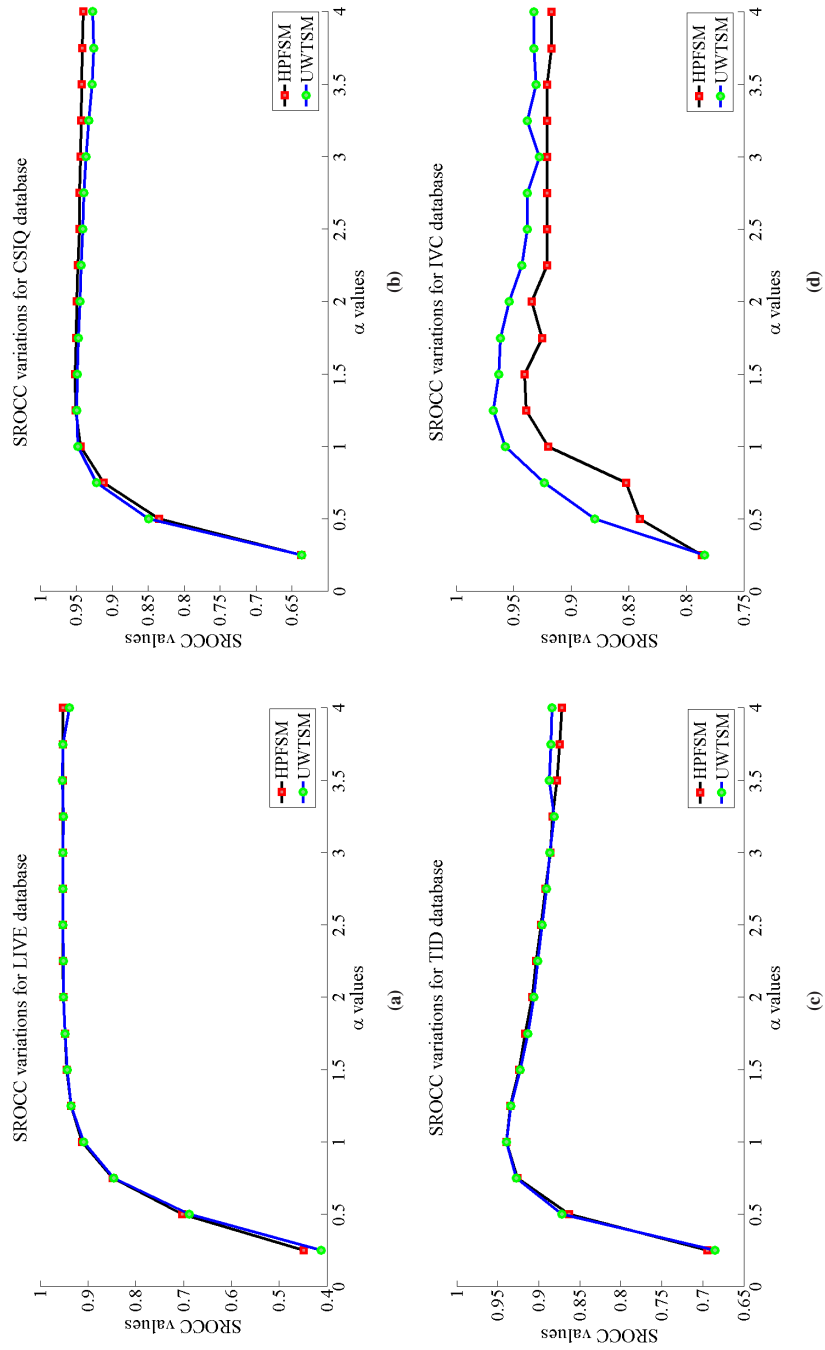
Figure 5.2: Variation of SROCC scores for different values of α using HPFSM and UWTSM.

experiment, the performance in four databases is presented. A comparative study with several existing methods is presented in the fourth experiment. Four databases are used for the experiments: LIVE database [94], CSIQ database [43], TID2008 database [79] and IVC database [4]. Pristine and corresponding blurred images are available in these databases along with the subjective scores provided by the human observers.

5.2.1 Different Blur Levels in Same Pristine Image

In the first experiment, we select a pristine image from the LIVE database and consider its several blurred versions. As per [94], varying levels of blur are applied on the pristine image 'Parrot' to generate four different images which are shown in the first row of Fig. 5.3. The figures in the first row are arranged in the increasing order of the corresponding DMOS values, signifying the degradation of perceptual quality resulting due to blur or loss of sharpness. In the first two images, the birds can clearly be identified as the part of the foreground. However, the distinction between the foreground and background gradually diminishes as the DMOS increases. The next two rows in the same figure shows the corresponding sharpness maps obtained using HPFSM and UWTSM respectively. A close look in each of these images show that the relative sharpness values at each pixel are presented in the sharpness maps. The range of the values in the maps are adjusted for better visualization. The maps produced by HPFSM and UWTSM are different as sharper regions are more discriminated by UWTSM when compared to HPFSM. In Fig. 5.3(a), the birds' outlines and eye regions are very sharp compared to the background. As we proceed over from Fig. 5.3(b) to Fig. 5.3(d), the sharpness difference between the backgrounds and the birds drops gradually signifying the loss of relative sharpness between the background and foreground. The scores from the proposed methods decrease in the same order as that of the DMOS values. This experiment shows that the proposed methods follow the rank order of the DMOS scores for different levels of blur applied on the same

image. This experiment also shows that for varying levels of blur applied on the same image, the maps produced are able to highlight relatively sharper regions.

5.2.2 Different Blur Levels for Different Pristine Images

The second experiment involves the performance evaluation of the proposed methods using different pristine images. The first part of the experiment is qualitative and second part is quantitative in nature. In the first part of the experiment, we test the performance on the suite of images provided with the implementation of S3 [101]. These images have different areas with different levels of sharpness. The images and corresponding sharpness maps are presented in Fig. 5.5. The first row of the figure presents the original images with the following rows representing the maps obtained from the proposed methods using HPF and UWT, and the ground truth respectively. However, one point needs to be clarified, while comparing the maps obtained from proposed method to the ground truth. The ground truths are generated using the evaluation of the subjects on $16{\times}16$ non-overlapping blocks of the image whereas the maps obtained from the proposed method are more localized. Therefore, a comparison can be made in the following way: the right side of the Fig. 5.5(a) 'Dragon' has relatively higher sharpness than that of the left side. The map from the proposed method supports that. As we can see, the proposed methods using HPF and UWT are able to identify the relative sharpness and blur in the images. In the Fig. 5.5(b), 'Flower', the flowers in the middle have considerably higher sharpness compared to the background. The map in Fig. 5.5(h) supports it but as it is localized, it is able to distinguish between the sharpness levels at the borders and middle portion of the flower. Fig. 5.5(n), obtained using UWT also reflects the sharper areas in the flower image. In the 'Monkey' image from Fig. 5.5(c), the left paws, fur closest to the monkey's face, monkey's eyes and fur close to the right leg are sharper areas. The same is supported by the map obtained using HPF in accordance with the ground truth. For the same image using UWT, the fur, eyes and paws are considered much

sharper than the rest of the images. As seen from the ground truth of the 'Orchid' image shown in Fig. 5.5(d), the petals in the middle have higher sharpness than that of the petals on the left. This is also captured by the maps obtained from the proposed approach. The 'Peak' image as shown in Fig. 5.5(e) has a blurred background, and the plants at the middle and left have more sharpness than the background. From the maps (shown in Figs. 5.5(k) and 5.5(q)) generated by the proposed methods, we can observe the same. Moreover, the plants at right have sharpness more than the background and less than the plants in middle. This is also supported by the ground truth. Finally, the 'Squirrel' image in Fig. 5.5(f) shows the wooden log and the squirrel has more sharpness compared to the background. The eye, borders of the ears and paws are sharper as shown by the proposed methods and the ground truth. This experiment makes clear that the maps generated using UWTSM demonstrate more contrast between the sharp and unsharp parts as compared to the maps generated using HPFSM.

In the second part of the experiment, we use images with varying levels of blur or lower sharpness generated from different pristine images. The LIVE database is chosen for this experiment. There are 29 reference or pristine quality images present in the LIVE database. Several images having varying Gaussian blur levels are generated from each of these reference images. The decrease of the value of the proposed methods with the increase of Gaussian blur is depicted in Fig. 5.4. Here, the values of the proposed methods are plotted against the standard deviation σ of the Gaussian blur applied to the pristine image. As seen from the figure, from each of the 29 reference images, the blurred images are derived by gradually increasing σ values. Thus, the subjective scores goes down gradually, indicating loss of sharpness and perceptual quality. The objective scores obtained from both HPFSM and UWTSM decreases gradually with the increasing σ values. Therefore, the objective scores are able to represent the loss of perceptual sharpness in each set of degraded images derived from the same pristine image. Though it is evident that perceived sharpness
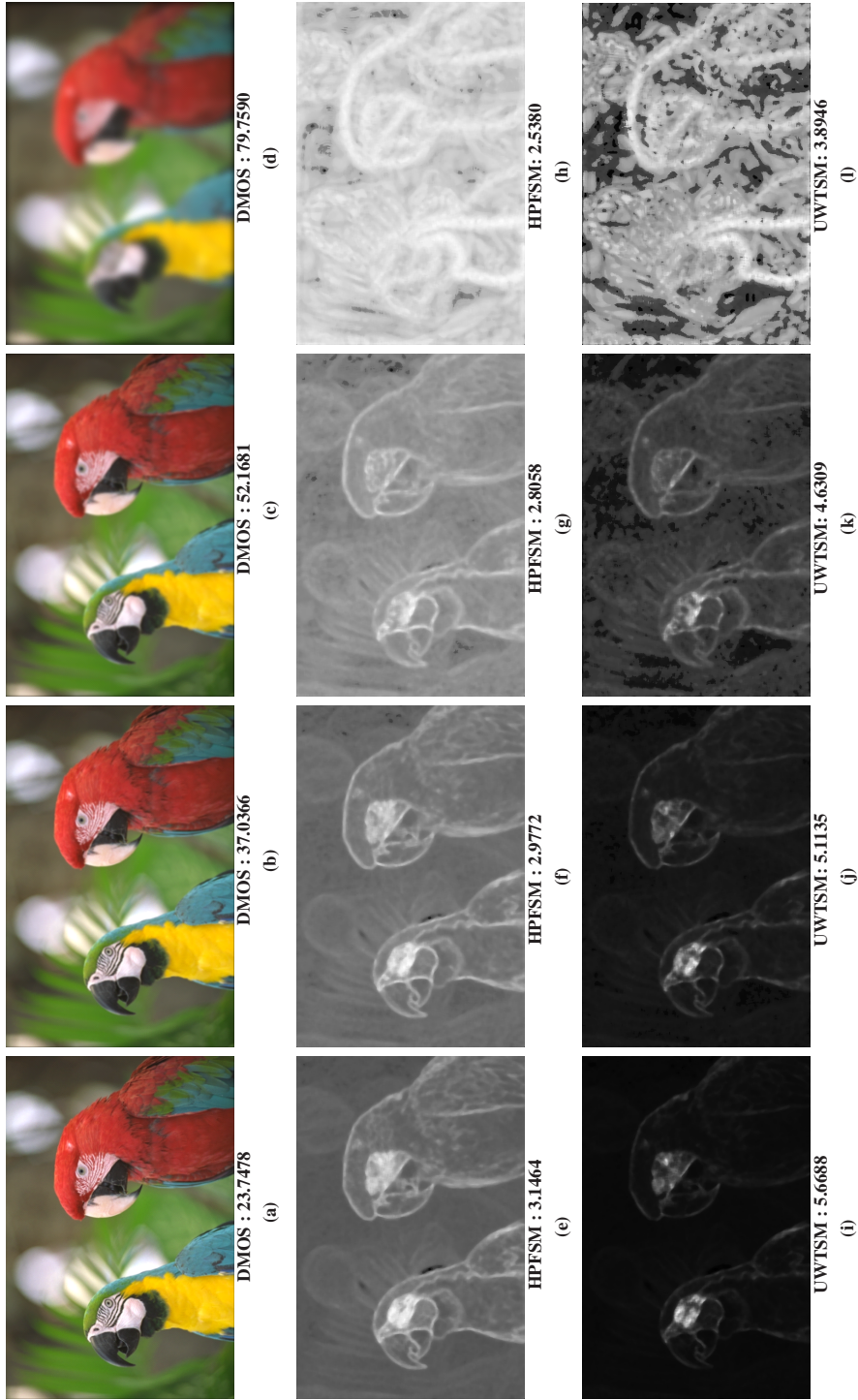
Figure 5.3: Sharpness maps and quality score for different levels of blur over the same pristine image from LIVE database [94] for HPFSM and UWTSM. Each column in the figure represents the query image, and the blur maps resulting from HPFSM and UWTSM respectively considering the direction from top to bottom.

depends a lot on the image content [28, 108], for the high levels of blur, the scores are low signifying greater degradation. We also find that for HPFSM and UWTSM, the ranges of sharpness score vary significantly.

5.2.3 Performance in Four Databases

In this experiment, the subjective scores in four databases are plotted against the objective scores obtained from the methods for different images as shown in Fig. 5.6. The scatter plot shows that a high degree of correlation exists between subjective values and the scores obtained by the proposed method. This experiment takes into account the performance across several images of varying content against the subjective scores. The scores mostly follow a monotonic variation with the subjective scores of all the blurred images taken together in a database for both HPFSM and UWTSM. A global quantitative evaluation is required to compare this performance with the existing methods and this is provided in the next section in comparison with the other methods.

5.2.4 Comparisons with Other Methods

We compare the proposed method in four databases against several FR and NR quality measures. Peak signal-to-noise ratio (PSNR) and SSIM [113] are the two FR-IQA techniques we compare with. We know that these FR-IQA techniques are moderately performing techniques in state-of-the-art FR-IQA, and hence we get a clue about how far behind is the NR techniques behind them. Considering the NR methods, we have divided them as NR-Specific and NR-General. In this context, NR-Specific means an NR quality measure that works only on blurriness or sharpness related distortion. The NR-specific methods used in comparison are : H-metric [132], Q-metric [133], JNBM [18], CPBD [75], S3 [108] and LPC-SI [28]. NR-General means the IQA technique is able to the assess the image quality for different distortions. BRISQUE [68]
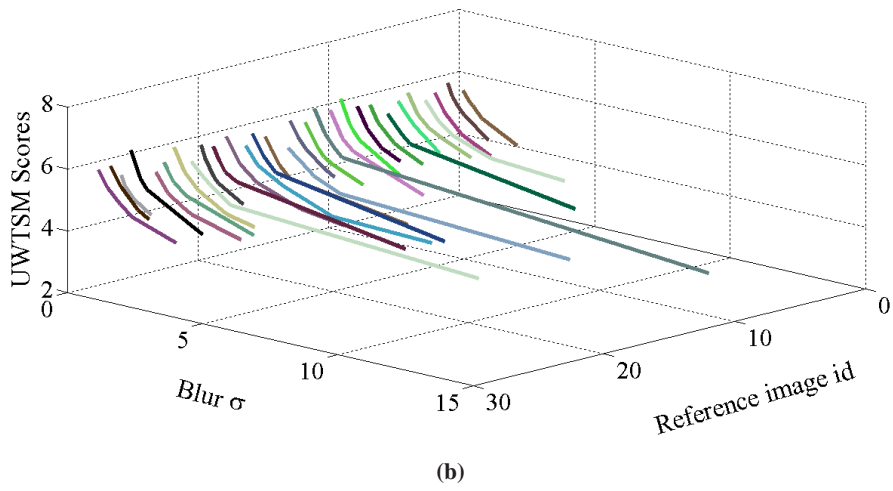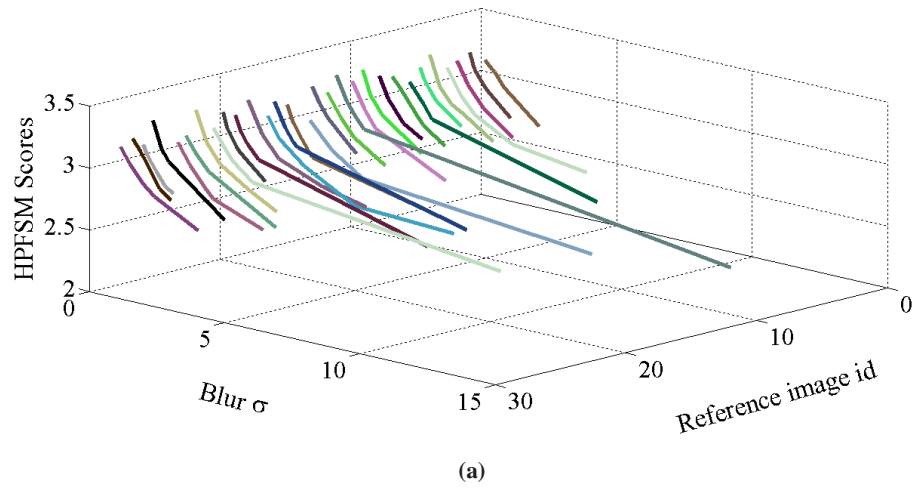
(a)

(b)

Figure 5.4: The variations of the objective scores with the Gaussian blur σ in LIVE database using (a) HPFSM and (b) UWTSM.
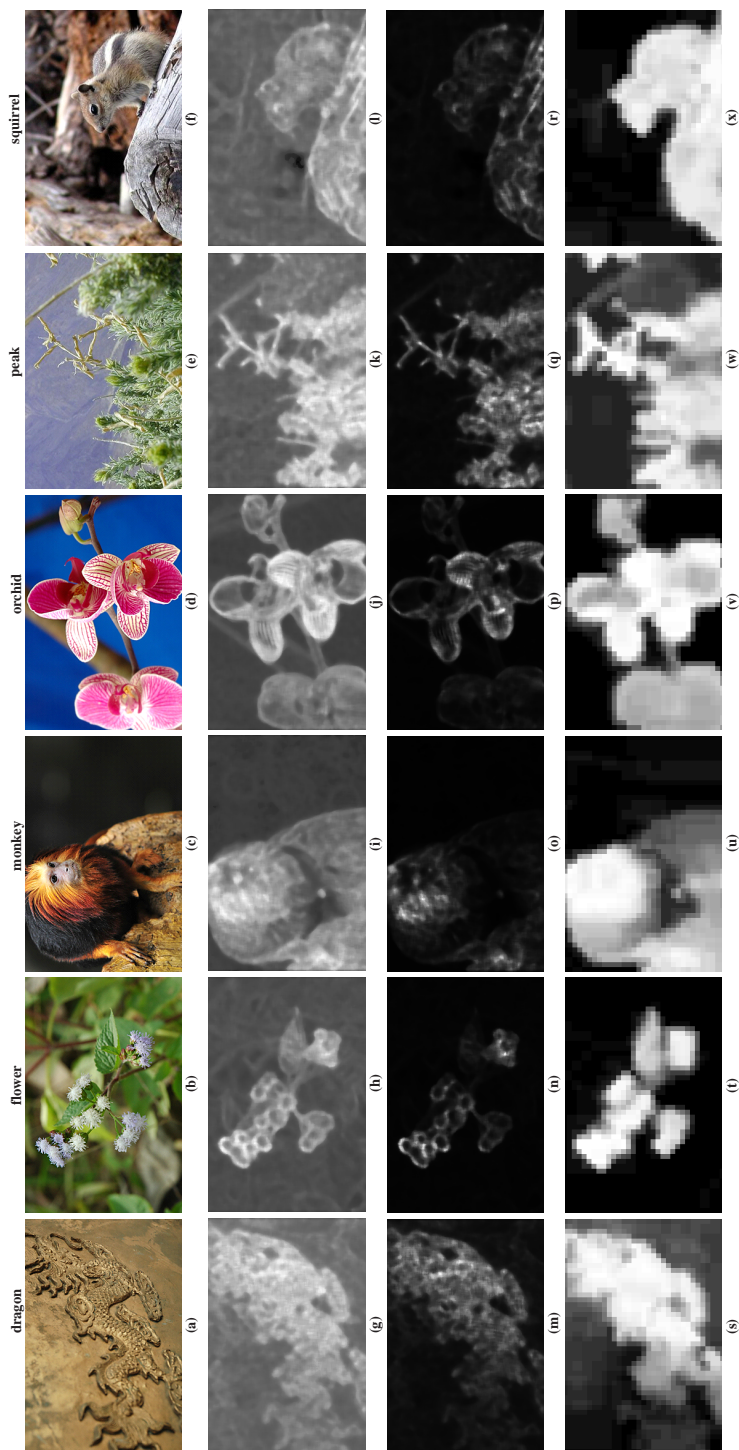
Figure 5.5: Sharpness maps for different images in [101]. Each of the six columns contains the query image. the sharpness map obtained from HPFSM and UWTSM respectively and finally the ground truth from top to bottom.
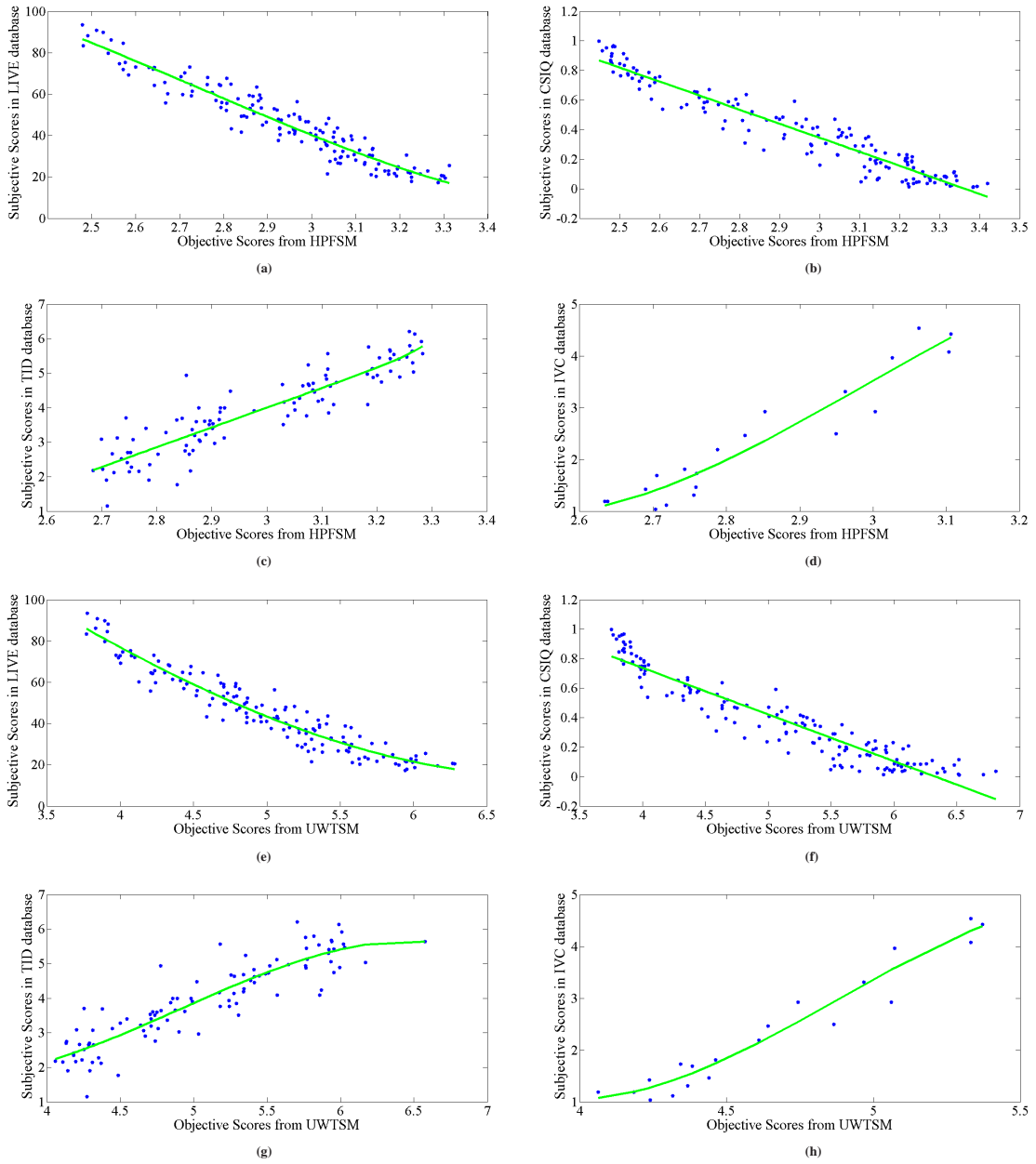
Figure 5.6: Scatter plots between the objective and subjective scores in blue. The green line represents the fitted values after applying the logistic function.

and BLIINDS [84] are the NR-General methods used for comparison here. The qualitative results obtained from the comparison process are presented in Table 5.1. The best performer's score in each database and for each evaluation measure is highlighted in bold and the best method among the NR-specific methods is underlined. For LIVE database, we find that SSIM has highest SROCC score, UWT has highest PLCC and lowest RMSE values and the proposed method provides highest KROCC value. In CSIQ databases, the proposed methods using HPFSM and UWTSM perform remarkably better than all the methods for all types of evaluation measures. For TID2008 database, SSIM is the best performer for all evaluation measures but considering NR-specific methods only, the proposed method improves the performance over the existing ones to a large extent. As already revealed in the Fig. 4.8 presented in Chapter 4, the best techniques in LIVE, CSIQ and TID2008 databases are VIF, GLD-SR and IGM respectively. For IVC database, UWTSM, LPC-SI and H-metric achieve the best performance using different evaluation measures.

We also note the average performances of the methods in all four databases are presented in Tables 5.2a and 5.2b. A direct computation of the average values for SROCC, PLCC and KROCC reveals the proposed UWTSM has the best performance among all the methods. However, when the values are weighted by the number of relevant images present in the corresponding dataset, SSIM shows the best performance with SROCC. UWTSM performs the best with PLCC while using KROCC, HPFSM is the best performer. The NR-general techniques work on a lot of distortions and hence achieving the best performance on a specific kind of distortion is difficult for them.

5.3   Chapter Summary

We presented a technique for automatic assessment of perceived sharpness in natural images. The technique emphasizes on the application of high frequency content with

Table 5.1: Comparison of HPFSM and UWTSM with the existing techniques in four databases.

Databases	Measure	FR		NR-Specific								NR-General	
		PSNR [113]	SSIM [113]	H-metric [132]	Q-metric [133]	JNBM [18]	CPBD [75]	S3 [108]	LPC-SI [28]	HPFSM	UWTSM	BRISQUE [68]	BLIINDS II [84]
LIVE (145 gblur images)	SROCC	0.7823	**0.9516**	0.6865	0.5482	0.7872	0.9182	0.9436	0.9389	0.9515	0.9510	Used for training	Used for training
	PLCC	0.7748	0.9483	0.6977	0.6559	0.7835	0.9119	0.9525	0.9315	0.9537	**0.9559**		
	KROCC	0.5847	0.8011	0.5080	0.3971	0.6071	0.7632	0.7996	0.7785	**0.8094**	0.8086		
	RMSE	11.6815	5.8634	13.2389	13.9424	11.5085	7.5852	5.6269	6.7352	5.5578	**5.4270**		
CSIQ (150 gblur images)	SROCC	0.9291	0.9245	0.7833	0.6528	0.7624	0.9182	0.9058	0.9068	**0.9495**	0.9444	0.9032	0.8765
	PLCC	0.9081	0.8903	0.7467	0.6562	0.7887	0.9119	0.8921	0.9255	0.9594	0.9449	0.9252	0.8930
	KROCC	0.7543	0.7657	0.6081	0.4860	0.5971	0.7632	0.7290	0.7197	**0.8053**	0.7972	0.7353	0.6783
	RMSE	0.1201	0.1328	0.1910	0.2165	0.1780	7.5852	0.1296	0.1086	**0.0809**	0.0939	0.1090	0.1290
TID2008 (100 gblur images)	SROCC	0.8682	**0.9596**	0.5275	0.3276	0.6667	0.8412	0.8418	0.8561	0.9079	0.9061	0.7989	0.8205
	PLCC	0.8666	**0.9555**	0.5553	0.2930	0.6535	0.8320	0.8540	0.8562	0.9043	0.9014	0.7388	0.8260
	KROCC	0.7328	**0.8288**	0.3398	0.2210	0.4951	0.6310	0.6124	0.6362	0.7365	0.7300	0.6229	0.6245
	RMSE	0.5953	**0.3465**	0.9766	1.1220	0.8921	0.6519	0.6110	0.6066	0.5011	0.5081	0.8239	0.6614
IVC (20 gblur images)	SROCC	0.7893	0.8691	**0.9541**	0.9323	0.6659	0.7690	0.8691	0.9398	0.9345	**0.9541**	0.8239	0.5262
	PLCC	0.8527	0.9149	0.9627	0.9360	0.7037	0.8469	0.9279	**0.9736**	0.9555	0.9728	0.8489	0.7806
	KROCC	0.6349	0.7090	0.8254	0.7831	0.4974	0.6138	0.7090	0.8042	0.8148	**0.8466**	0.6561	0.3979
	RMSE	0.6309	0.4695	0.3093	0.4097	0.8138	0.6201	0.4265	**0.2607**	0.3367	0.2646	0.6151	0.7136

proper exponent as the stimulus and combining it with the contrast. The logarithm of this combination is used to generate the sharpness map and sharpness score. The performance of the proposed method is demonstrated to provide improved performance over the state-of-the-art methods. The variation in the performance of the proposed technique is also analyzed for varying values of the exponent and for two types of high frequency content extraction methods using high pass filter and undecimated Wavelet transform. We have divided the experiments into several distinct parts in order to properly visualize the consistency of performance. A maximum pooling strategy has been adopted in our approach to arrive at the perceptual score. The future work lies in studying the effect of different pooling strategies for perceived sharpness evaluation.

Table 5.2a: Direct average comparison of HPFSM and UWTSM with existing techniques over four databases.

Measure	FR		NR-Specific								NR-General	
	PSNR	SSIM [113]	H-metric [132]	Q-metric [133]	JNBM [18]	CPBD [75]	S3 [108]	LPC-SI [28]	HPFSM	UWTSM	BRISQUE [68]	BLIINDS II [84]
SROCC	0.8422	0.9262	0.7379	0.6152	0.7206	0.8617	0.8901	0.9104	0.9359	**0.9389**	0.8420	0.7411
PLCC	0.8506	0.9273	0.7406	0.6353	0.7324	0.8757	0.9066	0.9217	0.9309	**0.9438**	0.8376	0.8332
KROCC	0.6767	0.7762	0.5703	0.4718	0.5492	0.6928	0.7125	0.7347	0.7915	**0.7956**	0.6714	0.5669

Table 5.2b: Weighted average comparison of HPFSM and UWTSM with existing techniques over four databases.

Measure	FR		NR-Specific								NR-General	
	PSNR	SSIM [113]	H-metric [132]	Q-metric [133]	JNBM [18]	CPBD [75]	S3 [108]	LPC-SI [28]	HPFSM	UWTSM	BRISQUE [68]	BLIINDS II [84]
SROCC	0.8564	**0.9398**	0.6961	0.5514	0.7434	0.8925	0.9018	0.9074	0.9395	0.9379	0.8587	0.8298
PLCC	0.8489	0.9275	0.6939	0.5821	0.7502	0.8895	0.9057	0.9132	0.9262	**0.9396**	0.8505	0.8599
KROCC	0.6841	0.7905	0.5189	0.4054	0.5712	0.7241	0.7246	0.7242	**0.7906**	0.7874	0.6878	0.6376

Chapter 6

Utilizing Image Scales for No-Training NR-IQA

A new approach to blind image quality assessment (BIQA), requiring no training, is proposed in this chapter. The approach is named as Blind Image Quality Evaluator based on Scales (BIQES) and works by evaluating the global difference of the query image analyzed at different scales with the query image at original resolution. The approach is based on the ability of the natural images to exhibit redundant information over various scales. A distorted image is considered as a deviation from the natural image and bereft of the redundancy present in the original image. The similarity of the original resolution image with its down-scaled version will decrease more when the image is distorted more. Therefore, the dissimilarities of an image with its low-resolution versions are cumulated in the proposed method. We dissolve the query image into its scale-space and measure the global dissimilarity with the co-occurrence histograms of the original and its scaled images. These scaled images are the low pass versions of the original image. The dissimilarity, called low pass error, is calculated by comparing the low pass versions across scales with the original image. The high pass versions of the image in different scales are obtained by Wavelet decomposition and their dissimilarity from the original image is also calculated. This dissimilarity, called high pass error, is computed with the variance and gradient histograms and weighted by the contrast sensitivity function to make it perceptually effective. These two kinds of dissimilarities are combined together to derive the quality score of the query image. This method requires absolutely no training with the distorted image, pristine images or subjective human scores to predict the perceptual quality but uses

the intrinsic global change of the query image across scales. The remaining parts of the chapter is arranged as follows. The motivation is discussed in Section 6.1. The details about different methods used are presented in Section 6.2. The proposed approach is discussed in Section 6.3. The experiments and related analyses are presented in Section 6.4. Finally, the chapter is concluded through chapter summary presented in Section 6.5.

6.1 Motivation

Recent approaches [46, 70] work without training on human scores as well on distorted images. The method in [70] generates a parameterized multivariate Gaussian model (MVG) based on NSS from high quality images by drawing quality aware features from it. For any distorted image, the distance between the NSS feature model and the MVG model fitted to the distorted image becomes the indicator of quality. This method delivers competitive performance with the state-of-the-art techniques in the LIVE database. Another approach that uses no modeling is [46]. They use simple relationship with the features of mean phase congruency, entropy and mean gradient magnitude to accomplish their tasks. However, prior to this, they use the first step of BIQI [71] to classify the distortions. Hence, the method implicitly involves training on distorted images. Though the performance of these techniques attempting to minimize training is not the best, they provide a methodology to continue research in the direction of no-training IQA. Any kind of training makes the IQA method dependent on the training ensemble and requires effort and time to generate the trained model. Motivated by these techniques, we attempt to use the multiscale information from the image to establish an NR-IQA technique, absolutely free from any training with human scores, pristine images or distorted images.

Table 6.1: List of symbols and their significance.

Symbol	Significance
IC_s	color image at scale s
IG_s	grayscale image at scale s
$\mathcal{I}$	Fourier transform of the image I
$AS_{I,z} : z \in [1, N]$	approximate subbands from N level Wavelet dec. of I
$DS_{I,z} : z \in [1, N]$	diagonal subbands from N level Wavelet dec. of I
$HS_{I,z} : z \in [1, N]$	horizontal subbands from N level Wavelet dec. of I
$VS_{I,z} : z \in [1, N]$	vertical subbands from N level Wavelet dec. of I
CH_I	color co-occurrence histogram for image I
GX_I	blockwise gradient along the horizontal direction of I
GY_I	blockwise gradient along vertical direction of I
SD_I	blockwise standard deviation of I
H_I	one-dimensional intensity histogram of I

6.2 Features and their Significance

As BIQES is not based on any learning, it uses several properties and effects of the different kinds of degradations on the image. In this section, we enlist all the different features that have been used to formulate the proposed quality measure. All of these features are affected by distortions and hence they have significant contribution in shaping up the final form of the proposed method. The first and foremost techniques are the ways to decompose the image into various scales. We have used simple scale-space technique for evaluating changes in low pass components and Wavelets to evaluate the high pass components. In course of explaining the features, we use symbols which are all listed in Table 6.1.

6.2.1 Scale-Space

One of the motivations behind the development of scale-space theory [51] is to imitate the multi-scale processing of images by human eyes. Natural images contain objects which are relevant at particular scales. An image might have objects and structures

that are relevant at different scales, but human eye is readily able to identify and process the information presented by it. Thus, processing an image at various scales adds flexibility to the processing technique. In decomposing image to various scales, it is necessary to disallow the formation of any spurious detail which is not due to the image of original size (called outer scale). At the same time, revelation of a structure at proper scale (called inner scale) is necessary. To establish this, convolution with Gaussian kernel and subsequent downsampling are required to form the proper scale-space of an image [37]. As one moves from the fine to coarser scales, the smoothing effect increases in the scales and the difference between the original and downsampled images increases. In spite of the smoothing, significant changes sustain throughout the coarser scales [37] as they are redundant in nature. However, for distorted images, the redundancy is lacking. As one proceeds over scales, this difference is likely to increase more for distorted images compared to the natural images. Using scale-space decomposition for any scale (among the inner scales), a smooth version of the original image is obtained. This version is dependent on the original image only. As the resulting images at different scales represent the low pass or smooth versions of original image, the comparison between these resulting images would mainly represent the comparison of low pass components of original image. Global features from all of the low pass versions at several scales are extracted, and compared with that of the image of original scale. Co-occurrence histograms (discussed in Section 6.2.4) extracted from the images at various scales, are used as the global features. To compare the high pass versions, we resort to the Wavelet decomposition.

6.2.2 Wavelet Decomposition

Application of Wavelet transform on an image decomposes the image into various subbands at different scales/levels. As pointed out in [82], occlusion (resulting due to the projection of 3D data to 2D space) is the rule of image formation and that restricts the Wavelet coefficients to be dependent with each other. Across various

scales and orientations, the joint and pairwise histograms of Wavelet coefficients for natural images are found to be non-Gaussian in nature [97]. But the maximum interdependency of Wavelet coefficients occurs across scales as demonstrated in [53]. The nature of dependency of Wavelet coefficients for distorted images is different than that of natural images [23]. Hence, we compare the features obtained from the Wavelet coefficients across scales using the horizontal, vertical and diagonal subbands with the features obtained from the original image to assess the quality. The distribution of variance and local gradients across scales of the natural images are found to be showing scale invariance [83]. Hence, we use these information across scales to obtain the perceptual quality of the image.

The high pass versions could have been calculated directly from the scale-space by upsampling the low pass version and finding its difference with the low pass version in just the scale above. However, upsampling process may introduce artifacts in the image. Therefore, Wavelet decomposition has been chosen and it is more effective due to available Wavelet statistics of the natural images. On the other hand, approximate subbands from the Wavelet decomposition are not used to compare the low pass versions of the image as they are not only dependent on the original image but also on the prior level of decomposition.

6.2.3   Fourier Transform

Fourier transform represents the global information of spatial image data in frequency domain. It is not possible to obtain local information from the transform. However, a global impression of the data is sometimes important as we show that log-spectra of natural images roughly abide by the power law. In Fig. 6.1, we can see remarkable change in the Fourier spectra due to different types of distortions of the same image. The local kurtosis of DCT power spectrum has been used earlier for the sharpness detection of images [6]. The diagram in Fig. 6.1 reveals that the kurtosis (κ) of log amplitude of the Fourier spectrum is also important. It has higher values for blur and
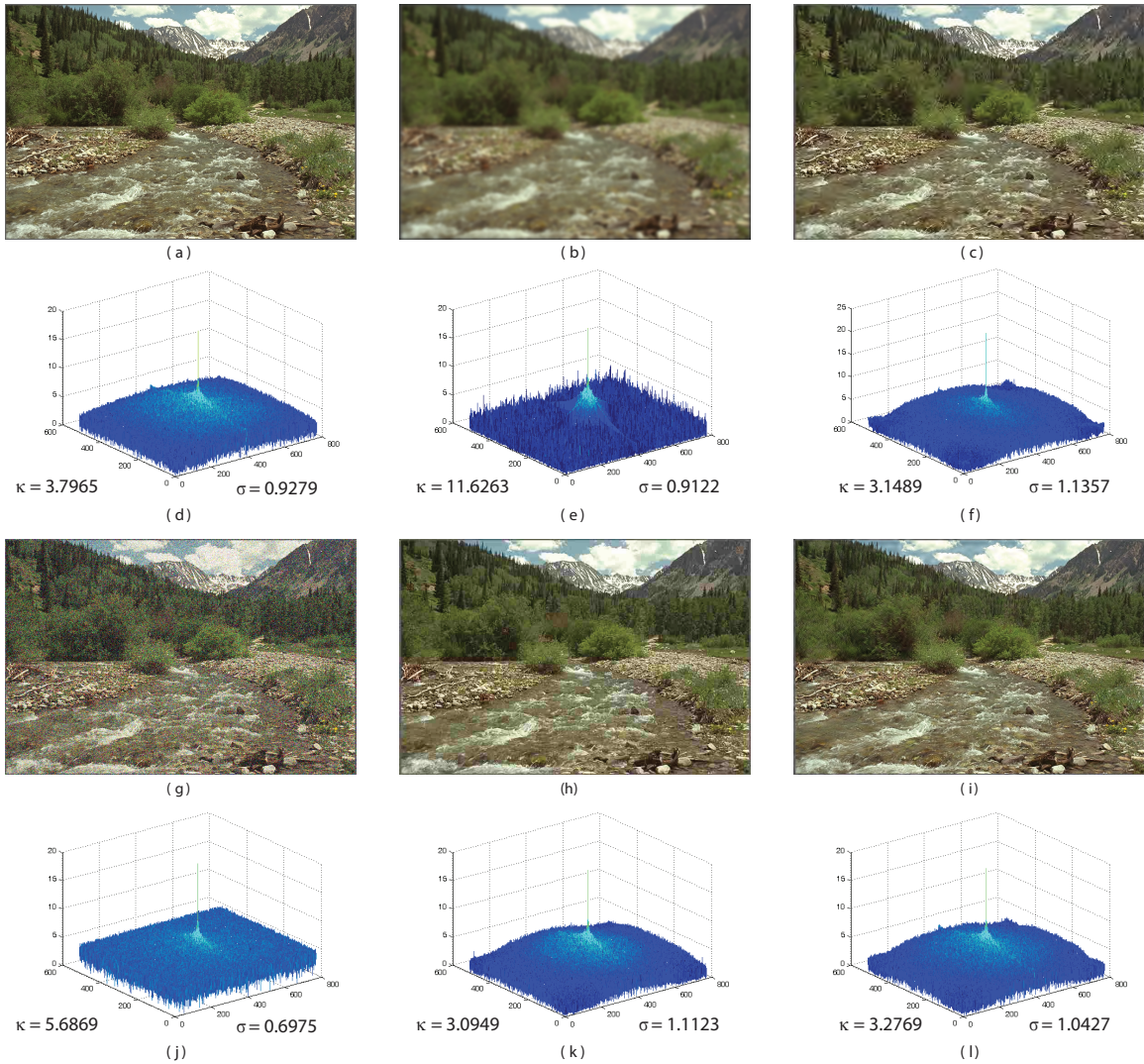
Figure 6.1: (a) Image 'Stream' and its distorted versions due to (b) blur, (c) jpeg2k, (g) wn, (h) jpeg, (i) Rayleigh Fading or fastfading [94]; (d)-(f) and (j)-(l) represent the log amplitude of the Fourier transform of the images (after conversion to grayscale) respectively. The kurtosis(κ) and standard deviation(σ) of each of spectrum are also provided.

noisy images. As shown in [88], these two distortions are greatly affected by the scale-space decomposition. Again, the standard deviation for the blurred and noisy images decreases, but for others the standard deviation increases. Thus, the ratio of kurtosis to standard deviation (σ) will be higher for blur and noisy images and this serves as important cue in the formulation of the quality measure. Higher values of kurtosis of the Fourier spectra signify that the high frequencies have comparable contribution to the spectra. In other words, the amplitude does not diminish highly with the increase of frequency. Lower standard deviation on the other hand shows that the peak is strong. As we proceed to higher scale values in scale-space, the high frequency information is lost mostly. The ratio (κ/σ), therefore, is supportive of the information loss suffered as we proceed over scales in the scale-space decomposition. On the other hand, detailed subbands in Wavelets decomposition, lose the low frequency information over increasing levels. The loss will be more, if the contribution of low frequency components are more in the spectrum. Hence, the standard deviation of log amplitude spectrum is supportive of the loss in detailed subbands over increasing levels in Wavelet decomposition.

6.2.4 Co-occurrence Histograms

Co-occurrence histograms can be regarded as the global representation of local information. Unlike image histograms which represent the frequency of occurrence of an intensity level in an image, co-occurrence histograms model the neighborhood of a specific intensity. They are found to generate effective features for finding image saliency [57]. Given a grayscale image I, the normalized co-occurrence histogram CH_I is calculated as

$$CH\_I(p, q) = \frac{n_I(p, q)}{\sum_{p_1} \sum_{q_1} n_I(p_1, q_1)}, \tag{6.1}$$

where $n_I(p, q)$ counts the number of times the intensity value q falls in the neighborhood of p in the image I. The spatial relationship between neighboring pixels is

captured by co-occurrence histogram which is two-dimensional. For a color image, co-occurrence histogram is extracted from each channel. The co-occurrence histograms of all three color channels taken together, is called color co-occurrence histogram. As pointed out in [57], it is relatively tolerant to scales for natural images. However, for the distorted images, the co-occurrence histogram varies significantly over the scales as revealed by the experiments. Thus, the comparisons of the low pass versions of the original image with the original image are carried out by the co-occurrence histograms across different scales. We show in Fig. 6.2, that for distorted images, the mean difference of the co-occurrence histograms increase for distorted images compared to the reference image.

6.3   Formulation of BIQES

The outline of the approach used in BIQES is presented in Fig. 6.3. We will present the method with color image (having 3 channels). Then, the method can easily be adapted to grayscale (single channel). This section is divided into three parts: calculating quality differences in low pass versions, calculating high pass difference from Wavelets and combining these two differences to generate the final form of the distortion measure.

6.3.1   Computation of Distortion from Low Pass Versions

The image at original scale is processed in two ways. At first, the co-occurrence histograms are extracted from the original image. For a color image IC, normalized color co-occurrence histogram CH_IC consisting of co-occurrence histograms from three channels as CH_IC1 ,CH_IC2 and CH_IC3 is formed. These co-occurrence histograms are also referred as CH_ICT where $T \in \{1, 2, 3\}$ represents each channel. In the next step, the image is subjected to scale-space decomposition by Gaussian filtering and subsequent downsampling for N levels. Thus, we have a set of N color
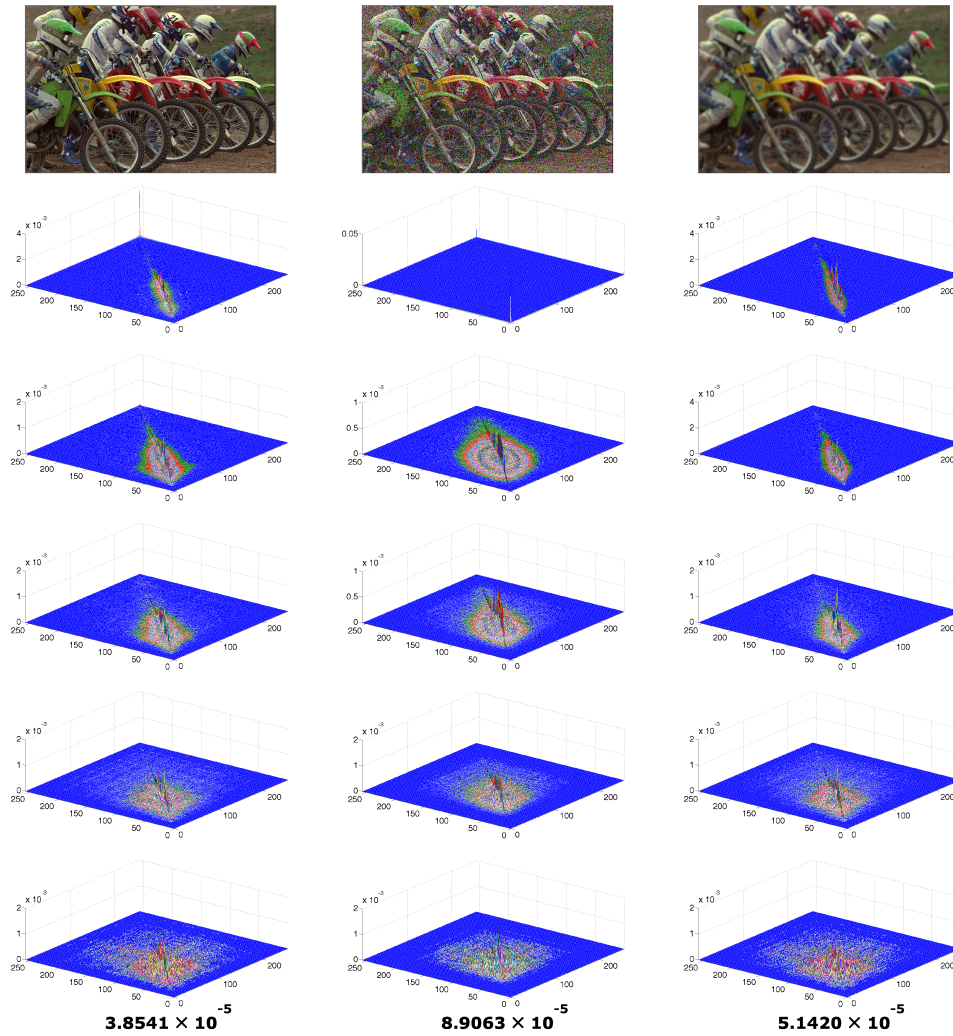
Figure 6.2: First row (left to right) Image 'Bikes' and its distorted versions due to white noise and blur respectively [94]; Second row presents the respective co-occurrence histogram from the first channel (red) of the color image shown in the top row. Third to last rows present the respective co-occurrence histogram from the first channel of the color image shown in the top row at subsequent four scales. The numerical value below each column is the mean of the total differences of the histograms at each scale and the histogram in the second row. The numerical values for the distorted images are greater than that of the reference image.
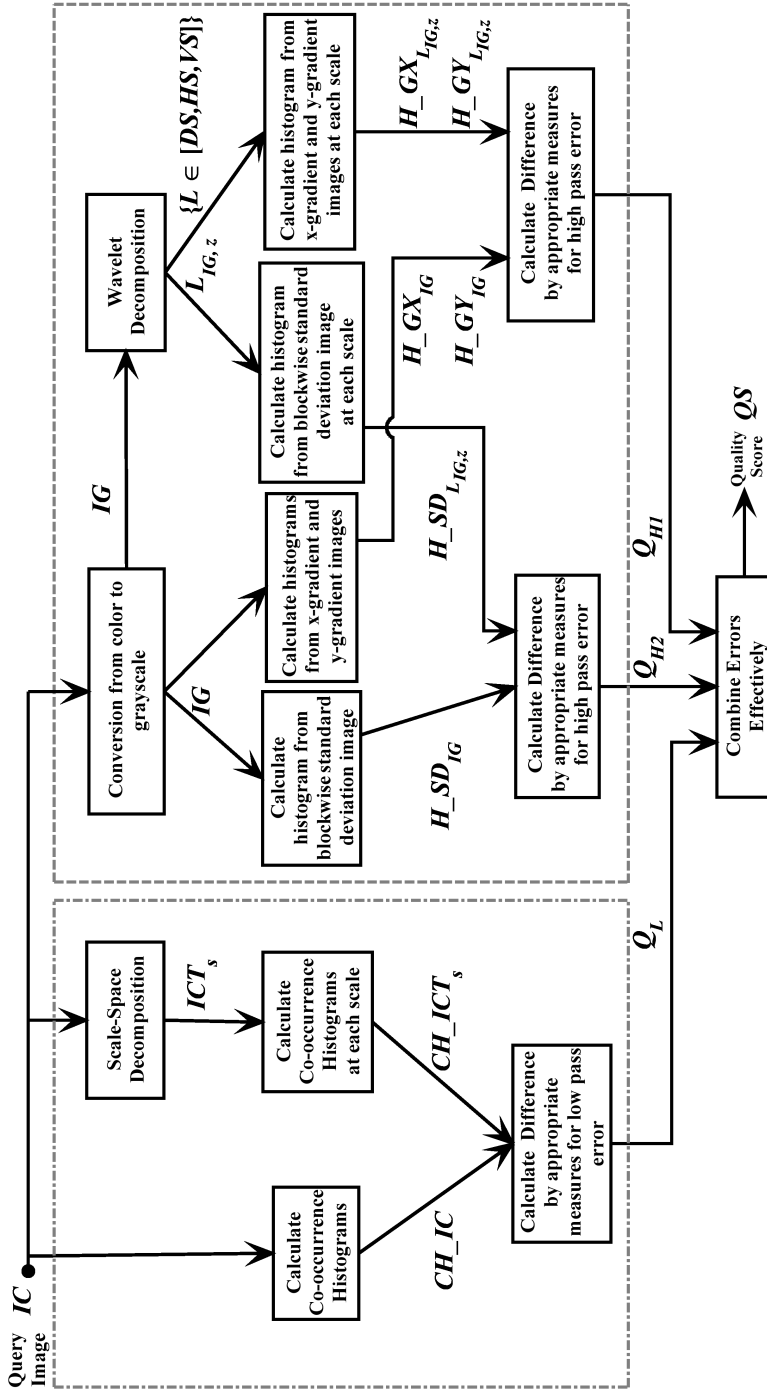
images, represented by IC_s where $s \in \{1, 2, .., N\}$. As already discussed in Section 6.2, both of the scale-space and color co-occurrence histograms should exhibit relatively higher scale invariance for natural images. For a distorted image, the difference of the histograms at various scales (from the histogram obtained at outer scale) will be more than that of a natural image. From each of these images, co-occurrence histograms CH_ICT_s where $T \in \{1, 2, 3\}$ are generated. Thus, the low pass difference or error at each scale s is calculated as

$$t_s(p, q) = \frac{1}{3} \sum_T |CH\_ICT(p, q) - CH\_ICT_s(p, q)|. \qquad (6.2)$$

Here, $|.|$ denotes absolute value of the argument. An observation regarding the difference of the histograms needs to be mentioned here. As the low pass versions are compared, considerable amount of information of the original image is retained, rather than being filtered out. So, the values of $t_s(p, q)$ are small. Before combining them over the scales, they are weighted by an appropriate factor γ_s at scale s. The explanation and way to calculate γ_s are explained in Section 6.3.3. The combined error from low pass versions at all scales is obtained as

$$Q_L(p, q) = \sum_s \gamma_s t_s(p, q). \qquad (6.3)$$

6.3.2 Computation of Distortion from High Pass Versions

After finding the difference in low pass versions, the next step is to find the difference in high pass versions. The color image is converted to grayscale (not required for grayscale images) to form IG. Next, the x-gradient and y-gradient images GX_{IG} and GY_{IG} are calculated using the simple pixel wise horizontal and vertical differences from the IG. Afterwards, the blockwise standard deviations are calculated to obtain SD_{IG}. From these three images, the histograms H_GX_{IG}, H_GY_{IG} and H_SD_{IG} are calculated as global features. Next, an N-level Wavelet decomposition is applied on IG. The CDF 9/7 Wavelet has been used for this purpose owing to its advantages over

the others [119]. At each level $z \in \{1, 2, .., N\}$, the detailed subbands obtained are represented as $DS_{IG,z}$, $HS_{IG,z}$ and $VS_{IG,z}$ and the approximate subband as $AS_{IG,z}$. From each of these subbands the x-gradient images, y-gradient images and standard deviation images are calculated. For example, in case of the the diagonal subbands $DS_{IG,z}$, these images are represented as $GX_{DS_{IG,z}}$, $GY_{DS_{IG,z}}$ and $SD_{DS_{IG,z}}$. From all of these detailed subbands' feature images, the 1D histograms are calculated as global features. For example, in case of the diagonal subbands $DS_{IG,z}$, the feature histograms are represented as $H\_GX_{DS_{IG,z}}$, $H\_GY_{DS_{IG,z}}$ and $H\_SD_{DS_{IG,z}}$. Thus, a part of the high pass difference or error is calculated from the gradient histogram images as

$$Q_{H1}(b) = 0.5\eta_z \sum_z (D\_GX_z(b) + D\_GY_z(b)) \tag{6.4}$$

where

$$D\_GX_z(b) = \max(\alpha1_z|H\_GX_{IG}(b) - H\_GX_{DS_{IG,z}}(b)|$$
$$, \alpha2_z|H\_GX_{IG}(b) - H\_GX_{HS_{IG,z}}(b)| \tag{6.5}$$
$$, \alpha2_z|H\_GX_{IG}(b) - H\_GX_{VS_{IG,z}}(b)|)$$

Here, max(.) finds the maximum of the elements and b denotes each bin of the one-dimensional histogram. η_z is vector denoting the relative weights applied to the subbands based on the experiments of weights for different scales [116] such that $\sum_z \eta_z = 1$. Also, $\alpha1_z$ and $\alpha2_z$ are derived from the contrast sensitivity function (CSF). We have used the CSF as presented in [131]:

$$C(f) = 2.6(0.0192 + 0.114\beta f)\exp(-(0.114\beta f)^{1.1}) \tag{6.6}$$

where f is the frequency expressed in cycles/degree and $\beta = 2$. The peak value of $C(f)$ occurs at $f = 4$, but the peak value is not 1. We denote the peak value as C_{max}. For a viewing resolution of n_p pixels/degree, the spatial frequency f_z of Wavelet decomposition level z will be

$$f_z = n_p 2^{-z} \text{cycles/degree}. \tag{6.7}$$

119

The viewing resolution [118] is calculated as

$$n_p = d_v d_r \tan(\pi/180) \tag{6.8}$$

where d_v is the viewing distance in centimeter and d_r is the display resolution is pixels/centimeter. The product $d_v d_r$ should at least be equal to the height (number of rows) h of the original image and in the databases used for experiments often the viewing distance is provided in terms of the screen or image height. Hence, the equation can be re-written as

$$n_p = n_r h \tan(\pi/180) \tag{6.9}$$

where n_r is the ratio of the viewing distance and image height. In this way, the viewing resolution is calculated for horizontal and vertical subbands. As shown in [43], we take into account the oblique effect for the diagonal subbands. Hence, the viewing resolution of the diagonal subband is $n_p/0.7$. Hence, $\alpha 1_z$ and $\alpha 2_z$ are evaluated as

$$\alpha 1_z = \frac{C(f_z/0.7)}{C_{max}} \text{ and } \alpha 2_z = \frac{C(f_z)}{C_{max}}. \tag{6.10}$$

In this way, the errors from the different subbands are weighed using contrast sensitivity based factors. The high pass error calculated using the standard deviation histogram images are presented as

$$Q_{H2}(b) = \eta_z \sum_z D\_S_z(b) \tag{6.11}$$

where

$$
\begin{aligned}
D\_S_z(b) = \max(&\alpha 1_z |H\_SD_{IG}(b) - H\_SD_{DS_{IG,z}}(b)| \\
,&\alpha 2_z |H\_SD_{IG}(b) - H\_SD_{HS_{IG,z}}(b)| \\
,&\alpha 2_z |H\_SD_{IG}(b) - H\_SD_{VS_{IG,z}}(b)|)
\end{aligned}
\tag{6.12}
$$

6.3.3 Combination of the Errors

In this section we perform three tasks. First, we explain the γ_s values required to calculate the error Q_L. Next, we combine the errors to generate the final form of the

proposed method. Thirdly, we discuss about the various parameters involved in the calculation.

Calculation of γ_s

As presented in the earlier section, we have used N levels of scale as well as N levels of Wavelet decomposition, though the variables s and z are used to denote each scale and level respectively. To minimize the computations, we work with the same value of scale and level at a time. This reveals two important factors. Firstly, while following the scale-space, we use Gaussian filtering and downsampling. Therefore, we discard all pixels at the borders. The Wavelet decomposition used, is the 'bior4.4'(same as CDF 9/7) presented in MATLAB. According to the implementation, the images at the same value of scale and level are not of identical sizes. Again, in the low pass error we compare the error resulting due to the original and low pass versions whereas for high pass error, the original image and the high pass images are compared. Thus, error values for the scale-space are much lower than that of the high pass error. To account for that, we use the ratio $(\gamma_{s,1})$ number of the elements used to form the co-occurrence histogram and the number of elements used to form high pass histograms. Secondly, the low pass images obtained from the scale-space and Wavelet decomposition undergo changes in the ranges of the coefficients. If the pixel values of the grayscale image are normalized between 0 and 1 by dividing the pixel values by 255 (for 8-bit grayscale images), scale-space preserves the range whereas Wavelet shifts the range from $-r_z$ to r_z, where r_z is a positive real number. At the same value of scale and level, energies of the low pass images also differ. This is also taken into account while designing γ_s through the incorporation of $\gamma_{s,2}$ provided in Eqn. 6.15. Therefore, the γ_s is calculated as

$$\gamma_s = \frac{1}{N}\gamma_{s,1}\gamma_{s,2}. \tag{6.13}$$

Dividing by $1/N$ ensures that γ_S values normalize Q_L over scales. γ_{s1} is calculated as

$$\gamma_{s,1} = \frac{\sum_p \sum_q n_{IC1_s}(p,q)}{\mathrm{num}(SD_{DS_{IG,z}})}.$$ (6.14)

where num(.) gives the number of elements of the argument. $\gamma_{s,2}$ is obtained as

$$\gamma_{s,2} = \sqrt{\frac{\mathrm{median}((AS_{IG,z}/r_z)^2)}{\mathrm{median}(\mathrm{cg}(IC_s)^2)}}$$ (6.15)

where cg(.) denotes the color to grayscale conversion (after division of the input color image by 255).

Final Form of the Method

After the errors Q_L, Q_{H1} and Q_{H2} are generated, they are combined to form the final quality score in the following manner. $\overline{Q_L}$, the mean of the error for low pass components Q_L multiplied by 100, is calculated. As discussed in Section 6.2.3, the ratio of kurtosis to standard deviation serves as an important visual cue and is supportive of the low pass error. Therefore, we calculate the two dimensional Fourier transform $\mathcal{IG}$ of the query image IG and find κ and σ from the log amplitude spectrum. The ratio (κ/σ) is multiplied with the low pass error. From the high pass errors Q_{H1} and Q_{H2}, we obtain total high pass error Q_H as

$$Q_H = (\mathrm{mean}(Q_{H1}) + \mathrm{std}(Q_{H1}) + \mathrm{mean}(Q_{H2}))/2.$$ (6.16)

where std(.) is the standard deviation and mean(.) denotes the mean operation. $\overline{Q_H}$, the mean of the error for high pass components Q_H multiplied by 100, is calculated. As we notice in Section 6.2.3, the standard deviation σ is indicative of the dispersion, and it increases for some distortions. The higher it is, the more significant should be the high pass error as it results due to the loss of low pass information. Hence, Q_H is multiplied by σ. Finally BIQES is formulated as,

$$QS = \kappa \frac{\overline{Q_L}}{\sigma} + \sigma \overline{Q_H}.$$ (6.17)

Discussion about the Parameters Involved

There are several parameters involved in the proposed method. First, the number of scales or levels N. We have used a constant $N = 4$ for all our experiments. The scale-space technique used Gaussian smoothing that requires a window size, mean and standard deviation. The Gaussian functions over the different scales have zero mean, with square windows of width $\{5, 7, 9, 11\}$ and the standard deviation values of $\{1, 2, 4, 8\}$. Next, in order to calculate the co-occurrence histogram, the neighborhood size needs to be mentioned. A square window of size 3×3 is used to calculate co-occurrence histogram.

The blockwise standard deviation is calculated for outer scale with a square block of length 11. Over the next 4 scales, the block size decreases as $\{9,7,5,3\}$. For the high pass error, we need to calculate the 1D histograms from the local standard deviation and gradient images. We need range and size of bins to form the histograms. Before Wavelet decomposition, the query image is converted to grayscale and normalized between 0 and 1. Since we used the 'bior4.4' filter from MATLAB, the maximum value of the coefficients in subsequent levels may increase. The maximum range after four scales is calculated as 35.11. Therefore, for gradient images the range of the histogram is [-70.22,70.22] and size of bins is $g\_hbin = 1/2000$. The range of the histogram for local standard deviation is lower and size of bins is higher depending on the experiments. The range of histogram for standard deviation image is [0,4] and size of bins is $sd\_hbin = 1/500$. The sizes of bins for standard deviation and gradient based histograms along with the proper scale value is decided from our analysis presented in Fig. 6.4. In this analysis, for scale values $N = 2, 3, 4$, the SROCC values in LIVE database are plotted for different combinations of bin sizes used for g_hbin and sd_hbin. The values used for g_hbin are $\{1/5000, 1/4500, 1/4000, ..., 1/500\}$. For each of these 10 values of g_hbin, sd_hbin takes values from $\{1/100, 1/200, ..., 1/1000\}$. Each unique combination of g_hbin and sd_hbin is represented by a unique index. Thus, we have 100 such indices. Therefore, the SROCC values are plotted against

these 100 indices in Fig. 6.4. As we see, with the increase of scale values, the performance gets better. The best performance is obtained for scale value $N = 4$, index number 65 pointing at combination of $g\_hbin = 1/2000$ and $sd\_hbin = 1/500$. Next comes η_z, which is used to weigh the high pass error over scales. Such weights are used in the experiments in [114]. Later, in [86], it is found that Gaussian distribution can also be used as weights. Therefore, a zero mean Gaussian vector of length 4 and standard deviation 1 is used to generate η_z values.

6.4 Experiments and Results

In this section, we present the experiments and the results of the proposed method. It is observed that, the hypothesis that over the scales the similarity between query images and the derived images at several scales decrease, is supported by most common distortions of additive noise, blur, jpeg2k and jpeg distortions. These are the most common types of image distortions available and several other IQA methods [29, 84, 122] have showed to work on these selected distortions only. Of course, this puts a restriction on the types of distortions. However, considering (a) the types of distortions available in the IQA databases do not form an exhaustive set, (b) BIQES can handle most common types of image distortions and (c) this is the first attempt on the hypotheses used for training free NR-IQA, the generality of the proposed method is not hampered. Four databases are used for the experiments: LIVE database [94], CSIQ database [43], TID2008 database [79] and MICT database [30]. In the first part of this section, we do the experiments on the the proposed method BIQES. In the next part, we provide a comparison study of the proposed method.

6.4.1 Experiments with the Proposed Method

The experiments in this section are carried out mainly in the LIVE database as it has all the four kinds of common distortions. We also involve the fastfading distortion of
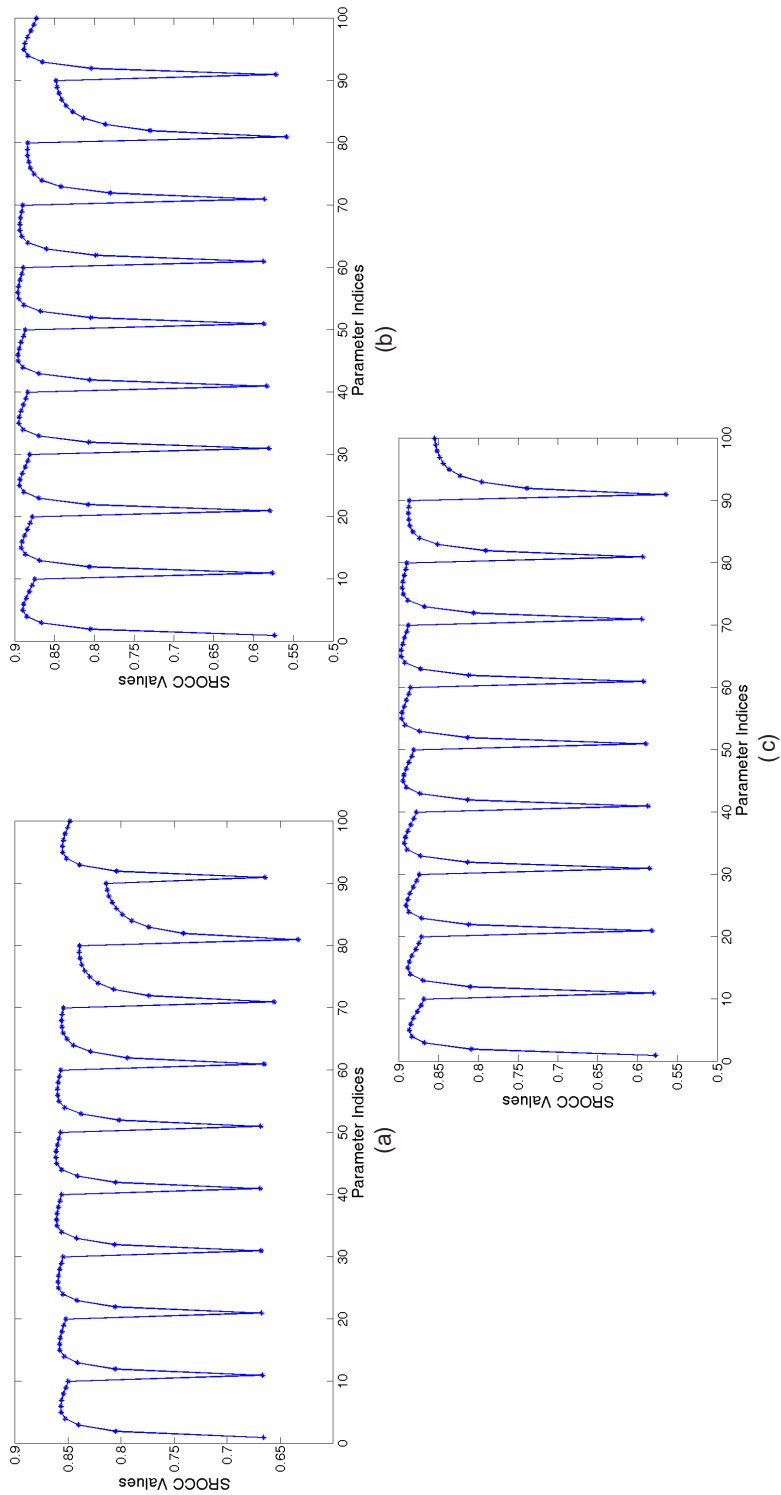
Figure 6.4: The changes in SROCC values for LIVE database using different number of scales (N) and sizes of bins used to compute the standard deviation and gradient based histograms. (a) $N = 2$ (b) $N = 3$ and (c) $N = 4$.

the LIVE database in these experiments. The experiments show the performance of the scores obtained from each of the low pass and high pass errors for all of the five types of distortions shown in the database. In the first experiment, the quality score resulting due to low pass error is checked against the subjective scores in the LIVE database. The results are presented in Fig. 6.5. The blue curves in figure represent the quality scores after mapping the scores obtained from the low pass errors with the DMOS. As expected, it is found that for additive white noise and blur, the correlation values are high, for jpeg, fastfading, they are moderate. However, the correlation is very low for jpeg2k images. Thus, the low pass error itself is not satisfactory measure of quality for jpeg2k images. We also find that this correlation is negative which means that it contradicts with the measure, whereas other distortions support it. Generally in literature, at the time of presenting these correlation values, absolute value is presented. Nevertheless, the sign has an important part to play in determining the overall performance of the quality measure in the database and in understanding its relationship with the individual distortions. The sign of the correlation will be again discussed in the later parts of the section. Also in this work, when no sign is indicated for a distortion in any database, it means that the four important distortions have correlation (SROCC) values of same sign. Thus, Fig. 6.5 shows that quality score due to low pass error is well correlated with some of the distortions. The quality score due to the high pass error is also checked against the subjective scores and the results are presented in Fig. 6.6. The correlation for the distortions like jpeg2k, jpeg and fastfading are high. The same for blur is low and additive Gaussian noise is negatively correlated with it. From the figures, the range of the values of the two quality scores also becomes clear. The score values due to low pass error are higher for additive Gaussian noise, blur and fastfading. For jpeg2k and jpeg, the score values are quite low. A look at the score values due to high pass error hints that the values for high pass error are very low for additive Gaussian noise and blur, but those values are high for jpeg2k, jpeg and fastfading. When these two scores are combined to form

the final score, we find from Fig. 6.7 that all of the distortions have high degree of correlation with the final scores. Also, all of the correlations have same sign showing that the final quality score has consistent representation for each type of distortion. The overall correlation is also shown to be high which assures that the ranges for high pass and low pass errors have been mutually compensated in the final score.

6.4.2 Comparison with Other Methods

We compare the proposed method with several existing FR-IQA and NR-IQA techniques such as PSNR, SSIM [113], BIQI [71], LBIQ [102], CBIQ [124], DIIVINE [72], BRISQUE [68], SRNSS [29], NIQE [70], LQF [69], QAC [122], and NRQI [46] in different experiments presented here. All of the NR-IQA techniques used training either with or without human subjective scores. The training and testing for these methods vary greatly. Here, we compare the scores reported in the respective papers with our methods in the LIVE database. Since all of the methods have used LIVE database, the comparison in LIVE database comprises of all of the aforesaid methods. As we can see from the comparison in the LIVE database, FR-IQA method SSIM perform better than the NR-IQA methods. The better performing NR-IQA methods use both human scores and distorted images for training in $4:1$ ratio from the database. QAC, LQF, NRQI and NIQE do not use human scores for training. Among these, QAC, LQF and NRQI use distorted and pristine images for training. However, NIQE uses only pristine images for training i.e learning MVG models from the pristine images. Hence, it is technically closest to the proposed method. Though, NIQE performs slightly better than BIQES in LIVE database, it uses training with pristine images while BIQES scores are totally training free. Also, BIQES perform better than QAC and NRQI in LIVE database. MICT databases has two distortions: jpeg2k and jpeg, BIQES performs better than NIQE and significantly better than QAC, PSNR and SSIM in this database. In CSIQ database, NIQE performs better than BIQES for four distortions. In TID2008 database, BIQES perform better than both of PSNR,
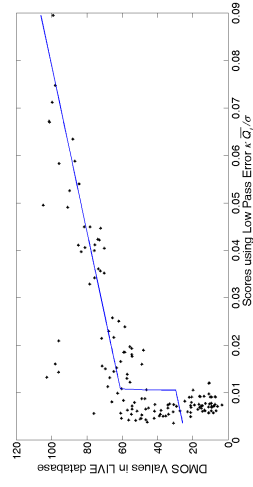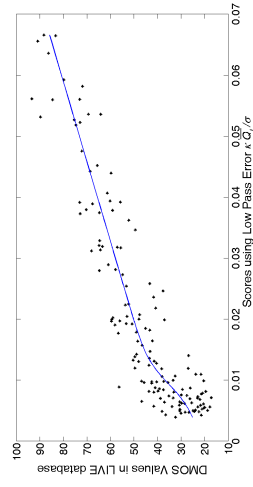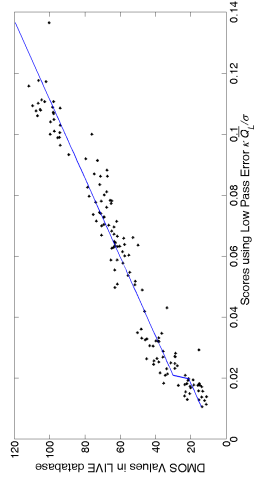
Figure 6.5: Performance on different types of distortions in LIVE database using the low pass error in the images having distortions: (a)jpeg2k (b)jpeg (c)wn (d)gblur and (e)fastfading.
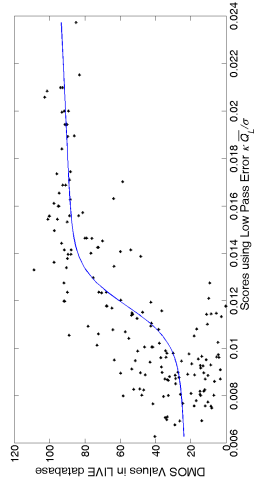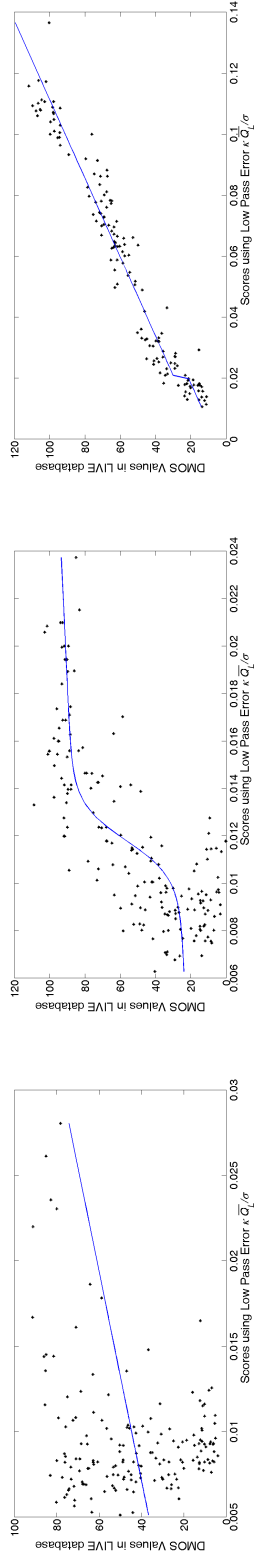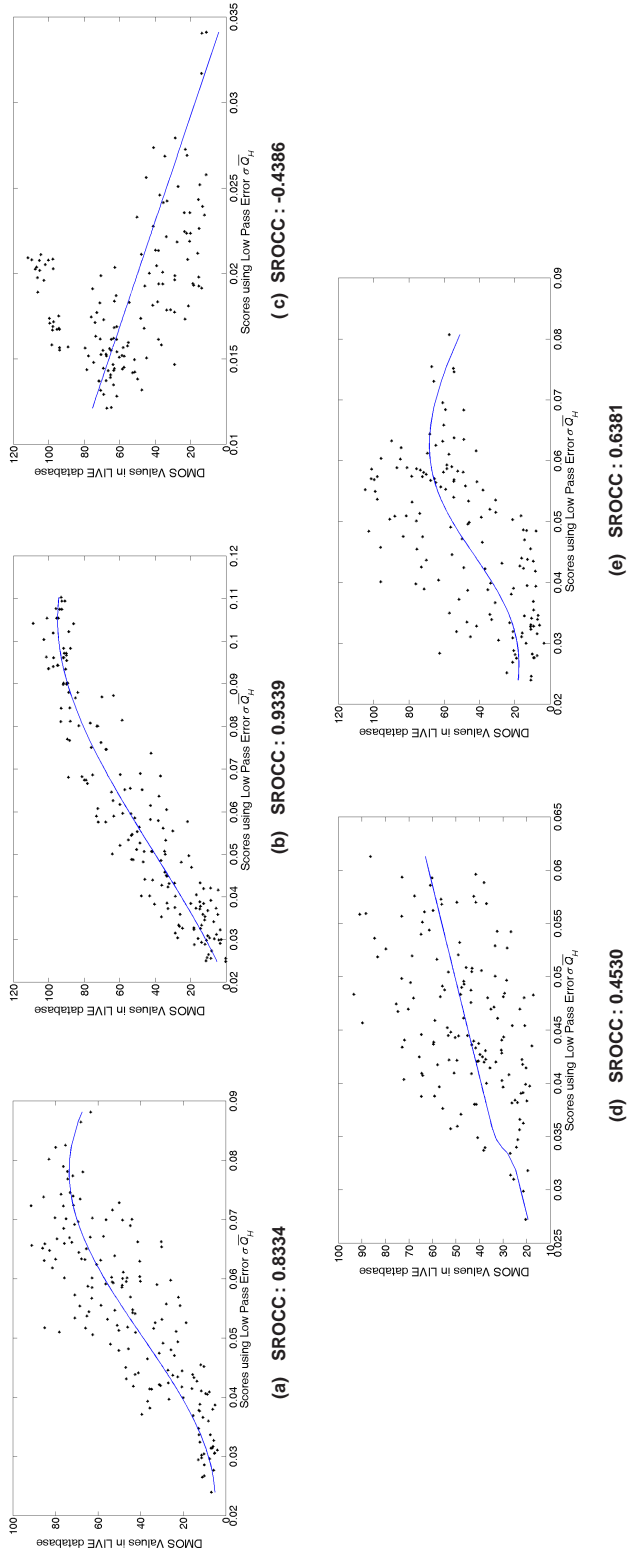
Figure 6.6: Performance on different types of distortions in LIVE database using the high pass errors in the images having distortions: (a) jpeg2k (b)jpeg (c)wn (d)gblur and (e)fastfading.
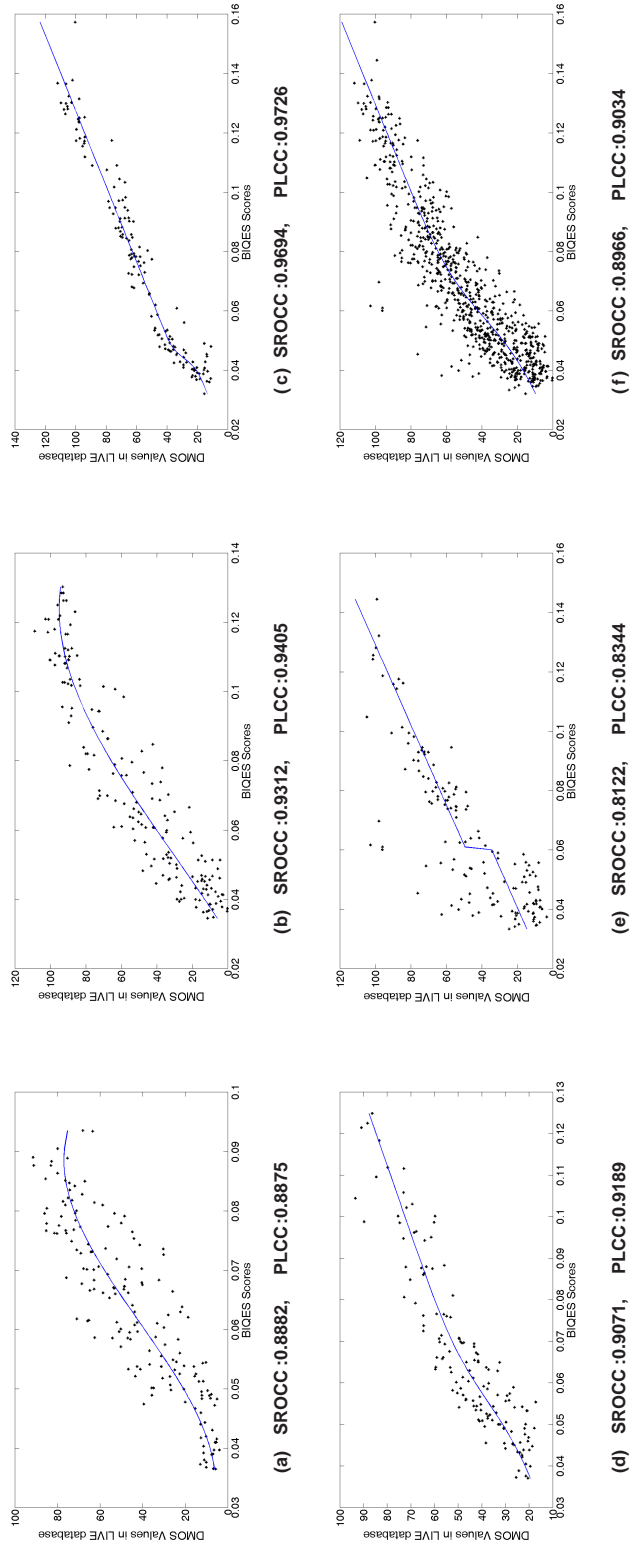
Figure 6.7: Performance of BIQES in LIVE database for distortions : (a)jpeg2k (b)jpeg (c)wn (d)gblur and (e)fastfading and (f) all distortions taken together.

NIQE and QAC taking four distortions together. It is also important to mention that the value of n_r, used to calculate the CSF is used as 2 in LIVE database, 3 in CSIQ and TID2008 databases (TID2008 database has varying viewing parameters, 3 is taken considering the mean condition) and 4 in MICT database depending on the parameters given in these databases.

6.4.3   A Look at the Common Limitations

We are going to look at some of the important analyses as brought out by our research. We notice some similarity in the failure patterns of the methods QAC, NIQE and BIQES. QAC and NIQE are trained without human subjective scores. QAC uses both pristine and distorted images for training and NIQE relies on pristine images for training while BIQES does not use any training. QAC does not rely on NSS features but works by forming local quality maps and NIQE measures deviations from the model fitted from training NSS properties of pristine images. BIQES combine low pass error and high pass error. In other words, these NR-IQA techniques differ a lot in their ways of functioning. All of them work on four common distortions. Their failures on certain omitted distortions are also tied by a common thread. Among the databases used in the experiments, all distortions of the LIVE and MICT databases are used but in CSIQ and TID2008 databases some distortions are dropped. A closer look at CSIQ database reveals that the distortions contrast and pink noise are dropped. In TID2008, 17 distortions are enlisted; however, only 4 of them are used. The performance of the aforesaid quality measures with the omitted individual distortions is discussed in Table 6.6. In CSIQ database, all of QAC, NIQE and BIQES perform poorly for contrast and fnoise. The negative sign here indicates that a particular distortion is inversely correlated with the IQA method compared to most of the distortions in the same database. Thus, inclusion of these distortions will decrease the overall performance of a method though it performs favorably in a set of distortions. In TID2008 database, all of the methods suffer from a sign

Table 6.2: Performance comparison of BIQES in the LIVE database.

	TYPE	Training with human opinion Scores	Training with pristine /distorted Images	SROCC	PLCC	Comments
PSNR	FR	N	N	0.8756	0.8722	all distorted images from LIVE database used for testing
SSIM	FR	N	N	0.9479	0.9449	all distorted images from LIVE database used for testing
BIQI	NR	Y	Y	0.8195	0.8205	train test ratio 4:1, median from 1000 such combinations
CBIQ	NR	Y	Y	0.8954	0.8955	train test ratio 4:1, median from 1000 such combinations
LBIQ	NR	Y	Y	0.9063	0.9087	train test ratio 4:1, median from 1000 such combinations
BLIINDS-II	NR	Y	Y	0.9124	0.9164	train test ratio 4:1, median from 1000 such combinations
DIIVINE	NR	Y	Y	0.9250	0.9270	train test ratio 4:1, median from 1000 such combinations
BRISQUE	NR	Y	Y	0.9395	0.9309	train test ratio 4:1, median from 1000 such combinations
SRNSS	NR	Y	Y	0.9304	0.9318	train test ratio 4:1, median from 1000 such combinations
QAC	NR	N	Y	0.8683	0.8637	trained using patches from ten images from the Berkeley Segmentation database, all distorted images from LIVE database used for testing
LQF	NR	N	Y	0.80	0.79	train test ratio 4:1 (in terms of disjoint reference image sets), median from 1000 such combinations
NRQI	NR	N	Y	0.8263	0.8350	all distorted images from LIVE database used for testing
NIQE	NR	N	Y	0.9062	0.9040	trained using pristine images from Berkeley database and free Flickr data all distorted images from LIVE database used for testing
BIQES	**NR**	N	N	**0.8966**	**0.9034**	all distorted images from LIVE database used for testing

Table 6.3: Performance comparison of BIQES in the MICT database.

	SROCC	KROCC	PLCC	MAE	RMSE
PSNR	0.6132	0.4443	0.6498	0.7746	0.9517
SSIM	0.7870	0.5922	0.7995	0.5649	0.7551
QAC	0.5189	0.3667	0.5388	0.8704	1.0563
NIQE	0.8094	0.6093	0.8185	0.5552	0.7211
BIQES	**0.8271**	**0.6361**	**0.8259**	**0.5191**	**0.7119**

Table 6.4: Performance comparison of BIQES in the CSIQ database.

	SROCC	KROCC	PLCC	MAE	RMSE
PSNR	0.9218	0.7489	0.9065	0.0830	0.1198
SSIM	0.8766	0.6808	0.8465	0.1048	0.1509
QAC	0.8415	0.6440	0.8734	0.1051	0.1377
NIQE	0.8698	0.6835	0.8855	0.0990	0.1319
BIQES	**0.8561**	**0.6584**	**0.8879**	**0.1020**	**0.1301**

Table 6.5: Performance comparison of BIQES in the TID2008 database.

	SROCC	KROCC	PLCC	MAE	RMSE
PSNR	0.7985	0.5982	0.7572	0.8139	1.0339
SSIM	0.9199	0.7506	0.9251	0.4675	0.6017
QAC	0.7967	0.6085	0.7829	0.7367	0.9871
NIQE	0.7832	0.5803	0.7940	0.7752	0.9627
BIQES	**0.8223**	**0.6209**	**0.8103**	**0.6991**	**0.9289**

Table 6.6: Limitations of QAC, NIQE and BIQES with some distortions.

Database	Distortion	QAC	NIQE	BIQES
CSIQ	contrast	-0.2241	0.2273	-0.3140
	fnoise	-0.0022	0.2988	0.3133
TID2008	awgn-pink	0.6624	0.6799	0.7021
	spatial-corr-noise	0.1353	0.7419	0.6662
	masked noise	0.6200	0.7577	0.6467
	high frequency noise	0.7985	0.8539	0.7750
	impulse noise	0.8188	0.6877	0.6662
	quantization noise	0.5690	0.8111	0.2351
	denoising	0.4198	0.6051	0.7935
	jpeg-trans-error	0.0190	0.1228	-0.3138
	jpeg2k-trans-error	0.3672	0.4934	0.5368
	pattern-noise	-0.0058	-0.0162	0.0287
	blockwise-distortion	0.3275	0.1555	0.4285
	intensity-shift	0.2704	-0.1371	0.2148
	contrast change	-0.2377	0.0379	-0.6252

reversal in two types of distortions. All of the methods perform poorly with pattern noise, JPEG transmission errors (jpeg-trans-error), JPEG2000 transmission errors (jpeg2k-trans-error) and image contrast change. Thus, these are some areas where these IQA methods are not sufficient enough to work. This is worth-studying as we find that all of these methods have performed without the aid of human subjective scores for the same set of distortions and identify their common problem areas. This is particularly important for BIQES as it is training free but has its pros and cons similar to the earlier methods.

6.5 Chapter Summary

We present a method of no reference image quality assessment. The method is absolutely training free as neither subjective human scores nor pristine or distorted

images are used to make the method aware of quality of distorted or pristine images. The quality score increases with the increase of distortions by evaluating the query image only. Natural images have redundancy over scales. Using only this information, the query image is processed in different scales and information loss over scales is primarily used to quantify the distortions. Some properties of natural images present in image scale-space, Wavelet transform and Fourier transform are exploited to formulate the quality score.

The proposed technique BIQES has been compared with the state-of-the-art and related IQA methods in four databases. It is found that NR-IQA techniques developed on subjective scores based training are best performers but a large proportion of images are required to train them. NR-IQA techniques that rely on pristine and distorted images but not on human subjective scores for training, have achieved comparable performance with the earlier ones. BIQES compares favorably with these methods without any training. Also, the common limitations of BIQES and other NR-IQA are pointed out and they may provide a direction of working towards a better technique for totally training free BIQA.

Chapter 7

Conclusion

In a nutshell, this dissertation presents some ideas in the field of IQA and its related problems. We started with the presentation of a relevant introduction to the problem of IQA and with the motivation to pursue the search for better solutions to the related problems. With the help of a discussion on the relevant literature related to the different areas, we broadly identified the problems and defined the course of the research work presented in the dissertation. After setting up the necessary base for presentation of our ideas in first two chapters, the next four chapters are developed using the main attributes of the our research. We have presented ideas in the areas of FR-IQA, specific-purpose and general purpose NR-IQA. In this concluding chapter, we enlist the contributions and limitations of our work and discuss the possible directions for future work.

7.1 Contributions and Limitations

We present the contributions and limitations of the our research works under three broad topics as follows.

7.1.1 FR-IQA

We have proposed three techniques for FR-IQA. Two of these techniques are modifications made on an existing framework by changing features. In the first technique, instead of the phase congruency features, phase deviation sensitive energy features are used to improve the overall performance of the technique. The second work uses

different spectral saliency based techniques in the same framework and studies their performance on six databases. The third technique proposes a novel framework with simple visual features. The salient points of these works are presented below as follows:

- FR-IQA techniques mainly aim to improve the generalized performance of any IQA method over several types of distortions. Hence, distortion-wise performance analysis of the methods are carried out in our works. We find that, so far, there exists no technique in the literature that can ensure good performance for all of the distortions in the vast set of distortions considered in our experiments. Most of the methods work well in the databases with less than 6 types of distortions. However, performance of these methods in databases having more types of distortions needs improvement. Also, a set of distortions that needs improved methods for perceptual quality evaluation are also found out. Hence improving the general ability of FR-IQA techniques remains an open problem.

- The first and second works, both reveal the parameter dependence of the aforesaid framework using saliency features.

- The first two proposed methods fix the framework and search for the best generalized method. Though performance in some databases improve with the proposed method, it is not guaranteed for every database. This may be attributed to the fact that every feature has inclination to some specific property of any image. When that property is highlighted, several other properties are subdued. However, the overall performance in all databases improves by 0.5% and this is important given the complexity of the task.

- Echoing the point mentioned earlier, the next proposed methods GLD-SR and GLD-PFT, improve the performance on CSIQ and IVC databases by 2% and 0.4% respectively. Also they improve the overall performance in terms of

weighted and direct averages computed over all six databases by 0.6%. These figures are significant given the fact they raise the performance in two databases and improve the overall performance compared to the existing techniques. The statistical analysis shows that, considering individual comparison with all 13 different techniques in six databases, the proposed techniques are able to provide improvement in more than 50% of the cases. Also, the framework introduced through GLD-SR and GLD-PFT has lesser dependence on the parameters. Hence, an effective and improved combination of visual features has resulted from the proposed framework.

7.1.2 Specific Purpose NR-IQA

We have proposed a blind perceptual sharpness evaluation approach in natural images. Since, blurring is one of the main reasons behind the loss of sharpness in any image, we evaluate the proposed approach in several IQA databases having blurred images. Besides the evaluation of perceptual sharpness, the approach also generates a sharpness map to highlight the sharper regions in any image. Several experiments are carried out to show the performance of the proposed methods HPFSM and UWTSM against different blur levels across different images. We show that the methods improve over the performance of the state-of-the-art perceptual sharpness detection techniques using quantitative analysis. The salient points of the proposed technique are summarized below.

- We have discussed in Chapter 2, Section 2.2 about the various types of visual features that have been used to detect and evaluate perceptual quality in to the presence of blur or sharpness. However, the proposed method uses simple combination of high frequency features and local standard deviation and combines them as per the psychometric function presented in [81].

- The method evaluates the presence and absence of blur and also evaluates the

perceived quality due to the presence and loss of sharpness.

- The performance of the method is analyzed for varying values of the exponent used in the process. The values of the exponent greater or equal to 1 are shown to improve the performance.

- Two types of high frequency features extracted using Gaussian high pass filter and Wavelets are shown to perform better than the state-of-the-art blind sharpness evaluation techniques. The proposed versions of the technique have improved 3% in average performance in terms of SROCC. This is significant as the ideal performance is 10.2% higher than the performance of the existing state-of-the-art techniques. As already seen with the FR-IQA methods, some of them have very high correlation with the subjective scores. Thus, blind sharpness evaluation needs more improvement to reach the high performance exhibited by some FR-IQA methods in some of the databases.

7.1.3 General Purpose NR-IQA

We have proposed a general purpose NR-IQA technique, called BIQES, for perceptual quality evaluation in natural images. It works for four distortions. It demonstrates competitive performance with the state-of-the-art methods. The salient points of the method are presented below.

- The proposed approach is a general purpose NR-IQA technique which does not incorporate any training with the pristine or distorted images or quality scores.

- The principal visual cue is the difference of propagation of image information through different scales between the pristine and distorted images. The multiscale information is captured by the scale space decomposition and Wavelet transform on the image.

- The method relies on image scale features for calculating with the high pass and low pass errors using co-occurrence histograms and Wavelet transform. These errors are combined using the kurtosis and standard deviation of Fourier amplitude spectrum to arrive at the objective score. It is shown using graphical analysis that each of these errors computed using various visual features has a part to play in shaping up the proposed approach. Thus, the visual feature combination is shown to have satisfactory success for training free BIQA as well.

- The property of being training free equips the method with the ability to eliminate time and effort for training. At the same time, it imposes restriction on the method to simultaneously handle a vast number of distortions. This phenomena is also observed with the methods that have tried to eliminate training with pristine images and human scores from their working procedures. Thus, the limitations of these methods are also identified.

7.1.4 General Summary

In general, when we look at the dissertation, we find a variety of features and their combinations have been used for the purpose. The main features used are phase deviation sensitive energy features, saliency maps based on different types of spectral saliency measures, local gradient, local variance, high frequency based features, low frequency based features, co-occurrence histograms, scale-space decomposition, Wavelet decomposition, Fourier transform and kurtosis. To summarize, we have used several image features throughout the different problems we have worked with. The working principles of all the proposed methods are based on plausible hypotheses rather than the actual functioning of the HVS, since it is unknown. Thus our approaches highlight that potential combination of features can be used in IQA and they support the overarching avenue to find solution for the respective sub-areas in IQA. They also find new opportunities in the sub-areas which are discussed next.

7.2 Scope for Future Work

With each of proposed works in the dissertation, we have possible future directions for carrying forward the results and findings of our work. As already mentioned, we have used several types of features and their combinations supported by different hypotheses in our work. The same features can be used in different ways to facilitate the application of different hypotheses. Thus, this work provides necessary evidence for exploring the wide set of features already existing in the field of image processing and computer vision today towards improving the state-of-the-art IQA and its different sub-areas. Next, we discuss in details the future directions of work revealed in each of the specific areas in IQA.

7.2.1 FR-IQA

For FR-IQA, we made efforts to enhance the generalized performance of existing techniques. Next, we have proposed a new framework by combining global and local distortion features to develop an improved IQA technique. The future directions of work are found as follows:

- Several other existing frameworks can be modified by using different features to improve their performance. As we know that the generalized performance of FR-IQA techniques is very important so that a particular technique can be used across several distortions altogether. By varying features in the existing techniques, it may be possible to improve their performance.

- We have used specific features for generating global and local distortion from the images. A possible future work may be directed towards using different types of features for generating global and local distortion measures from the images to improve the performance.

7.2.2   Specific Purpose NR-IQA

The proposed technique on perceptual sharpness detection improves over the performance of the state-of-the-art techniques. The future directions of work related to this research are found as follows:

- The high frequency content in the image has been extracted by two techniques. Several other features can be used for extracting the high frequency content and their effect can be studied for perceptual sharpness evaluation.

- A single pooling strategy has been applied to arrive at the final score. Application of different pooling strategies may be done to improve the performance.

- With our proposed approach, we have relied on a single psychometric model. Different psychometric models based hypotheses can be developed and implemented to improve the performance in blind perceptual sharpness evaluation.

7.2.3   General Purpose NR-IQA

For BIQA, we have presented an approach which is completely training free. It works on four types of distortions taken together. Some future directions of work related to this research are listed as follows:

- Different features may be used to improve the general applicability of the technique so that it can perform well across several distortions taken together.

- The work shows that completely training-independent IQA is feasible across selected distortions. However, research based on different hypotheses and features needs to be carried out in order to establish its potential.

References

[1] I. Avcibas, N. Memon, and B. Sankur. Steganalysis using image quality metrics. *IEEE Trans. on Image Processing*, 12(2):221–229, 2003. [10]

[2] C. F. Batten. Autofocusing and astigmatism correction in the Scanning Electron Microscope. Master's thesis, Univ. of Cambridge, U.K., 2000. [22]

[3] P. J. Bex and W. Makous. Spatial frequency, phase and the contrast of natural images. *Journal of the Optical Society of America A*, 19(6):1096–1106, 2011. [65, 91]

[4] P. Le Callet and F. Autrusseau. Subjective quality assessment-IVC database, 2005. http://www.irccyn.ec-nantes.fr/ivcdb/. [8, 96]

[5] C. Cavaro-Menard, L. Zhang, and P. Le Callet. Diagnostic quality assessment of medical images: Challenges and trends. In *European Workshop on Visual Information Processing*, pages 277–284, 2010. [10]

[6] J. Caviedes and F. Oberti. A new sharpness metric based on local kurtosis,edge and energy information. *Signal Processing : Image Communication*, 19(2):147–161, 2004. [23, 112]

[7] D. M. Chandler. Seven challenges in image quality assessment : Past, present and future research. *ISRN Signal Processing*, 2013:905685:1–905685:53, 2013. [11]

[8] D. M. Chandler and S. S. Hemami. VSNR: A Wavelet-based visual signal-to-noise-ratio for natural images. *IEEE Trans. on Image Processing*, 16(9):2284–2298, 2007. [8, 19]

[9] G.-H. Chen, C.-L. Yang, and S.-L. Xie. Gradient-based structural similarity for image quality assessment. In *IEEE Int. Conference on Image Processing*, pages 2929–2932, 2006. [18]

[10] M. J. Chen and A. C. Bovik. No-reference image blur assessment using multi-scale gradient. *Eurasip Journal on Image and Video Processing*, 2011:3, 2011. [23]

[11] Y. C. Chung, J. M. Wang, R. R. Bailey, S. W. Chen, and S. L. Lang. A non-parametric blur measure based on edge analysis for image processing applications. In *IEEE Conference on Cybernetics and Intelligent Systems*, volume 1, pages 356–360, 2004. [22]

[12] P. Corriveau. Video quality testing. In H. R. Wu and K. R. Rao, editors, *Digital Video Image Quality and Perceptual Coding*, pages 125–153. CRC Press, Taylor & Francis Group, 2006. [4]

[13] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas. The Blur Effect : Perception and estimation with a new no-reference perceptual blur metric. In *SPIE Elect. Symp. Conf. Human Vis. and Elect. Imag.*, volume 649201, pages 64920I–649201I–11, 2007. [22]

[14] S. Daly. The visible difference predictor: An algorithm for the assessment of image fidelity. In *Proc. SPIE*, volume 1616, pages 2–15, 1992. [18]

[15] E. Dumic, S. Grgic, and M. Grgic. New image quality measure based on Wavelets. *Journal of Electronic Imaging*, 19(1):011018, 2010. [19]

[16] S. Erasmus and K. Smith. An automatic focusing and astigmatism correction system for the SEM and CTEM. *Journal of Microscopy*, 127:185–199, 1982. [22]

[17] Y. Fang, W. Lin, B.-S. Lee, C.-T. Lau, Z. Chen, and C.-W. Lin. Bottom-up saliency detection model based on human visual sensitivity and amplitude spectrum. *IEEE Trans. on Image Processing*, 20(5):1185–1198, 2011. [64]

[18] R. Ferzli and L. J. Karam. A no-reference perceptual image sharpness metric based on the notion of just noticeable blur. *IEEE Trans. on Image Processing*, 21(7):717–728, 2012. [22, 100, 105, 107]

[19] D. Field and N. Brady. Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research*, 37(23):3367–3383, 1997. [23]

[20] D. J. Field. Relations between the statistics of natural images and the response propertes of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379–2394, 1987. [35]

[21] L. Firestone, K. Cook, N. Talsania, and K. Preston. Comparison of autofocus methods for automated microscopy. *Cytometry*, 12:195–206, 1991. [23]

[22] S. Gabarda and G. Cristóbal. Blind image quality assessment through anisotropy. *Journal of the Optical Society of America A*, 24(12):B42–B51, 2007. [22]

[23] X. Gao, F. Gao, D. Tao, and X. Li. Universal blind quality assessment metrics via natural scene statistics and multiple kernel learning. *IEEE Trans. on Neural Networks and Learning Systems*, 24(12):2013–2026, 2013. [112]

[24] P. Gorley and N. Holliman. Stereoscopic image quality metrics and compression. In *SPIE Stereoscopic Displays and Applications*, volume 6803, page 680305, 2008. [10]

[25] N. Graham. *Visual Pattern Analyzers*. Oxford University Press, New York, 1989. [92]

[26] Video Quality Experts Group. Methodology for the subjective assessment of the quality of television pictures. ITU-R BT.500-11, 2002. http://www.itu.int. [2]

[27] C. Guo, Q. Ma, and L. Zhang. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. [49, 54, 64]

[28] R. Hassen, Z. Wang, and M. M. A. Salama. Image sharpness assessment based on local phase coherence. *IEEE Trans. on Image Processing*, 22(7):2798–2810, 2013. [23, 24, 100, 105, 107]

[29] L. He, D. Tao, X. Li, and X. Gao. Sparse representation for blind image quality assessment. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1146–1153, 2012. [25, 124, 127]

[30] Y. Horita, K. Shibata, Y. Kawayoke, and Z. M. P. Sazzad. MICT image quality evaluation database [Online]. Available: http://mict.eng.u-toyama.ac.jp/mict/index2.html. [8, 124]

[31] X. Hou, J. Harel, and C. Koch. Image signature: Highlighting sparse salient regions. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(1):194–201, 2012. [50]

[32] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007. [20, 32, 48, 64, 72]

[33] J. Huang and Y. Q. Shi. Image-adaptive watermarking using visual models. *Electronics Letters*, 34(8):748–750, 1998. [10]

[34] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene anaysis. *IEEE Trans. Pat. Ana. Mach. Intell.*, 20(11):1254–1259, 1998. [62]

[35] B. Jähne, H. Haubecker, and P. Geibler. *Handbook of Computer Vision and Applications*. New York: Academic, 1999. [30, 72]

[36] R. A. Jarvis. Focus optimisation criteria for computer image processing. *Microscope*, 24(2):163–180, 1976. [22]

[37] J. J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984. [111]

[38] P. Kovesi. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, 1(3):1–26, 1999. [19, 29, 31, 35]

[39] A. Koz and A. A. Alatan. Oblivious spatio-temporal watermarking of digital video exploiting the human visual system. *IEEE Trans. on Circuits and Systems for Video Technology*, 18(3):326–337, 2008. [10]

[40] E. P. Krotkov. *Active Computer Vision by co-operative Focus and Stereo*. Springer-Verlag, 1989. [22]

[41] Y. Lai and C. J. Kuo. Image quality measurement using the Haar Wavelet. In *SPIE: Wavelet Applications in Signal and Image Processing V*, 1997. [19]

[42] E. C. Larson and D. M. Chandler. Categorical subjective image quality (CSIQ) database. http://vision.okstate.edu/csiq, 2009. [7, 66, 68, 70, 77]

[43] E. C. Larson and D. M. Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of Electronic Imaging*, 19(1):011006, 2010. [19, 20, 96, 120, 124]

[44] C. Li and A. C. Bovik. Content-partitioned structural similarity index for image quality assessment. *Signal Processing: Image Communication*, 25(7):517–526, 2010. [18]

[45] C. Li, A. C. Bovik, and X. Wu. Blind image quality assessment using a general regression neural network. *IEEE Trans. on Neural Networks*, 22(5):793–799, 2011. [26]

[46] C. Li, Y. Ju, A. C. Bovik, X. Wu, and Q. Sang. No-training, no-reference image quality index using perceptual features. *Optical Engineering*, 52(5):057003–1–057003–6, 2013. [26, 27, 109, 127]

[47] J. Li, M. D. Levine, X. An, X. Xu, and H. He. Visual saliency based on scale-space analysis in the frequency domain. *IEEE Trans. Pattern Anal. Mach. Intell.*, 35(4):996–1010, 2012. [50]

[48] S. Li, F. Zhang, L. Ma, and K. N. Ngan. Image quality assessment by separately evaluating detail losses and additive impairments. *IEEE Trans. on Multimedia*, 13(5):935–949, 2011. [20]

[49] Y. Li and C.-C. Jay Kuo. *Image Databses :Search and Retrieval of Digital Imagery*, chapter Introduction to Content-Based Image Retrieval-Overview of Key Techniques, pages 261–284. John Wiley & Sons, Inc., 2002. [90]

[50] Y. H. Lin and J. L. Wu. Quality assessment of stereoscopic 3d image compression by binocular integration behaviors. *IEEE Trans. on Image Processing*, 23(4):1527–1542, 2014. [10]

[51] T. Lindenberg. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):224–270, 1994. [110]

[52] A. Liu, W. Lin, and M. Narwaria. Image quality assessment based on gradient similarity. *IEEE Trans. on Image Processing*, 21(4):1500–1512, 2012. [19, 38, 63]

[53] J. Liu and P. Moulin. Information-theoretic analysis of interscale and intrascale dependencies between image Wavelet coefficients. *IEEE Trans. on Image Processing*, 10(11):1647–1658, 2002. [112]

[54] R. Liu, Z. Li, and J. Jia. Image partial blur detection and classification. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. [23]

[55] Y. Liu, J. Wang, S. Cho, A. Finkelstein, and S. Rusinkiewicz. A no-reference metric for evaluating the quality of motion deblurring. *ACM Trans. on Graphics*, 32(6):175:1–175:12, 2013. [10]

[56] Z. Liu and R. Laganière. Phase congruence measurement for image similarity assessment. *Pattern Recognition Letters*, 28(1):166–172, 2007. [43, 44, 45, 47, 57]

[57] S. Lu and J.-H. Lim. Saliency modeling from image histograms. In *European Conference on Computer Vision*, pages 321–332, 2012. [114, 115]

[58] W. Lu, K. Zeng, D. Tao, Y. Yuan, and X. Gao. No-reference image quality assessment in contourlet domain. *Neurocomputing*, 73(4–6):784–794, 2010. [25]

[59] J. Lubin. The use of psychophysical data and models in the analysis of display system performance. In A. B. Watson, editor, *Digital Images and Human Vision*, pages 163–178. The MIT Press, Cambridge, MA, 1993. [18]

[60] J. Lubin. A visual discrimination mode for system design and evaluation. In E. Peli, editor, *Visual Models for target Detection and Recognition*, pages 207–220. World Scientific Publishers, Singapore, 1995. [18]

[61] Q. Ma and L. Zhang. Image quality assessment with visual attention. In *International Conference on Pattern Recognition*, pages 1–4, 2008. [48]

[62] Q. Ma and L. Zhang. Saliency-based image quality assessment criterion. In *Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues (LNCS)*, volume 5226, pages 1124–1133, 2008. [19]

[63] Q. Ma and L. Zhang. Saliency-based image quality assessment criterion. In *International Conference on Intelligent Computing,*, pages 1124–1133, 2008. [48, 64]

[64] X. Marichal, W. Ma, and H. J. Zhang. Blur determination in the compressed domain using dct information. In *IEEE Int. Conference on Image Processing*, pages 386–390, 1999. [23]

[65] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi. A no-reference perceptual blur metric. In *IEEE Int. Conference on Image Processing*, pages 57–60, 2002. [22]

[66] J. Miao, D. Huo, and D. L. Wilson. Quantitative image quality evaluation of mr images using perceptual difference models. *Medical Physics*, 35(6):2541–2553, 2008. [10]

[67] A. Mittal, A. K. Moorthy, and A. C. Bovik. Automatic parameter prediction for image denoising algorithms using perceptual quality features. In *SPIE Human Vision and Electronic Imaging*, volume 8291, page 82910G, 2012. [10]

[68] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *IEEE Trans. on Image Processing*, 21(12):4695–4708, 2012. [25, 26, 100, 105, 107, 127]

[69] A. Mittal, G. S. Muralidhar, and A. C. Bovik. Blind image quality assessment without human training using latent quality factors. *IEEE Signal Processing Letters*, 19(2):75–78, 2012. [26, 127]

[70] A. Mittal, R. Soundararajan, and A. C. Bovik. Making a 'completely blind' image quality analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2013. [26, 109, 127]

[71] A. K. Moorthy and A. C. Bovik. A two-step framework for constructing blind image quality indices. *IEEE Signal Processing Letters*, 17(5):513–516, 2010. [25, 27, 109, 127]

[72] A. K. Moorthy and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Trans. on Image Processing*, 20(12):3350–3364, 2011. [25, 127]

[73] M. C. Morrone and D. C. Burr. Feature detection in human vision: A phase dependent energy mode. In *Royal Society London B*, volume 235, pages 221–245, 1988. [32]

[74] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6(5):303–313, 1987. [32]

[75] N. Narvekar and L. J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). *IEEE Trans. on Image Processing*, 20(9):2678–2683, 2011. [22, 100, 105, 107]

[76] A. Ninassi, O. LeMeur, P. LeCallet, and D. Barba. On the performance of human visual system based image quality assessment metric using Wavelet domain. In *SPIE Human Vision and Electronic Imaging*, volume 6806, pages 680610–680610–12, 2008. [19]

[77] A. Oliva, A. Torralba, M. S. Castelhano, and J. M. Henderson. Top-down control of visual attention in object detection. In *International Conference on Image Processing*, pages 253–256, 2003. [62]

[78] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti. A no-reference quality metric for measuring image blur. In *Int. Symposium on Signal Processing and its Applications*, pages 469–472, 2003. [22]

[79] N. Ponomarenko, V. Lukin, A. Zelensky, K. Eziazarian, M. Carli, and F. Battisti. TID2008- A database for evaluation of full reference image quality metrics. *Advances of Modern Radioelectronics*, 10:30–45, 2009. [7, 96, 124]

[80] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin. On between-coefficient contrast masking of DCT basis functions. In *Int. Conference on Video Processing and Quality Metrics for Consumer Electronics*, 2007. [16, 44]

[81] J. G. Robson and N. Graham. Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, 21(3):409–418, 1981. [93, 138]

[82] D. L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398, 1996. [111]

[83] D. L. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6):814–818, 1994. [112]

[84] M. Saad, A. C. Bovik, and C. Charrier. Blind image quality assessment: A natural scene statistics approach in the dct domain. *IEEE Trans. on Image Processing*, 21(8):3339–3352, 2012. [25, 104, 105, 107, 124]

[85] N. G. Sadaka, L. J. Karam, R. Ferzli, and G. P.Abousleman. A no-reference perceptual image sharpness metric based on saliency weighted foveal pooling. In *IEEE Int. Conference on Image Processing*, pages 369–372, 2008. [23]

[86] A. Saha, G. Bhatnagar, and Q. M. J. Wu. SVD filter based multiscale approach for image quality assessment. In *International Conference on Multimedia and Expo Workshop*, pages 43–48, 2012. [124]

[87] A. Saha, G. Bhatnagar, and Q. M. J. Wu. Mutual spectral residual approach for multi-focus image fusion. *Digital Signal Processing : A Review Journal*, 23(4):1121–1135, 2013. [10]

[88] R. Samadani, T. A. Mauer, D. M. Berfanger, and J. H. Clark. Image thumbnails that represent blur and noise. *IEEE Trans. on Image Processing*, 19(12):363–373, 2010. [114]

[89] B. Schauerte and R. Stiefelhagen. Predicting human gaze using quaternion DCT image signature saliency and face detection. In *Workshop on the Applications of Computer Vision*, 2012. [50]

[90] B. Schauerte and R. Stiefelhagen. Quarternion-based spectral saliency detection for eye fixation prediction. In *European Conference on Computer Vision*, 2012. [50]

[91] H. R. Sheikh and A. C. Bovik. Image information and visual quality. *IEEE Trans. on Image Processing*, 15(2):430–444, 2006. [19]

[92] H. R. Sheikh, A. C. Bovik, and G. Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. on Image Processing*, 14(12):2117–2128, 2005. [19]

[93] H. R. Sheikh, M. F. Sabir, and A. C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. on Image Processing*, 15(11):3440–3451, 2006. [9]

[94] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. LIVE image quality assessment database release 2. http://live.ece.utexas.edu/research/quality. [6, 37, 39, 96, 99, 113, 116, 124]

[95] A. Shnayderman, A. Gusev, and A. M. Eskicioglu. An SVD-based grayscale image quality measure for local and global assessment. *IEEE Trans. on Image Processing*, 15(2):422–429, 2006. [19]

[96] J. Silvestre-Blanes. Structural similarity image quality reliability: Determining parameters and window size. *Signal Processing Elsevier*, 91(4):1012–1020, 2011. [19]

[97] E. P. Simoncelli. Modeling the joint statistics of images in the Wavelet domain. In *SPIE 44th Annual Meeting*, pages 188–195, 1999. [112]

[98] E. E. Smith and S. M. Kosslyn. *Cognitive Psychology : Mind and Brain.* Pearson/Prentice Hall, 2007. [62]

[99] O. H. Schade Sr. Image quality: A comparison of photographic and television systems. *SMPTE Journal*, 96(6):567–648, 1987. [11]

[100] J.-L. Starck, J. Fadili, and F. Murtagh. The undecimated Wavelet decomposition and its reconstruction. *IEEE Trans. on Image Processing*, 16(2):297–309, 2007. [91]

[101] Online Supplement. S3:A spectral and spatial measure of local perceived sharpness in natural images. http://vision.okstate.edu/S3. [xviii, 97, 102]

[102] H. Tang, N. Joshi, and A. Kapoor. Learning a blind measure of perceptual image quality. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 305–312, 2011. [25, 127]

[103] P. C. Teo and D. J. Heeger. A perceptually tuned sub-band image coder with image dependent quantization and post-quantization data compression. In *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pages 1945–1948, 1989. [18]

[104] P. C. Teo and D. J. Heeger. Perceptual image distortion. In *Proc. SPIE*, volume 2179, pages 127–141, 1994. [18]

[105] H. Tong, M. Li, H. Zhang, and C. Zhang. Blur detection for digital images using Wavelet transform. In *IEEE Int. Conference on Mutltimedia and Expo*, pages 17–20, 2004. [23]

[106] E. Tsomko and H. J. Kim. Efficient method of detecting globally blurry or sharp images. In *Int. Workshop on Image analysis for Multimedia Interactive Services*, pages 171–174, 2008. [22]

[107] Video Quality Expert Group (VQEG). Final report from the video quality experts group on the validation of objectivemodels of video quality assessment II [Online]. Available: http://www.vqeg.org/ 2003. [5, 9]

[108] C. T. Vu, T. D. Phan, and D. Chandler. S_3: A spectral and spatial measure of local perceived sharpness in natural images. *IEEE Trans. on Image Processing*, 21(3):934–945, 2012. [21, 23, 100, 105, 107]

[109] P. V. Vu and D. M. Chandler. A fast Wavelet-based algorithm for global and local image sharpness estimation. *IEEE Signal Processing Letters*, 19(7). [23]

[110] Z. Wang. SSIM index for image quality assessment [Online]., 2003. Available: http://www.ece.uwaterloo.ca/ z70wang/research/ssim/. [29, 69]

[111] Z. Wang and A. C. Bovik. *Modern Image Quality Assessment.* Morgan and Claypool Publishers, 2006. [2, 5, 6, 10, 11, 16, 18]

[112] Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, 2009. [15]

[113] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612, 2004. [18, 29, 100, 105, 107, 127]

[114] Z. Wang and Q. Li. Information content weighting for perceptual image quality assessment. *IEEE Trans. on Image Processing*, 20(5):1185–1198, 2011. [15, 16, 19, 48, 124]

[115] Z. Wang, L. Lu, and A. C. Bovik. Foveation scalable video coding with automatic fixation selection. *IEEE Trans. on Image Processing*, 12(2):243–254, 2003. [10]

[116] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multi-scale structural similarity for image quality assessment. In *IEEE Asilomar Conference on Signals, Systems, and Computers*, pages 1398–1402, 2003. [18, 119]

[117] A. B. Watson. DCTune: A technique for visual optimization of dct quantization matrices for individual images. In *Society for Information Display Digest of Technical Papers*, volume XXIV, pages 946–949, 1993. [18]

[118] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor. Visibility of Wavelet quantization noise. *IEEE Trans. on Image Processing*, 6(8):1164–1175, 1997. [18, 120]

[119] D. Wei, J. Tian, Jr. R. O. Wells, and C. S. Burrus. A new class of biorthogonal Wavelet systems for image transform coding. *IEEE Trans. on Image Processing*, 7(7):1000–1013, 1998. [119]

[120] S. Winkler. *Digital Video Quality: Vision Models and Metrics*. John Wiley and Sons, 2005. [4]

[121] J. Wu, W. Lin, G. Shi, and A. Liu. Perceptual quality metric with internal generative mechanism. *IEEE Trans. on Image Processing*, 22(1):43–54, 2013. [20]

[122] W. Xue, L. Zhang, and X. Mou. Learning without human scores for blind image quality assessment. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 995–1002, 2013. [26, 124, 127]

[123] K. Yang and H. Jiang. Optimized-SSIM based quantization in optical remote sensing image compression. In *Int. Conference on Image and Graphics*, pages 117–122, 2011. [10]

[124] P. Ye and D. Doermann. No-reference image quality assessment using visual codebooks. *IEEE Trans. on Image Processing*, 21(7):3129–3138, 2012. [26, 127]

[125] P. Ye, J. Kumar, L. Kang, and D. Doermann. Unsupervised feature learning framework for no-reference image quality assessment. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1098–1105, 2012. [26]

[126] H. Yeganeh, M. Rostami, and Z. Wang. Objective quality assessment for image super-resolution: A natural scene statistics approach. In *IEEE Int. Conference on Image Processing*, pages 1481–1484, 2012. [10]

[127] K. Zeng and Z. Wang. Perceptual evaluation of image denoising algorithms. In *IEEE Asilomar Conference on Signals, Systems, and Computers*, pages 1351–1355, 2013. [10]

[128] L. Zhang and H. Li. SR-SIM: A fast and high performance IQA index based on spectral residual. In *IEEE Int. Conference on Image Processing*, pages 1473–1476, 2012. [12, 20, 32, 51, 52, 54, 64]

[129] L. Zhang, L. Zhang, X. Mou, and D. Zhang. FSIM: A feature similarity index for image quality assessment. *IEEE Trans. on Image Processing*, 20(8):1–26, 2011. [12, 15, 20, 38, 40]

[130] N. Zhang, A. Vladar, M. Postek, and B. Larrabee. A kurtosis-based statistical measure for two-dimensional processes and its applocation to image sharpness. In *Section on Physical and Engineering Sciences of American Statistical Society*, pages 4730–4736, 2003. [23]

[131] J. Zhu and N. Wang. Image quality assessment by visual gradient similarity. *IEEE Trans. on Image Processing*, 21(3):919–932, 2012. [119]

[132] X. Zhu and P. Milanfar. A no-reference sharpness metric sensitive to blur and noise. In *IEEE Int. Workshop on Quality of Multmedia Experience*, pages 64–69, 2009. [23, 100, 105, 107]

[133] X. Zhu and P. Milanfar. Automatic parameter selection for denoising algorithms using a no-reference measure of image content. *IEEE Trans. on Image Processing*, 19(12):3116–3132, 2010. [100, 105, 107]

Appendix A : Copyright Permissions

IEEE Permission to Reprint

In reference to IEEE copyrighted material which is used with permission in this dissertation, the IEEE does not endorse any of University of Windsors products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www .ieee.org/publications standards/publications/rights/rights link.html to learn how to obtain a License from RightsLink.

**ELSEVIER LICENSE
TERMS AND CONDITIONS**

Feb 12, 2015

This is a License Agreement between Ashirbani Saha ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Ashirbani Saha
Customer address	University of Windsor
	Windsor, ON N9B3P4
License number	3566560154931
License date	Feb 12, 2015
Licensed content publisher	Elsevier
Licensed content publication	Signal Processing
Licensed content title	Perceptual image quality assessment using phase deviation sensitive energy features
Licensed content author	Ashirbani Saha,Q.M. Jonathan Wu
Licensed content date	November 2013
Licensed content volume number	93
Licensed content issue number	11
Number of pages	10
Start Page	3182
End Page	3191
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Title of your thesis/dissertation	Investigating potential combinations of visual features towards improvement of full-reference and no-reference image quality assessment
Expected completion date	Feb 2015

Estimated size (number of pages)	185
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment

terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above: Permission granted is limited to the author accepted manuscript version* of your paper.

*Accepted Author Manuscript (AAM) Definition: An accepted author manuscript (AAM) is the author's version of the manuscript of an article that has been accepted for publication and which may include any author-incorporated changes suggested through the processes of submission processing, peer review, and editor-author communications. AAMs do not include other publisher value-added contributions such as copy-editing, formatting, technical enhancements and (if relevant) pagination.

You are not allowed to download and post the published journal article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version. A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx. As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting.

18.**Posting to a repository:** Authors may post their AAM immediately to their employer's institutional repository for internal use only and may make their manuscript publically available after the journal-specific embargo period has ended.

Please also refer to Elsevier's Article Posting Policy for further information.

19. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only.. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

20. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

Elsevier Open Access Terms and Conditions

Elsevier publishes Open Access articles in both its Open Access journals and via its Open Access articles option in subscription journals.

Authors publishing in an Open Access journal or who choose to make their article Open Access in an Elsevier subscription journal select one of the following Creative Commons user licenses, which define how a reader may reuse their work: Creative Commons Attribution License (CC BY), Creative Commons Attribution – Non Commercial - ShareAlike (CC BY NC SA) and Creative Commons Attribution – Non Commercial – No Derivatives (CC BY NC ND)

Terms & Conditions applicable to all Elsevier Open Access articles:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation.

The author(s) must be appropriately credited.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: You may distribute and copy the article, create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), to

164

include in a collective work (such as an anthology), to text or data mine the article, including for commercial purposes without permission from Elsevier

CC BY NC SA: For non-commercial purposes you may distribute and copy the article, create extracts, abstracts and other revised versions, adaptations or derivative works of or from an article (such as a translation), to include in a collective work (such as an anthology), to text and data mine the article and license new adaptations or creations under identical terms without permission from Elsevier

CC BY NC ND: For non-commercial purposes you may distribute and copy the article and include it in a collective work (such as an anthology), provided you do not alter or modify the article, without permission from Elsevier

Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

·   Promotional purposes (advertising or marketing)

·   Commercial exploitation (e.g. a product for sale or loan)

·   Systematic distribution (for a fee or free of charge)

Please refer to Elsevier's Open Access Policy for further information.

21. **Other Conditions**:

v1.7

Vita Auctoris

NAME	: Ashirbani Saha
YEAR OF BIRTH	: 1983
PLACE OF BIRTH	: INDIA
EDUCATION	

2015 : **Doctor of Philosophy**
Electrical and Computer Engineering
University of Windsor, Windsor, Ontario, Canada

2010 : **Masters of Applied Science**
Electrical and Computer Engineering
University of Windsor, Windsor, Ontario, Canada

2006 : **Bachelors of Engineering**
Electronics and Telecommunications Engineering
Bengal Engineering and Science University, Howrah, India