

Optimum Framework Design for No Reference Image Quality Assessment based on Generic Statistical Modeling

Thesis submitted

In fulfilment of the requirement for the award of degree of

DOCTOR OF PHILOSOPHY

Submitted by

Jayashri Vitthalrao Bagade

Registration Number - 950903029

Under the Supervision of

Dr. Kulbir Singh

Professor

Electronics & Communication Engineering
Thapar Institute of
Engineering & Technology, Patiala

Dr. Yogesh Dandawate

Professor

Electronics & Telecommunication
Vishwakarma Institute of
Information Technology, Pune



THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

COMPUTER SCIENCE AND ENGINEERING DEPARTMENT
THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY, PATIALA - 147004
PUNJAB, INDIA
March 2019

Certificate

I hereby certify that the work which is being presented in this thesis entitled **"Optimum Framework Design for No Reference Image Quality Assessment based on Generic Statistical Modeling"** for the award of degree of Doctor of Philosophy submitted to Computer Science and Engineering Department of Thapar Institute of Engineering & Technology, Patiala is an authentic record of my own work carried under the supervision of Dr. Kulbir Singh and Dr. Yogesh Dandawate. The thesis refers the work of other researchers which are duly listed in the reference section.

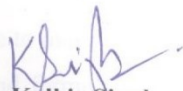
The matter presented in this thesis has not been submitted for the award of any other degree of this or any other university.



Jayashri Vitthalrao Bagade

Reg. No. 950903029

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.



Dr. Kulbir Singh

Professor

Electronics & Communication Engineering,
Thapar Institute of
Engineering & Technology, Patiala



Dr. Yogesh Dandawate

Professor

Electronics & Telecommunication,
Vishwakarma Institute of
Information Technology, Pune

Acknowledgement

Without motivation, encouragement, and right guidance it would not be possible for me to carry out this research work successfully. I take this opportunity to acknowledge and thank all those who directly or indirectly supported me and contributed in getting this research work completed.

First and foremost, I would like to extend my sincere gratitude to my mentors, Dr. Yogesh H. Dandawate and Dr. Kulbir Singh, who always encouraged me throughout the research work.

I would like to thank my supervisor Dr. Yogesh H. Dandawate for his exemplary support, advice, and mentoring. He was deeply involved, and contributed immensely throughout this research work. He patiently addressed my questions and kept me motivated even during the tough times in my pursuit of the Ph.D. I must appreciate the time and energy he invested in scrupulously reviewing my research papers and sharing his wisdom to make the content more accurate and comprehensive. He was always there to guide me throughout this research and writing this thesis. I express my sincere gratitude towards him.

Similarly, my profound gratitude to my supervisor, Dr. Kulbir Singh, for having faith in my work and supporting me throughout this research work with his knowledge whilst allowing me the room to work in my own way. I am overwhelmed and deeply obliged by his readiness and willingness to help. His exceptional guidance immensely helped me to put my efforts in right direction for getting the research work published. He always ensured that I am not only adequately involved in my research but also ensure constructive and unique outcomes.

I would like to make a special mention of Dr. A. S. Abhyankar, who helped me in initiation and formulation of the problem statement that became the base of the research. He always supported me with his impeccable mastery in this subject and his vision.

I also want to thank Dr. S. N. Londhe for encouraging me to register and pursue my doctoral study in this esteem institute.

Conducting regular progress presentations helped me follow the research plan and complete the work in given timeline. My sincere thanks to Dr. Manindar Singh, Head of Department, Dr. Rajesh Khanna, Dr. A. K. Verma, Dr. Inderveer Channa, and Dr. Bhatia for their valuable suggestions during these presentations. I am really fortunate and honoured to have such personalities as supervisors during the doctoral program.

I would like to thank Dr. P. Gopalan, Director, Thapar Institute of Engineering and Technology for his motivation and support throughout this work.

I express my special gratitude towards Director Dr. Mrs. B. S. Karkare, the management team of my institute Vishwakarma Institute of Information Technology, Prof. N. P. Pathak, Head of Depart of Information Technology, VIIT, and my colleagues, for extending their support whenever required.

Last but not the least, I thank to my family and friends for being always supportive and caring towards me.

I have been blessed with a loving son, Chetas, who might have missed me sometimes during his transformation into a young boy from a kid. I felt sorry for him. He has shown great understanding and never ever shown any childhood insistence during my doctoral study. I truly appreciate and thank him. A special acknowledgement goes to my husband, Devendra who has always been supportive in every way which helps me to focus on my goal. Without his encouragement this work would not have been a reality.

Jayashri Vitthalrao Bagade

Abstract

Since its inception, a digital image undergoes various image processing tasks and applications where different distortions are introduced in an image. These distortions are required to be quantified for quality estimation. In the real-world applications, image quality has to be estimated in the absence of the pristine image. No Reference Image Quality Assessment (NRIQA) methods play a vital role in such situations.

The digital images may have suffered from different distortions. No single image feature can model the presence of the distortions in an image. Therefore, a generic approach of combining parametric and other level-two features is proposed. This hypothesis is tested with the combination of existing block based, second-order statistical, and natural scene statistics based features. In the subsequent experimentations textural, block based, and shape adaptive wavelet transform features are implemented. The experiment further extended to derive a feature vector through a fusion of scale invariant feature transform key points and curvelet coefficients. Spatial covariance is also implemented as a feature.

This study proposes a framework for NRIQA to estimate the overall image quality. In order to implement the framework that can address different distortions, it is trained on the feature vector created by combining the above features. A combined feature vector grows in size with every appended feature and may raise the dimensionality curse problem. Therefore, to better the efficiency, the combined feature vector is optimized with principal component analysis.

The relationship between extracted features, subjective score, and the predicted score is complex and non-linear. Therefore, a machine learning based approach is proposed for designing the framework. SVM and ANN are employed as the quality prediction tools in the initial experimentations. It has been observed that the artificial neural network remains undertrained when it is fed with the feature values with lesser variations, and the feature magnitudes are not proportional to the degradation. Therefore, the neuro-wavelet model is developed. Neurofuzzy classifier is also implemented. The performance of the metrics are evaluated through LCC and SROCC. NFC and Neuro-wavelet model shows better accuracy for optimized combined feature metric. The proposed framework is developed to work with gray images and tested on LIVE, TID2008, and TID2013 datasets.

List of Publications

Published:

[P1] Jayashri V. Bagade, Kulbir Singh, Y. H. Dandawate, “No-Reference Image Quality Assessment based on Distortion Specific and Natural Scene Statistics based Parameters: A Hybrid Approach”, Malaysian Journal of Computer Science, vol. 32, pp. 31-46, 2019 (**SCI Indexed**).

[P2] Jayashri Bagade, Kulbir Singh, Yogesh Dandawate, “Generalized reference-less image quality assessment framework using texture energy measures and pattern strength features”, Turkish Journal of Electrical Engineering & Computer Sciences, vol. 25, pp. 3947-3959, 2017 (**SCI Indexed**).

[P3] Jayashri V. Bagade, Kulbir Singh, Yogesh H. Dandawate, "No reference image quality assessment using block-based and frequency domain statistical features: A machine learning approach", International Journal of Communication Networks and Distributed Systems, vol.12, no. 1, pp. 95-112, 2014.

[P4] J. V. Bagade, Y. H. Dandawate, Kulbir Singh, "No reference image quality assessment using block based features and artificial neural network", Global Trends in Information Systems and Software Applications, Springer Berlin Heidelberg Publisher, 2012, pp. 128-138

Communicated:

[P1] Jayashri V. Bagade, Kulbir Singh, Y. H. Dandawate, “No-Reference Image Quality Assessment using Neuro-Wavelet Model”, communicated to The Imaging Science Journal. (**SCI Indexed**).

[P2] Jayashri Bagade, Kulbir Singh, Yogesh Dandawate, “No-Reference Image Quality Assessment with Block based and Shape Adaptive Discrete Wavelet Features using Neuro-Wavelet Model” communicated to Sadhana Academy Proceedings in Engineering Sciences. (**SCI Indexed**).

[P3] Jayashri V. Bagade, Kulbir Singh, Y. H. Dandawate, “No-Reference Image Quality Assessment using Fusion Metric”, communicated to Multimedia Tools and Applications. (**SCI Indexed**).

List of Abbreviations

AD	Average Difference
ANN	Artificial Neural Network
BIQA	Blind Image Quality Assessment
BIQI	Blind Image Quality Index
BLIINDS	BLind Image Integrity Notator using DCT Statistics
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator
CBPN	Circular Back Propagation Network
CFM	Convolution Feature Map
CNN	Convolutional Neural Network
CPBD	Cumulative Probability Blur Detection
CQ	Correlation Quality
CRRM	Computational Ringing Region Map
CWT	Continuous Wavelet Transform
DCT	Discrete Cosine Transform
DESIQUE	DERivative Statistics-based Image QUality Evaluator
DFT	Discrete Fourier Transform
DIIVINE	Distortion Identification-based Image Verity and INtegrity Evaluation
DMOS	Differential Mean Opinion Score
DNT	Divisive Normalization Transform
DoG	Difference of Gaussian
DWT	Discrete Wavelet Transform
FF	Fast Fading
FR	Full Reference
FRIQA	Full Reference Image Quality Assessment
FSIM	Feature Similarity Index Metric
GAP-RBF	Growing And Pruning Radial Basis Function
Gblur	Gaussian blur
GGD	Generalized Gaussian Distribution
GLBP	Generalized Local Binary Pattern

GLCM	Gray Level Co-occurrence Matrix
GMF	Gaussian Membership Function
GRNN	General Regression Neural Network
HDTV	High Definition TeleVision
HVS	Human Visual System
ICA	Independent Component Analysis
IF	Image Fidelity
IG	Image Gradient
IPTV	Internet Protocol TeleVision
IQA	Image Quality Assessment
ITU	International Telecommunication Union
JNB	Just Noticeable Blur
JPEG	Joint Photographic Experts Group
KLD	Kullback Leibler Divergence
LBP	Local Binary Pattern
LCC	Linear Correlation Coefficient
LIVE	Laboratory of Image and Video Engineering
LMSE	Laplacian Mean Square Error
LoG	Laplacian of Gaussian
LSB	Least Significant Bit
LTP	Local Ternary Pattern
MD	Maximum Difference
MDID	Multiply Distorted Image Database
MF	Membership Function
ML	Maximum Likelihood
MLBP	Multiscale Local Binary Pattern
MLPBPNN	Multi-Layer Perceptron Back Propagation Network
MOS	Mean Opinion Score
MSE	Mean Square Error
MSSIM	Mean Structural Similarity Index Metric
MVG	Multi Variate Gaussian
NAE	Normal Absolute Error

NBM	Noticeable Blockiness Map
NFC	NeuroFuzzy Classifier
NK	Normal Cross Correlation
NR	No Reference
NRIQA	No Reference Image Quality Assessment
NSS	Natural Scene Statistics
NWT	Neuro-Wavelet Technique
PCA	Principal Component Analysis
PDF	Probability Distribution Function
PEM	Perceptual Edge Map
PLCC	Pearson's Linear Correlation Coefficient
PSNR	Peak Signal to Noise Ratio
QoE	Quality of Experience
QoS	Quality of Service
RDCT	Recognized Discrete Cosine Transform
RMSE	Root Mean Square Error
RR	Reduced Reference
RRIQA	Reduced Reference Image Quality Assessment
SADWT	Shape Adaptive Discrete Wavelet Transform
SAM	Spectral Activity Measure
SC	Structural Content
SD	Standard Deviation
SFM	Spatial Frequency Measure
SIFT	Scale Invariant Feature Transform
SROCC	Spearman Rank Order Correlation Coefficient
SS	Standard Score
SSIM	Structural Similarity Index metric
SVD	Singular Value Decomposition
SVM	Support Vector Machine
SVR	Support Vector Regression
TID	Tampere Image Database
UQI	Universal Quality Index

USFFT	Unequally Spaced Fast Fourier Transform
VQEG	Video Quality Experts Group
WN	White Noise
ZC	Zero Crossing rate

List of Figures

Figure 1.1	: Categorization of image quality assessment methods	3
Figure 1.2	: Thesis Contribution	10
Figure 3.1	: Schematic representation of the proposed framework	39
Figure 3.2	: Sample images from the created dataset.....	41
Figure 3.3(a)	: Sample images from LIVE dataset.....	43
Figure 3.3(b)	: Sample images from TID2008 dataset	44
Figure 3.4	: The basic architecture of ANN	47
Figure 4.1	: Examples of higher objective score computed by a mathematical model for distorted images than the subjective score of pristine images.....	53
Figure 4.2	: Schematic representation of methodology for existing block based feature metric [43] using ANN	54
Figure 4.3	: Regression plot of ANN using block based features [43]. MOS (target) are plotted against predicted scores (output)	55
Figure 4.4	: Schematic representation of methodology for a combination of block based features [43] and statistical features using ANN	56
Figure 4.5	: Regression plot of ANN using a combination of block based [43] and statistical features. MOS (target) are plotted against predicted scores (output)	57
Figure 4.6	: Schematic representation of methodology for combination of the block based [43], statistical [76], and DCT based NSS feature [103] using ANN and SVM.....	60
Figure 4.7	: Regression plots of ANN for training and testing datasets using combined feature metric	63
Figure 4.8	: Loss of texture details in “Tiger.bmp”. (a)Pristine image, (b)compressed with quality factor 50, (c)quality factor 10, and (d)50% blur variants.	64
Figure 4.9	: Sample image captured by phone camera	65
Figure 4.10	: Mask Convolution, $Pixels = N = L \times L$, $y_i = \sum_{i=1}^N x_i \times m_i$	67
Figure 4.11	: Statistic Evaluation, $Pixels = N = C \times C$, $z_i = f(y_1, y_2, \dots, y_n)$...	68
Figure 4.12	: Reference image “bikes.bmp” from LIVE dataset and its different distortion variants - (a)Original, (b)JPEG, (c)JPEG2000, (d)Gblur, (e)WN, and (f) FF with their respective subjective scores	68

Figure 4.13	: Plots showing variations for (a)mean and (b)standard deviation of nine energy maps for “bikes.bmp” and its distortion variants from LIVE dataset numbered as 1-original, 2-JPEG, 3-JPEG2000, 4- Gblur, 5-WN, and 6-FF	69
Figure 4.14	: Plots showing variations for (a)SD, (b)skewness, and (c)kurtosis of LBP for “bikes.bmp” and its distortion variants from LIVE dataset numbered as 1-original, 2-JPEG, 3-JPEG2000, 4-Gblur, 5-WN, and 6-FF	72
Figure 4.15	: Schematic representation of methodology for textural feature metric	73
Figure 4.16	: Regression plots of ANN for training and testing datasets using textural feature metric.....	76
Figure 5.1	: DMOS, IG, SFM, and SAM for reference image “Plane.bmp” and its compressed variants with different bit rates	81
Figure 5.2	: NeuroFuzzy Technique.....	82
Figure 5.3	: A typical Gaussian membership function.....	83
Figure 5.4	: Schematic representation of methodology for block based feature metric	85
Figure 5.5	: Regression Plots of ANN for training and testing datasets using block based features for JPEG images	87
Figure 5.6	: Reference image “Monarch.bmp” from LIVE dataset with different distortions (a)Original, (b)JPEG (c)JPEG2000, (d)Gblur, (e)WN, and (f)FF with their respective subjective scores	90
Figure 5.7	: Plot showing variations in mean, SD, SVD, and shape features computed for SADWT coefficients of “Monrach.bmp” and its distorted variations	91
Figure 5.8	: Regression plot for the undertrained ANN for LIVE dataset.....	91
Figure 5.9	: Algorithm of neuro-wavelet model	93
Figure 5.10	: Multilevel decomposition by wavelet.....	94
Figure 5.11	: Schematic representation of methodology for combination of SADWT and DWT features	95
Figure 5.12	: Regression Plots of NWT classifier using a combination of SADWT and DWT features	97
Figure 5.13	: Regression Plots of NWT classifier using block based features for JPEG images.....	98
Figure 6.1	: (a) SIFT feature point map and (b) curvelet coefficients map for “Monarch.bmp” from LIVE dataset	102

Figure 6.2	: Flowchart of the fusion of SIFT key points and Curvelet coefficients.....	105
Figure 6.3	: Histograms of feature values (a) Original image “bikes.bmp”, (b) Gaussian blur, and (c) JPEG2000 from LIVE dataset	106
Figure 6.4	: Comparison of the feature values calculated using spatial covariance for original image “bikes.bmp” and its variants from LIVE dataset.....	108
Figure 6.5	: Schematic representation of methodology for performance evaluation of a combination of fusion feature and spatial covariance feature.....	109
Figure 6.6	: Regression plots of NWT using fusion feature metric	111
Figure 6.7	: Regression plots of NWT using spatial covariance metric.....	113
Figure 6.8	: Regression plots of NWT using a combination of fused feature metric and spatial covariance metric	115
Figure 7.1	: Schematic representation of the proposed framework	121
Figure 7.2	: Evolution of the proposed framework.....	122
Figure 7.3(a)	: Regression plots of NWT using combined feature metric_(a)before and (b)after dimension reduction by PCA for LIVE dataset. DMOS (target) are plotted against predicted scores (output).....	125
Figure 7.3(b)	: Regression plots of NWT using combined feature metric_(a)before and (b)after dimension reduction by PCA for TID2008 dataset. MOS (target) are plotted against predicted scores (output).....	125
Figure 7.3(c)	: Regression plots of NWT using combined feature metric. (a)before and (b)after dimension reduction by PCA for TID2013 dataset. MOS (target) are plotted against predicted scores (output).....	125
Figure 7.4	: Regression plots of NWT using existing combined statistical features (given in Appendix-I)	132

List of Tables

Table 2.1	: Summary of the survey of the literature on the development in distortion specific approach	21
Table 2.2	: Summary of the survey of the literature on the development in feature extraction and machine learning approach.....	26
Table 2.3	: Summary of the survey of the literature on the development in NSS based approach.....	32
Table 3.1	: Summarised characteristics of image datasets LIVE, TID2008 and TID2013.	45
Table 3.2	: Preparation of Training and testing datasets for LIVE, TID2008, and TID2013	46
Table 4.1	: Performance evaluation of block based feature metric [43]	54
Table 4.2	: Performance evaluation of combination of block based [43] and statistical features	57
Table 4.3	: DMOS and MOS values as SVM classes	61
Table 4.4(a)	: Performance evaluation of combined feature metric using ANN and SVM on LIVE dataset	61
Table 4.4(b)	: Performance evaluation of combined feature metric using ANN and SVM on TID2008 dataset.....	62
Table 4.4(c)	: Performance evaluation of combined feature metric using ANN and SVM on TID2013 dataset.....	62
Table 4.5	: Comparison of the proposed metric with state-of-the-art metrics for LIVE dataset.....	63
Table 4.6(a)	: Performance evaluation of textural feature metric using ANN and SVM on LIVE dataset	74
Table 4.6(b)	: Performance evaluation of textural feature metric using ANN and SVM on TID2008 dataset.....	75
Table 4.6(c)	: Performance evaluation of textural feature metric using ANN and SVM on TID2013 dataset.....	75
Table 4.7	: Performance comparison between proposed metric and state-of-the-art NR metrics on LIVE dataset.....	75
Table 5.1	: Preparation of NFC clusters based on the subjective scores.....	86
Table 5.2	: Performance evaluation of block based features using ANN and NFC through LCC and SROCC for JPEG images from LIVE, TID2008, and TID2013 datasets	86

Table 5.3	: Performance comparison of the proposed metric with state-of-the-art metrics on LIVE dataset	87
Table 5.4(a)	: Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on LIVE dataset.....	96
Table 5.4(b)	: Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on TID2008 dataset	96
Table 5.4(c)	: Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on TID2013 dataset	96
Table 5.5	: Performance Comparison of the proposed combined feature metric with state-of-the-art metrics on LIVE dataset.....	97
Table 5.6	: Performance evaluation of block based features using NWT through LCC and SROCC for JPEG images from LIVE, TID2008, and TID2013 datasets.....	98
Table 6.1(a)	: Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on LIVE dataset	110
Table 6.1(b)	: Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on TID2008 dataset.....	110
Table 6.1(c)	: Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on TID2013 dataset.....	110
Table 6.2(a)	: Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on LIVE dataset.....	112
Table 6.2(b)	: Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2008 dataset.....	112
Table 6.2(c)	: Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2013 dataset.....	112
Table 6.3(a)	: Performance Evaluation of combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on LIVE dataset.....	113
Table 6.3(b)	: Performance Evaluation of the combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2008 dataset	114
Table 6.3(c)	: Performance Evaluation of the combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2013 dataset	114

Table 6.4	: Performance comparison of all three proposed metrics with state-of-the-art metrics through LCC and SROCC on LIVE, TID2008, and TID2013 datasets.....	115
Table 7.1(a)	: Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on LIVE dataset.....	124
Table 7.1(b)	: Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on TID2008 dataset.....	124
Table 7.1(c)	: Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on TID2013 dataset.....	124
Table 7.2(a)	: Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on LIVE dataset.....	127
Table 7.2(b)	: Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on TID2008 dataset.....	128
Table 7.2(c)	: Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on TID2013 dataset.....	129
Table 7.3	: Performance comparison of the combined feature metrics formed at different stages of evolution of the framework using NWT through LCC on LIVE, TID2008, and TID2013 datasets.....	130
Table 7.4	: Performance comparison of the existing combined statistical feature metrics (Appendix-I) using NWT through LCC on LIVE, TID2008, and TID2013 datasets.....	131
Table 7.5	: Performance comparison of the proposed framework using NFC and NWT with state-of-the-art frameworks through LCC and SROCC on LIVE, TID2008, and TID2013 datasets.....	133

Table of Contents

Abstract	v
List of Publications	vi
List of Abbreviations.....	vii
List of Figures	xi
List of Tables.....	xiv
 Chapter 1. Introduction	 1
1.1 Categorization of Image Quality Assessment	2
1.1.1 Subjective Assessment.....	3
1.1.2 Objective Assessment	4
1.1.2.1 Full Reference Image Quality Assessment.....	4
1.1.2.2 Reduced Reference Image Quality Assessment	5
1.1.2.3 No Reference Image Quality Assessment	6
1.2 Contribution.....	8
1.3 Organization of The Thesis	9
 Chapter 2. Literature Survey	 11
2.1 A Brief Discussion on Literature Survey of FRIQA and RRIQA	11
2.2 A Survey of No Reference Image Quality Assessment.....	12
2.2.1 Distortion Specific NRIQA.....	13
2.2.2 Feature Extraction and Machine Learning based NRIQA	22
2.2.3 Natural Scene Statistics based NRIQA.....	27
2.3 Research Gaps	33
 Chapter 3. Methodology.....	 36
3.1 Motivation	37
3.2 Problem Definition	38
3.3 Research Methodology	39
3.4 Image Datasets used for The Experimentation.....	40
3.4.1 Image Dataset Created and Used for Analysis of Existing NR Metrics.....	40
3.4.2 Laboratory of Image and Video Engineering (LIVE).....	41
3.4.3 Tampere Image Database 2008 (TID2008).....	43

3.4.4	Tampere Image Database 2013 (TID2013).....	44
3.4.5	Summarized Comparison of LIVE, TID2008, and TID2013 ...	45
3.4.6	Preparation of Training and Testing Datasets from LIVE, TID2008, and TID2013 Image Datasets	45
3.5	Performance Evaluation	46
3.6	Study of Artificial Neural Network Classifier.....	47
Chapter 4.	No Reference Image Quality Assessment using Textural Features	50
4.1	Analysis of Existing NR Metrics.....	51
4.1.1	Study of The Existing Block based NR Metric.....	51
4.1.1.1	Block based features	52
4.1.1.2	Block based and Statistical Features	55
4.1.1.3	Combination of Block based, Statistical, and NSS based Features.....	58
4.2	Design of Textural Feature Metric	64
4.2.1	Texture Analysis	65
4.2.2	Feature Extraction	66
4.2.2.1	Laws' Energy Measures	66
4.2.2.2	Pattern Strength	70
4.2.3	Methodology and Results.....	73
4.3	Summary.....	77
Chapter 5.	Design of NRIQA Metrics using Block based and Shape Adaptive Discrete Wavelet Features	78
5.1	Design of Block based Feature Metric	79
5.1.1	Neurofuzzy Classifier.....	81
5.1.2	Methodology and Results.....	85
5.2	Shape Adaptive Discrete Wavelet Transform Features.....	88
5.2.1	Design of Shape Adaptive Discrete Wavelet Transform Feature.....	88
5.2.2	Neuro-Wavelet Technique	92
5.2.3	Methodology and Results.....	94
5.3	Summary.....	99
Chapter 6.	Design of NRIQA Metric using Fusion of Curvelet and SIFT Features.....	100
6.1	Feature Extraction.....	101
6.1.1	Fusion Feature Vector	101

6.1.1.1	Scale Invariant Feature Transform (SIFT)	102
6.1.1.2	Curvelet Transform.....	104
6.1.1.3	Implementation of A Fusion Feature Metric	104
6.1.2	Spatial Covariance	106
6.2	Methodology and Results	108
6.2.1	Performance Evaluation of Fusion Metric	109
6.2.2	Performance Evaluation of Spatial Covariance Metric.....	111
6.2.3	Performance Evaluation of Combined Feature Metric	113
6.3	Summary.....	116
Chapter 7.	Optimum Framework for NRIQA	117
7.1	Design of The Proposed Framework for NRIQA.....	118
7.1.1	Principal Component Analysis (PCA)	119
7.1.2	Methodology	120
7.1.3	Performance Evaluation and Results	123
7.2	Summary.....	133
Chapter 8.	Conclusion and Future Scope	134
8.1	Summary.....	134
8.2	Conclusion.....	137
8.3	Future Scope	137
References		139
Appendix I		152
Appendix II		155

Chapter 1.

Introduction

Digital technology and Electronics have seen enormous advancements in last few decades. Digital revolution, specifically the internet, marked the beginning of a new phase in this information era. This development paved the way for the new inventions and large-scale manufacturing of digital electronic devices. High availability, affordability, and usability of digital gadgets like smartphones and digital imaging devices have accelerated the use of these gadgets among the people. With the widespread use of high-speed internet and increased use of various social networking websites such as Facebook, Instagram, and Youtube, the volume of data being shared and transmitted over the internet has been increasing by leaps and bounds. Inventions of Internet Protocol TeleVision (IPTV), internet protocol telephony, High Definition TeleVision (HDTV), and ever-increasing demand of high-resolution contents such as HD and 4K has increased the requirement of image processing and transmission applications. Bandwidth and storage availability is still the major hurdle faced by the industry to meet the ever-increasing expectations of promised visual experience. This challenge has been addressed to some extent by applying various compression techniques to multimedia data. Compression algorithms are developed to reduce file size which results in a reduction of storage space requirement, communication time, and transmission bandwidth requirement. Although the compression helps to achieve better efficiency and minimizes the resource requirements, it introduces blocking and ringing artifacts. Transmission also adds channel noise which adversely affects the image quality.

Blocking and ringing artifacts introduced in an image by Joint Photographic Experts Group (JPEG) compression is an example of such compression induced distortion. JPEG compression technique uses block based transform coding. JPEG coder divides an image into blocks of $n \times n$ size which are then transformed and quantized. The abrupt change across the block boundaries causes the introduction of blocking artifact in the image and results into the image quality degradation [1].

In JPEG and JPEG2000 compression techniques, high-frequency components are attenuated due to the quantization process. This process of attenuation causes the loss of sharp details and introduces the blurring effect in an image.

The relative motion of an object and camera may cause a blur to the captured images during the image acquisition process. In image processing applications, use of smoothing filters also adds blur in the resulted images.

Images are also affected by the noise introduced during analogue to the digital conversion process and bit errors introduced in image data during the transmission.

Quality of a digital image can thus be affected by various kinds of degradations during processes such as image acquisition, processing, restoration, compression, and transmission. In today's consumer-centric multimedia industry, Quality of Service (QoS) has the utmost importance, and Quality of Experience (QoE) offered to the end user is one of the essential means of assessing it. It is essential to quantify these degradations to improve visual experience [2]. For improved QoE, image processing applications are required to identify such degradations without human intervention. Reliable Image Quality Assessment (IQA) algorithms thus play an important role in various image processing systems [2]. Real-time remote medical assistance, diagnosis, and treatment, security surveillance, various scientific expeditions are the few examples where image quality assessment is essential.

1.1 Categorization of Image Quality Assessment

IQA techniques attempt to measure the degradation in the image quality. Image quality estimation methods are broadly classified in two categories as- subjective IQA and objective IQA. In the subjective IQA, human subjects estimate the image quality whereas, objective IQA techniques use algorithms for estimating the perceived image quality. Based on the availability of an original image or its information, the objective IQA methods are the usually classified in three categories as- Full Reference Image Quality Assessment (FRIQA), Reduced Reference Image Quality Assessment (RRIQA), and No Reference Image Quality Assessment (NRIQA) as shown in figure 1.1 [3].

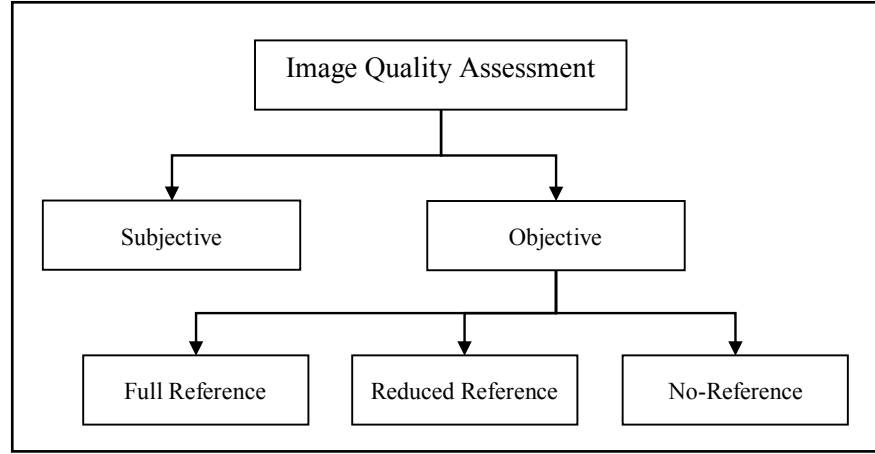


Figure 1.1: Categorization of image quality assessment methods

1.1.1 Subjective Assessment

The subjective evaluation method is valued as the most reliable means of assessing image quality since human observers are the end users of many image processing applications. In the subjective assessment method, usually, a large number of human observers are presented with a set of images for quality assessment. Observers are grouped into two groups, expert subjects and non-expert subjects [4, 5]. Observers perceive images, assess and rate the quality of images on the scale of 1 to 10 called as opinion score, where scale from 1 to 10 represents bad to the excellent score. Mean Opinion Score (MOS) is obtained by averaging these individual opinion scores.

Differential Mean Opinion Score (DMOS) is another method used for the subjective assessment. It is calculated as the difference between the score of the reference image and the score assigned to the test image by the subject. Higher the DMOS value, more is degradation to the image.

Subjective assessment methods are widely used and regarded as reliable, but they are expensive, labourious, and time-consuming [4, 5]. Also, subjective assessment methods cannot provide instantaneous quality assessment due to the involvement of a large number of subjects which limits its applicability in the real-time image processing applications. Image quality assessment methods developed using objective techniques provide a viable solution in such scenario.

1.1.2 Objective Assessment

The main aim of objective IQA is to build a computational model to assess image quality precisely without human intervention. Objective assessment of image quality requires matching with the human perception of quality which is not possible without an understanding of the Human Visual System (HVS). Many researchers have tried to develop objective image quality assessment metric that mimics the functioning of HVS and thus can take advantage of known characteristics of it. However, HVS is a very complex system and not fully understood yet.

Objective IQA methods are classified broadly as Full Reference (FR), Reduced Reference (RR), and No Reference (NR) methods which are discussed below.

1.1.2.1 Full Reference Image Quality Assessment

The Full Reference Image Quality Assessment (FRIQA) metrics determine the quality of a distorted image with reference to the original image available all the time which is considered to be of perfect quality or distortion-free [4, 5, 6].

FRIQA metrics developed initially were based upon various mathematically defined measures. Each of these metrics may have its advantages and shortcomings. Structural Content (SC) and Average Difference (AD) are little consistent. Correlation Quality (CQ), Normal Cross Correlation (NK), Laplacian Means Square Error (LMSE), and Maximum Difference (MD) are pixel-based measures and are consistent. Image Fidelity (IF) and Normal Absolute Error (NAE) also show good consistency [7, 8]. Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) metrics are used widely due to their simplicity. However, they are also criticized for not associating with the perceived quality.

Understanding the statistical models for the identification of natural images resembling the human brain activity is necessary to define HVS based and statistically based features [9, 10, 11]. Though statistically defined quality measures have low complexity, they do not always offer correct evaluation. Various experimentations conducted for image quality evaluation demonstrated the fact that these metrics may specify appropriate values for an image due to the statistical nature even though it is not perceptually good [4, 5]. This emphasizes the need for the metrics strongly associating with human perception.

Further development in FRIQA is focused on HVS based metrics. Contrast sensitive function, light adaption, contrast masking and foveated vision are the psychophysical features of HVS [4]. Changes in the structural information due to the distortions are assessed and quantified by HVS. Mean Structural Similarity Index Metric (MSSIM) tries to simulate this very similar function of image assessment and quantification. It is statistically defined in terms of luminance, contrast, and structural content [12]. The Universal Quality Index (UQI) is developed in 2004 as an extension to the Structural Similarity Index Metric (SSIM) [13].

In the majority of the applications and devices being used in recent days do not have access to the reference images. FR metrics cannot be used in such applications. Also, FRIQA metrics are neither feasible nor practical for the applications constrained by the bandwidth availability. Despite of these disadvantages of FRIQA metrics, their strong advantage is that they show excellent accuracy and hence are very useful in laboratory benchmarking of the image and video processing algorithms; such as different compression schemes.

1.1.2.2 Reduced Reference Image Quality Assessment

Reduced Reference Image Quality Assessment (RRIQA) methods use the partial information available about the reference image to evaluate image quality. A set of certain features of the pristine image extracted before the compression is either embedded in the image or transmitted separately as additional information [14]. A similar set of features is extracted from the possibly distorted image at the receiving end. These extracted features are then compared with the features of the reference image received along with the image.

Careful selection of appropriate features is essential for reliable quality estimation. These features should convey an efficient summary of the reference image and successfully reflect the various distortions introduced in the image. However, the amount of feature information transmitted is constrained by the available transmission data rate. So, this transmission data rate can affect the performance achievement of RRIQA algorithm as it depends on the quantity of feature data transmitted. Higher the data rate available more is the amount of transmitted features information and more accurate is the quality estimation [15, 16, 17].

The situations, where transmission bandwidth availability is limited, feature data cannot be transmitted along with the actual image, which is usually huge in size. Such situations formulate the design of RR methods similar to the NR methods which may have less accuracy due to the added difficulty level. RRIQA is considered as the better negotiation between FRIQA and No Reference Image Quality Assessment (NRIQA) as the algorithms are still being developed and evolved for effective and efficient IQA models.

1.1.2.3 No Reference Image Quality Assessment

NRIQA approach aims towards the development of a computational model that can quantitatively evaluate image quality accurately and automatically. NRIQA is a challenging problem in the area of digital image processing [2].

In many real-world applications, reference image or even partial information about the reference image could not be made available, and only distorted images or images with artifacts are available for quality evaluation. Few common examples of such applications are –

- Images processed and transmitted/broadcasted over the social media,
- Photos captured, processed, and stored in digital cameras,
- Remote printing,
- Synthetic or artificial 2D/3D images created using some photorealistic rendering software where no reference image ever exists [18], etc.

Although human being can easily assess the quality of an image just by a glance or observation without the need of a reference image, development of an objective NR image quality assessment metric is quite a complex task due to the limited perceptive available about HVS.

NR techniques developed so far follow three trends: Distortion specific approach, Feature extraction and machine learning approach, and Natural Scene Statistics (NSS) approach [19].

In the distortion specific approach, algorithms are designed to quantify particular presumed distortion in isolation of the other distortions. Examples of such NRIQA algorithms are blockiness evaluation, blur estimation and ringing effects measure.

Distortion specific NRIQA schemes deal with the specific distortion only and do not handle other types of distortions which is the major disadvantage of distortion specific approach.

In the machine learning approach, features describing the distortions are extracted from the images. These features are then used to train the machines through techniques like neural networks and Support Vector Machine (SVM) to evaluate the quality score. The effectiveness of machine learning approach depends on how much information the features convey. More appropriate the features extracted, more reliable is the quality assessment.

NSS approach considers a natural image as a small subset of a complete set of possible images. NSS methods calculate the distance between the distributions of the distorted image and the subset of natural images to evaluate the quality score. This approach tries to determine how the statistics of an image is affected by the distortions. NSS approach relies on extensive statistical modeling and reliable generalization of the model.

However, an image can have different distortions induced by different image processing applications. Such images cannot be assessed correctly by evaluating it for any particular distortion type. A framework generalized to predict the image quality independent of the distortion type and image processing application can have the potential to evaluate such images. This framework should be robust and must work in a plug and play environment.

Thus, the proposed research work aims toward the development of a framework that works with the optimized feature vector generated by combining the different parametric and other level-two features. Optimization of a feature vector can be achieved using Principal Component Analysis (PCA). The proposed framework uses fuzzy classifier or equivalent for the quality prediction. This framework is proposed for the quality evaluation of gray images for different types of distortions with the highest possible accuracy.

Deep-learning technique is being used recently for NRIQA. Instead of handcrafted features, the network is designed to extract the features along with the quality prediction. The framework based on this concept shows good performance. Still, this concept is in its infant stage and has challenges such as resource demanding and

time consuming. With the improvement in the algorithms addressing these challenges, real time systems can utilise these metrics. However, it is not the part of this study.

1.2 Contribution

Below are the main contributions of the work:

- Textural information is represented in the high-frequency domain. The processes such as image acquisition and compression introduce distortions which result into loss of textural details. Textural feature metric is designed to analyse the texture information using texture energy measures and pattern strength.
- JPEG images are widely used in various image processing applications. These images suffer from the blockiness. NR metric is designed using image gradient, spatial frequency and spectral activity measures to evaluate the blockiness.
- An image may consist of different objects. These objects are also subjected to the distortions. NRIQA metric is thus defined using shape-adaptive discrete wavelet transform for the quality evaluation of an image.
- NSS based NR metric is designed by combining Scale Invariant Feature Transform (SIFT) coefficients and curvelet coefficients. The intensities of SIFT feature points are changed due to the distortion. SIFT discards the points which lie in low contrast region and on the edges. Whereas, Curvelet transform is capable of identifying these points and hence curvelet coefficients are fused with SIFT feature points to define the metric.
- During the experimentation, it has been observed that the variations in feature magnitudes are lesser and not proportional to the amount of degradation. Therefore, Artificial Neural Network (ANN) remained untrained despite a sufficient number of training exemplars. The Neuro-Wavelet Technique (NWT) which improves the performance of the framework is thus employed. Further, to overcome the drawbacks of ANN, Neurofuzzy Classifier (NFC) is used for quality prediction.

- A generic framework capable of accommodating different features is designed. Adding the features increases the size of the feature set. Therefore, the framework is optimized using an optimization technique.

1.3 Organization of The Thesis

Majority of existing NRIQA metrics are focused on specific distortion. A framework for estimating the overall image quality irrespective of the distortion type is proposed in this thesis. Pictorial representation of the thesis contribution is shown in figure 1.2. The study is presented through eight chapters which are summarized below.

Chapter two presents the detailed survey of published literature in the area of NRIQA. This chapter begins with the overview of the literature focused around the distortion specific NRIQA approaches, followed by the feature extraction and machine learning based approaches, and NSS based NRIQA approaches at the end. The chapter is concluded with the research gaps identified from the literature survey.

Chapter three describes the methodology for the proposed work. The motivation behind this work and the problem statement is discussed in this chapter. The architecture and working of ANN classifier, the public image datasets used in the experimentations are also discussed.

Chapter four describes the determination of NRIQA metric using textural parameters. An NRIQA metric defined as the combination of existing block based, statistical and NSS based features is presented. Further, the necessity of texture analysis and design of textural metric is also explained. Finally, results obtained for this metric are depicted.

Chapter five illustrates the design of NRIQA metric that uses block based and shape adaptive wavelet features. Design of the neuro-wavelet model is also discussed in the chapter. Experimental results are exhibited to showcase the performance of these metrics.

Chapter six describes the articulation of NR metric using a fusion of SIFT key points and curvelet coefficients. Spatial covariance metric is also implemented and combined with the fusion metric.

Chapter seven describes the design of an optimum framework for NRIQA. It also explains the use of principal component analysis technique for optimization of the framework. Experimental results for the neuro-wavelet model and neuro-fuzzy classifiers are also presented.

Finally, chapter eight concludes the presented work and discusses the future scope and directions which may assist to improve the performance of NRIQA framework further.

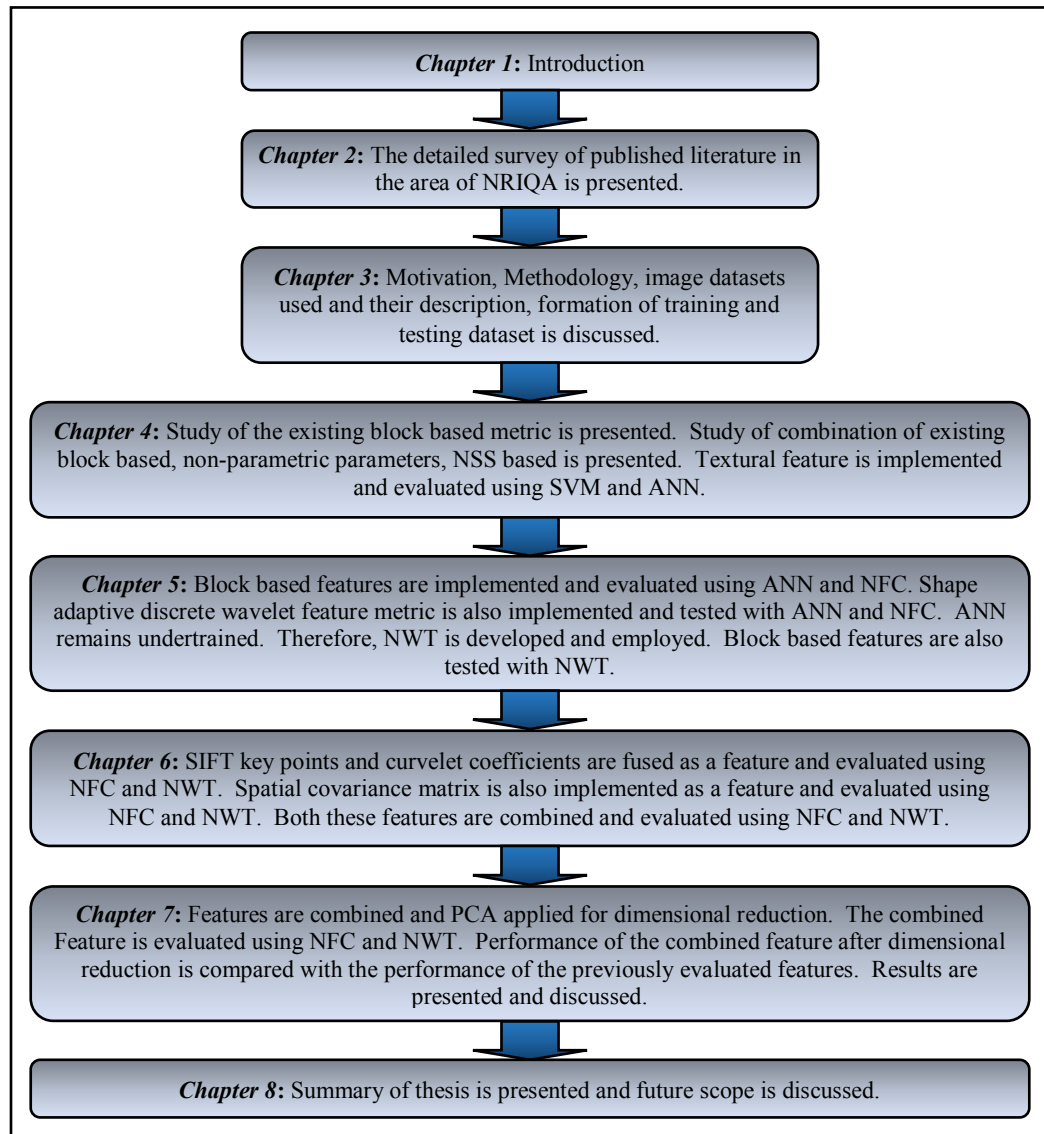


Figure 1.2: Thesis Contribution

Chapter 2.

Literature Survey

Ever evolving digital imaging applications provide great opportunities to the researchers to come up with various novel approaches for efficient evaluation of the image quality. The available literature is scrupulously studied to understand the research carried out so far and methodologies designed by the researchers in the area of image quality assessment. Before presenting the survey of the literature related to NRIQA, a brief discussion on literature surveyed on FRIQA and RRIQA is presented below.

2.1 A Brief Discussion on Literature Survey of FRIQA and RRIQA

The published literature illustrates that initial research was mainly focused on the development of FR metrics. HVS based MSSIM metric proposed by Z. Wang *et al.* [12, 13] came as a breakthrough in FR metric development. The proposed HVS based metric was designed to extract structural features from an image that changes with the distortions. Many up-gradations to MSSIM metric were followed up for the performance improvement [20, 21, 22, 23]. Researchers further extended the work to deal with non-structural degradations, such as geometric degradation. Few of the pioneer development need to be mentioned here are implementation of MSSIM in the wavelet domain by C. L. Yang *et al.* [24], FR metrics developed based on gradient similarity [25, 26], Feature Similarity Index Metric (FSIM) developed by L. Zhang *et al.* [27], Singular Value Decomposition (SVD) based FR metrics [28, 29, 30], etc. M. Pedersen *et al.* [31] contributed majorly with the most extensive evaluation carried out on FR metrics to know their merits and demerits. The authors used six different image databases including Laboratory of Image and Video Engineering (LIVE) and Tampere Image Database 2008 (TID2008).

A. Chetouania *et al.* [32] discussed the limitations of conventional metrics designed irrespective of the degradation type. Authors stated that the efficiency of these metrics depends on the particular degradation present in the image and the image

content. Authors proposed estimation of the degradation type present in the image before the quality estimation. A Chetouania *et al.* [33] proposed a novel approach towards the development of a framework in FRIQA. Authors worked with the existing FR metrics but allow the framework to select “most appropriate” IQA metric based on “degradation of classification scheme”.

FRIQA assume that reference image is available all the time for evaluation of a test image which is practically not possible in real-world applications. RRIQA models facilitated the quality evaluation in the absence of the reference images. RR models utilize partial information extracted from the reference image and transmitted besides the image as additional information. These features can reflect various types of degradations. R. Soundararaja *et al.* [15] implemented an RR metric that compares the entropies of the pristine image and its distorted variant for evaluating the changes in the image data. Divisive Normalization Transform (DNT) is employed to articulate different RR features by some researchers [16, 17, 34]. L. Ma *et al.* [35] developed a novel metric for RRIQA using Recognized Discrete Cosine Transform (RDCT). Researchers further put forth their efforts to develop RR metrics based on NSS [36, 37, 38, 39].

The performance of the RRIQA metrics depends on the quantity and quality of feature data, which demands additional transmission and computation resources. Sometimes the feature data can be huge. Therefore, RRIQA is not convenient in various practical situations. Also, with the growing use of the internet, social networking applications, digital imaging and printing devices, where availability of reference image or its partial information is nearly impossible; demand for effective and efficient NRIQA algorithms has begun increasing.

2.2 A Survey of No Reference Image Quality Assessment

NRIQA, also called as Blind Image Quality Assessment (BIQA), methods evaluate the image quality without any information about the reference or pristine image. However, developing the NRIQA model is not an easy task and requires novel ideas with successful implementations through complex algorithms [40, 41]. Apart from the efficiency, lesser time complexity, adaptability to different distortions; the NRIQA has the challenge to develop the metrics which are consistent with HVS. The HVS is trained since the childhood of an individual with high quality images.

It builds the models of high quality images and trains itself for quality estimation of distorted images by applying these models. NRIQA model is supposed to function similar to that of HVS in assessing the image quality. Researchers have taken varied approaches and proposed different NRIQA models that are consistent with HVS.

The literature survey shows that researchers have opted for different approaches for developing the NRIQA metrics. The survey of the literature is categorized into three sections based on these approaches as Distortion specific, Feature extraction and machine learning based, and NSS based NRIQA.

2.2.1 Distortion Specific NRIQA

Literature shows that most of the earlier NRIQA models are distortion specific, i.e., they presume that the distortion which affects the image is known [42]. These NRIQA models are focused on the assessment of specific types of distortion such as blocking, ringing, and blurring.

Image processing applications, transmission, certain lossy compression techniques, etc. can induce distortion in the image. The widely used JPEG and JPEG2000 are examples of such lossy compression techniques. JPEG compression technique uses block based Discrete Cosine Transform (DCT). The discontinuity at block boundaries occurs due to the block based quantization. The blocks are quantized independently which introduces blocking artifact in an image. Whereas images compressed using Discrete Wavelet Transform (DWT) based JPEG2000 compression scheme suffer from blurring and ringing artifacts. Undoubtedly, distortion specific NR methods are useful in assessing the image quality by quantifying specific types of distortions such as the blockiness.

In a pristine or original image, low frequency elements are allocated with maximum energy while high frequency elements possess a very small amount of energy. However, block boundaries in the image affected by the blockiness exhibit considerable energy at high frequency. Real edges in the image also exhibit a similar characteristic and may be erroneously read as blockiness by the NRIQA methods specifically developed for blockiness. Thus distortion specific NRIQA methods might lead to inaccurate assessment in certain conditions.

Z. Wang *et al.* [43] proposed an NRIQA metric which predicts the quality of JPEG images. The NRIQA method proposed by the authors work by extracting effective features from an image. Blockiness is estimated by averaging the difference determined across the boundaries of blocks of size 8×8 . Decreased signal activity indicates the introduction of blurriness. Thus, the average of the absolute difference between neighbouring pixels in the block is computed. Authors used the nonlinear parameterized mathematical model for predicting the quality. The proposed model is computationally efficient and requires less memory. Experimental results demonstrated good accuracy for two image datasets which were created for this study. However, the model gets overtrained when both the datasets are combined for the training. This might adversely affect the generality of the model.

An NRIQA metric proposed by S. Suthaharan [44] attempts to estimate the distortion using a combination of features addressing the blocking artifact and undistorted image edges. The proposed model follows multi-neural channel concept of HVS and uses Standard Score (SS) formula. Bit pattern of Least Significant Bits (LSB) is another feature extracted to determine the regions affected by JPEG compression. A normalized visually significant blocking artifact metric is computed. The parameterized mathematical model computes the quality score. Results illustrated that this approach works well with JPEG images from LIVE dataset.

S. Lee *et al.* [45] proposed an NRIQA metric that deals with the blocking artifact. Firstly, disconnections among the pixels along the block boundaries are inspected. Any disconnection or discontinuity among the pixels is considered to be indicative of blocking distortion. Histogram of the gradient is used to identify the block boundary. Mean of pixels along the vertical and horizontal boundaries gives a quality score. This approach shows good results for JPEG images from LIVE dataset. However, proposed NR metric applies to JPEG images only. Authors recommended further improvement in the NR metric to deal with other distortions.

An NRIQA metric proposed by C. Chen *et al.* [46] evaluates blockiness in the JPEG images. It first calculates the absolute difference between adjacent pixels in the horizontal direction. This difference is normalized, averaged along each column, and 1D Discrete Fourier Transform (DFT) is applied to derive vertical blockiness measure. Similarly, horizontal blockiness measure is defined. Average of both the measures defines the feature for quality prediction. The results show good accuracy

for known as well as unknown block size. However, equal weights are assigned to DFT coefficients at different DFT frequencies while computing Mean Square of DFT at the peak locations, which is the limitation of this metric.

S. Amiri *et al.* [47] developed an NR metric for JPEG images based on localized discrete cosine transform. Motion blur filter is applied to an input image to obtain the sharpness map. An image may have low frequency information block called as a uniform block and high frequency information block called as a non-uniform block. Non-uniform blocks can be differentiated from uniform blocks by applying threshold value on sharpness map which is called as binary sharpness map. This metric only considers the non-uniform blocks because DCT coefficients of these blocks are modified due to the compression. Elements of the blocks are divided by DC value in order to normalize the DCT coefficients. The quality is predicted using logistic non-linearity transform. This metric shows good correlation between objective and subjective scores. However, the model loses its generality because of the consideration of threshold value empirically.

Blockiness and improper luminance change the structure of an image. Y. Zhan *et al.* [48] proposed the features to measure this change within an individual block in the image. The gradients of an image are normally used to depict the structure. The proposed metric uses the gradient information to extract the features to estimate blockiness and luminance change. Blockiness is measured by averaging the amount of discontinuities along the four boundaries of an individual block. The luminance change is calculated by averaging the gradients inside the block. The features are computed locally for 8×8 non-overlapping blocks, whereas the block values are averaged to obtain the global feature. The mathematical model predicts the quality score. Though this metric shows good performance, some blocks have zero value for both the features due to the high compression ratio or flat areas in an image. These blocks are not considered for quality prediction which is the limitation of this metric.

JPEG2000 images suffer from blur and ringing degradations. Z. M. Parvez Sazzad *et al.* [49] developed an NR metric for JPEG2000 images using spatial features namely edge information and pixel distortion. Edge information is computed as Zero Crossing (ZC) rate in vertical and horizontal directions. The histogram measure is computed with and without edge preserving filter. The pixel distortion

is estimated as Standard Deviation (SD) of a central pixel within 5×5 neighbourhood pixels and absolute difference measure of a central pixel from the second closest neighbourhood pixels. The mathematical model predicts the quality score using logistic function to establish the nonlinear relation between the extracted features and the subjective score.

L. Liang *et al.* [50] proposed a good approach for JPEG2000 images. Proposed NR metric estimates blur and ringing artifacts through statistical features derived from the gradient profiles along the strong edges in an image. The model predicts the image quality with 94% accuracy for LIVE and 86% for TID2008 image datasets.

J. Zhang *et al.* [51] developed a model based on the structural activity in an image. Structural activity is computed in two stages. First, the structural activity weights are calculated followed by the local structural activity. The global structural activity is obtained by combining these two features.

J. Zhang *et al.* [52] used kurtosis feature for assessing the quality of JPEG2000 images. In JPEG2000 compression, high frequencies get attenuated. As a result, blur distortion is introduced in the image. The Probability Distribution Function (PDF) of blurry image illustrates a deeper tail than the natural image. Therefore, kurtosis increases with the increase in a blur. Average of the local kurtosis values are computed for overall quality estimation. This metric is simple and computationally efficient. The two-stage NR metric proposed by H. Liu *et al.* [53] deals with the ringing artifact present in the image. In the first step, the ringing detection method is defined, and in the second step, ringing annoyance is quantified. Firstly, the edges relevant for ringing are extracted to obtain Perceptual Edge Map (PEM). Then perceived ringing regions are identified to yield a Computational Ringing Region Map (CRRM). The detected region in the CRRM is then split into the ringing objects to quantify the ringing annoyance. The activity of each ringing object is calculated and compared with the activity in the neighbouring background. The algorithm works well for JPEG and JPEG2000 images.

The image with blur distortion suffers the loss of spatial details which affects the sharpness at edges. There are various reasons that can introduce blur in the images. Relative motion of an object and camera during the image acquisition process, use of smoothing filters in the image processing application are a couple of the examples.

In the compression techniques, high frequency elements are pruned due to the quantization which results in a blur. Conventional blur NR metrics generally concentrate on blur degradation occurred due to the compression [54]. Moreover, the ringing artifact is introduced in an image because of the quantization of DCT coefficients or DWT coefficients at high frequency.

P. Marziliano *et al.* [55, 56] developed a model to assess blur for FRIQA and NRIQA. It also assesses the ringing artifact for FRIQA. Proposed model analyses the neighbouring regions and edges. Estimation of ripple around edges is used to evaluate ringing. Edge is identified by determining the start and end point from the gradient image. Spreading of an edge is estimated to assess blur. FR model computes ringing artifact as the width of the ring around edges. This metric performs well as compared to the widely used PSNR metric. Autofocusing on cameras, coding techniques, and network resource management are the applications where this metric is useful. The accuracy of FRIQA and NRIQA metric is 87% and 73% respectively. The author suggested an extension to the proposed FRIQA metric to evaluate ringing artifact without reference metric.

A. Chetouani *et al.* [57] proposed a method that utilizes constraint on the blurring discriminationability of HVS. Authors added a small amount of blur and estimated its effect on the image quality. Using some HVS characteristics perceptual map is obtained for the two images. Quality index is computed by extracting geometrical features from the blur map. This metric shows good performance with 85% accuracy.

R. Ferzli *et al.* [58] integrated the concept of Just Noticeable Blur (JNB) into probability summation model and proposed an NR metric for evaluating the sharpness and blurriness. The proposed metric computes the probability of blur detection for predicting the relative blurriness present in an image. The algorithm detects the edge using a Sobel operator. Number of edge pixels is counted for each block. Edge block is determined from the number of edge pixels below the threshold value. The local contrast of the block is estimated to determine the JNB width. Further, the overall distortion is quantified by pooling all the local values that are computed. The proposed sharpness metric is the inverse of the overall distortion. Authors suggested that the performance of the algorithm needs to be improved for large blur values and the model is to be extended for noise. N. Narvekar *et al.* [59]

also proposed an NRIQA metric based on the concept of JNB with Cumulative Probability Blur Detection (CPBD). Authors developed a probabilistic framework based on the idea of the sensitivity of human to blur at the various contrast. The probability of blur at every edge can be quantified using the framework. This information at every edge is combined for the whole image for computing the overall quality score.

F. Kerouh *et al.* [60, 61] worked on blur assessment using wavelet transform. Wavelet decomposition is used for the edge analysis. High frequency coefficients of each decomposition level are extracted which are combined to obtain edge map. In wavelet transform, a pixel is assumed as blurred if the disparity between that pixel value and its neighbour is less than the fixed threshold. Blur assessment is done at different resolution levels. Quality is predicted using linear regression. This model uses an edge detection technique for blur assessment. Unlike blur metrics, this metric is free from the possibility of false detection of uniform areas as a blur. Application of multi-resolution analysis has an advantage of small image block size and less execution time.

L. Li *et al.* [62] proposed an algorithm to estimate blur using discrete orthogonal moments. Edges of an image are spread as an effect of blur which results in the shape alteration. Discrete moments are used as shape descriptors. The gradient of the image is divided into the blocks of equal size for which Tchebichef moments are computed.

The energy of each block is calculated as the sum of the squared non-DC moments. The algorithm calculates the quality score using the visual saliency model by normalizing the moment energy with the block variance. The reported accuracy for this algorithm is 96%.

C. Lim *et al.* [63] used gradient magnitudes and exact Zernike moments to develop an NRIQA metric to address Gaussian blur. The authors defined the mathematical relation between low order exact Zernike moments of the Gaussian spread point function and the standard deviation of the Gaussian distribution. SVM is used as a quality predictor which predicted the quality with 96% accuracy for LIVE. C. Vu *et al.* [64] presented an algorithm to determine sharpness using spectral and spatial properties. Authors calculated the slope of magnitude spectrum and total spatial

variations for each block to obtain spectral sharpness map and spatial sharpness map. By applying geometric mean, these two maps are combined to obtain the final sharpness map. The scalar index value is obtained to predict overall quality. Results show that this metric works well for blur images.

P. Vu *et al.* [65] developed a sharpness metric in the wavelet domain. Decomposition of an image is carried by wavelet transform up to three levels. Log energy of each sub-band at each DWT level is computed. Further, the total energy is calculated at each decomposition level. Finally, three scalar index values are computed for three decomposition levels. A similar algorithm is applied blockwise to estimate the local sharpness.

R. Hassen *et al.* [66] proposed a sharpness assessment metric using local phase coherence. The input image is decomposed into three scale sub-bands using steerable pyramid decomposition. Local phase coherence and the phase of complex wavelet coefficients show a consistent relationship across scales in the area of image features. The degree of local phase coherence at each spatial location is estimated as a derived feature for the assessment. The proposed metric correlates well with subjective assessment.

Partially blurred images can be assessed as good quality images by HVS, although contrast, saturation and other features of an image are affected. E. Mavridaki *et al.* [67] proposed quality assessment metric for partial blurred images in the frequency domain. As a result of a blur, high frequencies are attenuated. Therefore, the proposed model uses Fourier transform to evaluate the distribution of frequencies. Authors divided the image into nine blocks. The power spectrum is estimated for the blocks and the whole image as well. The feature vector is formed as five bin frequency histogram of these ten spectra. This feature vector is used to train SVM to estimate partial blur in test images. This metric shows good performance for naturally blurred images as well as artificially blurred images.

For assessing the blur distortion in colour images, M. Qureshi *et al.* [68] proposed a framework using higher order singular values. It first carried out the tensor analysis to represent the correlation among different colour components. Higher order single value decomposition is used as tensor decomposition technique. This metric shows good correlation with the subjective score.

Many researchers used DCT statistics for developing the NRIQA model. M. Saad *et al.* [19] defined the quality index by observing statistics of local DCT coefficient of an image. Two HVS based features, namely image structure and contrast, are computed. The local DCT contrast is defined as the ratio between non-DC DCT coefficient magnitude in the local patch and DC coefficient of the patch. The local contrast values from all patches of the image are then pooled together by averaging the computed values to obtain a global image contrast value. Structure features are derived locally from the local DCT frequency coefficients. To capture the statistical traits of the DC coefficient histogram, its kurtosis is computed. The kurtosis of each DCT image patch is computed. The resulting values are pooled together by averaging the lowest tenth percentile of the obtained values to determine the global image kurtosis value. Quality is predicted using a probabilistic model with 79% accuracy.

The noticeable blockiness NRIQA metric is proposed by G. Zhai *et al.* [69] for block-based DCT coded images. It determines the changes along the block boundaries to obtain block discontinuity map. Authors considered the effect of luminance adaptation and texture masking on the blocks. Local maps are combined using a nonlinear operator, and overall masking map is obtained. The overall masking map and discontinuity map are integrated to obtain Noticeable Blockiness Map (NBM). NBM is used for estimation of the quality score. The accuracy of this model is 96% for JPEG images.

T. Brando *et al.* [70] proposed DCT based quality metric which estimates original DCT coefficients from their quantized values available at the decoder. Association between distribution parameter at the neighbouring DCT frequencies is computed using linear prediction method. The results are then combined through Maximum Likelihood (ML) parameter estimation. For the quality assessment, PSNR is estimated using Laplacian parameter estimation. Accuracy for this metric is 97% for LIVE dataset.

For assessing the quality of JPEG, JPEG2000, noisy, and blurred images, J. Shen *et al.* [71] proposed a hybrid approach. Authors combined DCT, wavelet, and curvelet transforms. Curvelet is used to evaluate noise and blur present in an image. The wavelet transform is used to evaluate artifacts present in JPEG2000 images and DCT coefficient is used to evaluate the blockiness of JPEG images. Although a combined

approach, it evaluates artifacts individually rather than applying features simultaneously to the single image. Following table 2.1 briefly summarizes the representative research work carried out in the distortion specific NRIQA.

Table 2.1: Summary of the survey of the literature on the development in distortion specific approach

SN.	Author	Features	Prediction Model	Remarks
1.	Z. Wang <i>et al.</i> [43]	Difference determined across the block boundaries	Parameterized nonlinear regression model	Database dependent, deals with JPEG images only.
2.	S. Suthaharan [44]	Undistorted image edges, bit pattern of LSB	Parameterized Mathematical Model	Does not work well with highly compressed images, deals only with JPEG images.
3.	S. Lee <i>et al.</i> [45]	Histogram of the gradient	Statistical model	Deals only with JPEG images.
4.	Z. Sazzad <i>et al.</i> [49]	Spatial features	Logistic function	Deals only with JPEG2000 images.
5.	L. Liang <i>et al.</i> [50]	Statistical features of the gradient	Mathematical model	Blur, ringing, and JPEG2000.
6.	J. Zhang <i>et al.</i> [51]	Structural activity	Mathematical model	Deals only with blocking artifacts.
7.	J. Zhang <i>et al.</i> [52]	Kurtosis	Mathematical model	Ringing and JPEG2000.
8.	H. Liu <i>et al.</i> [53]	Perceptual edge map, computational ringing region map.	Mathematical model	Ringing, JPEG, and JPEG2000.
9.	P. Marziliano <i>et al.</i> [56]	Neighbouring regions and edges	Mathematical model	Deals only with JPEG2000 images.
10.	R. Ferzli <i>et al.</i> [58]	Just noticeable blur	Probabilistic summation model	Smaller blur values.
11.	N. Narvekar <i>et al.</i> [59]	JNB with a cumulative probability distribution	Mathematical model	Deals only with blur.
12.	F. Kerouh <i>et al.</i> [61]	Wavelet transform	Linear regression	Deals only with blur.
13.	M. Saad <i>et al.</i> [19]	Image structure and contrast	Probabilistic Model	Time consuming.
14.	G. Zhai <i>et al.</i> [69]	Changes along the block boundaries	Nonlinear operator	Blockiness, only for JPEG.
15.	T Brando <i>et al.</i> [70]	Estimation of original DCT from quantized values	Laplacian parameter estimation	Time consuming.

The features extracted and quality prediction model used in the methods are given in the above table. Development of the distortion specific NRIQA mainly focuses

on JPEG, JPEG2000 and blur. Mathematical models are used for quality prediction. These metrics show good performance. Some worked with DCT based features.

2.2.2 Feature Extraction and Machine Learning based NRIQA

The mathematical model predicts the quality score based on the model parameters which are estimated using subjective scores as a training data set. The prediction of the model sometimes does not correlate with the subjective score and leads to the incorrect prediction due to the nonlinear relationship between input features, predicted score, and subjective score as the model parameters are same for all the test images. Some researchers suggested a machine learning approach to deal with the issue [72, 73].

S. Suresh *et al.* [74] proposed a novel approach using extreme machine learning classifier. Overall quality is predicted by defining the functional relationship between subjective score and HVS based features such as edge length, edge amplitude, background activity, and background luminance. This model works well for JPEG images from LIVE dataset with 70% accuracy. The model is capable of handling the blocking artifact only. R. Venkatesh Babu *et al.* [75] proposed a novel neural network approach using the above features. Authors proposed Growing And Pruning Radial Basis Function (GAP-RBF) network to resolve the function approximation problem perceived by the neural network. The model demonstrated a very good correlation between subjective analysis and objective analysis.

P. Gastaldo *et al.* [76] developed an NR metric with non-parametric features based on the hypothesis that natural images are complex to be analyzed at the global level. Authors used three types of locally extracted features such as the first-order histogram of image blocks, features obtained from the co-occurrence matrix, and features obtained in the frequency domain. These feature vectors are used to train the Circular Back Propagation Network (CBPN) which predicts the quality. Images preprocessed with contrast stretching method are used for the feature extraction. This method shows good correlation between objective and subjective analysis with 86% accuracy.

A. Chetouani *et al.* [77] proposed a novel approach for blur and ringing artifacts using neural weighting scheme. The authors considered six existing IQA metrics to handcraft input features to ANN. The level of ringing and blur artifacts is estimated

firstly using ANN. The quality index is computed in the next step by combining blur and ringing measures using approximated weights during the learning process. This approach shows promising results.

F. Shao *et al.* [78] developed an algorithm to combine blocking, ringing, blur, and noise distortions with a machine learning approach for JPEG2000 images. Gradient histogram is used to evaluate blur, while frequency amplitude is used to compute the quantity of white noise. Further, DCT coefficient contrast and wavelet decomposition are used to evaluate the blockiness and ringing effects respectively in JPEG2000 images. Support Vector Regression (SVR) model is trained on the extracted features. C. Li *et al.* [79] proposed an NR metric using features such as phase congruency, entropy and gradient of the image. Authors used a General Regression Neural Network (GRNN) to predict quality score by finding the functional relationship between features and MOS.

L. Liu *et al.* [80] presented a two-stage framework for NR quality assessment; distortions are identified in the first stage followed by the quality assessment in the second stage. The spatial and spectral entropies are used as features. The image is downsampled by a factor of 2 and decomposed with three scales for the multiscale analysis. The block based entropy is calculated locally. Spatial entropy is a function of the probability distribution of the local pixel values while spectral entropy is a function of the probability distribution of local DCT coefficients. These features are pooled using percentile pooling method. Mean and skew values are computed for spatial and spectral entropies as final feature values. SVM is used to predict the image quality. The model uses pre-processed images for the feature extraction.

S. Wang *et al.* [81] proposed blur estimation using extreme learning machine model. The results of this model demonstrate that the extreme learning machine model has better learning performance and faster learning speed as compared to SVM. The proposed metric uses gradient information for blur assessment. The input image is convolved with Prewitt filter to obtain the gradient map. The gradient magnitudes are modelled with generalized Gaussian distribution. The shape, scale parameters, and mean of distribution features are used for quality prediction.

M. Zhang *et al.* [82] proposed an NR metric based on the joint statistics of Generalized Local Binary Pattern (GLBP). Laplacian of Gaussian (LoG) filter is

applied to an image to decompose it into multi-scale sub-band images. These sub-band images are encoded by applying GLBP operator, and features are developed by concatenating GLBP histograms of each sub-band image. The SVR model is used to map feature vectors to the subjective score. The linear correlation coefficient between objective and subjective score is 0.95 and 0.93 for LIVE and TID2008 dataset respectively.

F. Rezaie *et al.* [83] developed an NR metric using Local Binary Pattern (LBP) in the wavelet domain. In this metric, one level of the two-dimension discrete wavelet transform is applied to the input image to obtain four sub-bands namely LL, LH, HL, HH. These sub-bands are determined with the LBP operator separately. The histograms of these sub-bands are concatenated to define the feature and SVR is used to predict the quality score. The correlation between subjective and objective score determined by the metric is 0.93 for Live and 0.89 for the TID2008 dataset.

P. Freitas *et al.* [84] also proposed an NR metric using LBP. Authors, in their approach, compared texture information of a test image with the texture information of the set of texture images. Csiszar-Morimoto distance norm is used to determine the distance between histograms of local binary patterns of the test image and histograms of local binary patterns of the texture images used for training. The distance measures are utilized by the support vector regression model to predict the quality of an image. The linear correlation coefficient between subjective score and the objective score is 0.97 for LIVE2 database. P. Freitas *et al.* [85] further extended the metric using Local Ternary Pattern (LTP). LTP is the generalization of LBP; former shows the performance improvement over the later. However, as LTP is less sensitive to noise in the uniform region, multiple LTP channels are proposed to extract the texture information.

P. Feritas *et al.* [86] proposed another metric based on Multiscale LBP (MLBP) operator. MLBP operator is applied to extract the textural features. The parameters, radius and sampling points of LBP are varied to generate MLBP. Using multiple channels, determined by varying the radius parameter, histograms of LBP are computed. Authors extended their work with new salient LBP [87]. Authors used LBP to extract texture based features. The features are weighted by the saliency of image regions, which are calculated using a visual attention computational model.

The weighted histogram of features is given as input to the regression model for estimating the image quality.

S. Corchs *et al.* [88] proposed an NR metric to assess the quality of JPEG-distorted images. This metric is a combination of different metrics proposed by the other researchers. Authors combined eleven different metrics to assess the quality of JPEG images. SVM is employed for quality prediction. Results show good accuracy for a different number of classes. X. Shang *et al.* [89] proposed image quality assessment based on joint quality-aware features namely multiscale histograms of oriented gradient, multiscale edge information, and colour histograms. Further, the distortion indices are obtained by computing the difference between each feature of the reference image and distorted image. These indices are mapped to the subjective score using SVM. Authors stated that the metric showed a good correlation between subjective and objective scores.

L. Kang *et al.* [90] proposed Convolutional Neural Network (CNN) for NRIQA. In this approach, image patches are given as input to the CNN. The architecture of this network comprises of one convolution layer with min and max pooling, fully connected two hidden layers, and one output layer. Feature learning and regression are combined into one optimisation process within the network configuration. The linear correlation coefficient between subjective score and the objective score is 0.95 for LIVE database.

S. Bianco *et al.* [91] proposed DeepBIQ for BIQA by average pooling the scores predicted on multiple sub-regions of the pristine image. Authors demonstrated different design choices such as – use of feature extracted from pre-trained CNNs as generic image descriptors, use of fine-tuned features for estimating the image quality.

X. Lv *et al.* [92] proposed Convolution Feature Map (CFM) driven scheme for image quality assessment. Distorted images are filtered by convoluting with the learned kernel to generate convolution maps. A forward NSS layer performs on CFM to extracts NSS features. The trained SVR model maps the output of NSS layer to image quality.

H. Zeng *et al.* [93] developed a novel image quality representation scheme, Probabilistic Quality Representation (PQR), to describe the subjective score

distribution of each image. Authors transformed each scalar MOS into a vectorized PQR using the difference between the MOS value and each quality “anchor” value. This PQR representation for training samples is able to define a more robust loss function which speeds up the training process and achieve high prediction accuracy.

J. Bagade *et al.* [72] proposed a machine learning approach for the measurement of blocking artifact. Block based features, such as average of absolute difference between the pixels in the block, zero crossing rate, spatial frequency measures, and spectral activity measures are used for the quality prediction. Extracted features are used to train the back propagation neural network. The result shows that the metric correlates with the subjective score. Authors, in their extended work [73] proposed NR image quality assessment metric for evaluation of JPEG images using block-based features and statistical first and second order features are extracted in the frequency domain. These features and corresponding MOS values are used to train the back propagation neural network.

The representative research work carried out in the feature extraction and machine learning approach is briefly summarized below in table 2.2. The table depicts the extracted features and quality prediction model used by the researchers.

Table 2.2: Summary of the survey of the literature on the development in feature extraction and machine learning approach

SN.	Author	Features	Prediction Model	Remarks
1.	S. Suresh et al. [74]	Edge length, edge amplitude, background activity, and background luminance	Extreme machine learning	Distortion specific
2.	R. Venkatesh Babu <i>et al.</i> [75]	Same as above	GAP-RBF	Distortion specific
3.	P. Gastaldo <i>et al.</i> [76]	First order histogram, co-occurrence matrix, frequency domain	CBPN	Pre-process before feature extraction
4.	C. Li <i>et al.</i> [79]	Phase congruency, entropy, and gradient	GRNN	Less accuracy for Fast Fading images
5.	S. Wang <i>et al.</i> [81]	Gradient information	Extreme machine learning	Distortion specific
6.	M. Zhang <i>et al.</i> [82]	Joint statistics of Generalized Local Binary Pattern	SVR	Distortion generic.
7.	P. Freitas <i>et al.</i> [85]	Multiscale LBP operator	SVR	Distortion generic.
8.	S. Corchs <i>et al.</i> [88]	Eleven existing metrics	SVM	Distortion specific

This section presents the survey of literature on the feature extraction and machine learning approach. The work presented by P. Gastaldo using circular back propagation network is known as the pioneer work in this area. Different types of prediction models are implemented and used for quality predictions. The performance of this approach is based on feature extraction. In this approach, researchers mostly worked with distortion specific metrics. The researchers have recently focused on the development of distortion generic metrics.

2.2.3 Natural Scene Statistics based NRIQA

Designing an NRIQA model that can address different types of distortions is a complex and difficult task. Literature illustrates that natural scene statistics has immense prospective to resolve the problem of generalized NRIQA.

It is considered that natural images generate a small subset of a vast set of all possible images and show related statistical properties. The deviation is observed in these properties due to the distortions in that image. NSS based metrics quantify these distortions in the absence of the reference image.

Generally, normalized luminance coefficients of natural image pursue Gaussian distribution whereas this characteristic deviated due to the introduction of degradation. Based on this hypothesis, A. Mittal *et al.* [94, 95] developed an NSS based metric - “Blind Image Spatial Quality Evaluator (BRISQUE)”. The model uses the statistics of luminance coefficients that are normalized locally. Quality of the image is predicted using support vector regression. The model is computationally efficient as it is built without the transform. It demonstrates good accuracy.

A. Mittal *et al.* [96] proposed another excellent NSS based model “Complete Blind Quality Analyzer” for NRIQA. This model works with little knowledge about the degradation. Model extracts the quality-aware NSS based statistical features from the natural image and fits with Multi Variate Gaussian (MVG). Quality prediction is achieved by finding the distance between the distribution of the natural image and MVG distribution of the distorted image. This model with the machine learning approach demonstrated comparable results.

A. Moorthy *et al.* [97] proposed an excellent two-step framework based on the NSS model to assess image quality without reference. Distortions are identified first, and the degradation is evaluated in the following step. Probability weighted sum is used to predict the quality score. Wavelet coefficients of the natural image are fitted with Generalized Gaussian Distribution (GGD). Therefore, the coefficients of the distorted image are fitted with GGD. The classification of distortions is achieved using SVM. The feature vector is comprised of mean, variance, and shape parameters of different sub-bands in the wavelet domain. SVR is applied to calculate the quality with a similar set of features. Authors illustrated that this model is consistent with HVS and exhibits good results in distortion identification. Still, there is a scope for improvement for distortions like JPEG2000 and fast fading.

Y. Zhang *et al.* [98] proposed an extension to Distortion Identification-based Image Verity and INtegrity Evaluation (DIIVINE) metric called as a C-DIIVINE metric which uses complex Gaussian scale mixture model. Wavelet coefficients are fitted using three distribution models. The magnitude of wavelet coefficients are modeled with the complex generalized Gaussian distribution, relative magnitudes of coefficients are modeled with GGD, and relative phases of coefficients are modeled with the wrapped Cauchy distribution. Authors calculated the shape feature for all these distributions. Complex wavelet structural index is also computed to evaluate the distortion of correlations across the image scales. SVM is used to predict the quality score. Though the results show very good performance with 94 % accuracy; the metric is computationally demanding due to the use of transform and distributions.

Y. Jhang *et al.* [99] proposed an NSS based metric called DERivative Statistical based Image QUality Evaluator (DESIQUE). This model works in two stages, distortion identification followed by distortion specific evaluation. Model extracts features in frequency as well as the spatial domain. Log Gabor filters extract high frequency elements of an image in the frequency domain. Spatial features are computed as log derivative statistics using a pair of pixels. These features are fitted with GGD and parameters of the distribution are used for quality prediction. This model is computationally efficient and shows an edge in the performance when compared to the existing NRIQA methods.

D. Zhang *et al.* [100] used Independent Component Analysis (ICA) to decompose an image into independent elements. Distribution of each element is fitted with GGD. ICA generic model is built for the large sample sets of natural scenes which results in the set of filters. Each image is decomposed into different independent components by applying these filters. Features of the distorted image are fitted with GGD. Kullback-Leibler Divergence (KLD) is applied to compute the difference between modeled and distorted features. This model works well for white noise and Gaussian blur images. Authors suggested further experimentation with the model using ICA based statistical model.

H. Sheikh *et al.* [101] proposed an NR metric using a statistical model of natural images in the wavelet domain. This metric evaluates the wavelet coefficients of natural images in a given sub-band and their relationship with other wavelet coefficients across scales and orientations. Each sub-band vector is nonlinearly transformed and linearly combined using linear predictor to predict the quality score. This metric obtain 93% accuracy for JPEG2000 images.

M. Chen *et al.* [102] have developed an NSS model using multi-resolution decomposition. Features are extracted from gradient histogram. The distance between the gradient of natural image and the distorted image is calculated. Quality is predicted using SVM. The accuracy of the model is 96% for Gaussian blur images from LIVE dataset. M. Saad *et al.* [103] proposed DCT based NSS features. Authors worked with three features namely coefficient of frequency variation, energy sub-band measure ratio, and orientations. These features are obtained block wise, and the feature values are averaged for quality estimation. Reported accuracy of this approach is 82% for LIVE dataset. Y. Fang *et al.* [104] developed a model for contrast distorted images by extracting moments and entropy features. SVR model is used to predict the overall quality. Experimental results show promising performance with 80% accuracy. This model is distortion specific. The authors suggested the model need to be further extended for a wide variety of applications.

K. Gu *et al.* [105] designed a metric using free-energy-based brain theory. Authors defined the features based on free energy principle and structural degradation model, gradient magnitude, and structural information. Mean Subtracted Contrast Normalized (MSCN) coefficients are modeled using GGD. The quality is predicted

using the SVR model. Authors suggested the scope for the improvement in this metric by defining the features to address fast fading distortion.

L. Zhang *et al.* [106] presented a different and novel approach called the opinion-unaware method. NSS features are derived from the distribution of locally mean subtracted and contrast normalized coefficients which characterize structure distortion. The authors also derived gradient features for capturing the contrast distortion. To compute multi-scale and multi-orientation image properties Log-Gabor filters are applied to the image and the statistical features are derived from filter responses. Statistical features are derived from the image intensity distribution in logarithmic scale opponent colour space to define colour distortions. Features derived from the patches of the pristine image are modeled using multivariate Gaussian model. Bhattacharyya distance norm is used to measure the quality of each image patch, and overall quality is measured by average pooling. This metric shows good results when all features are used in combination. Results show that it has moderate computational efficiency.

L. Liu *et al.* [107] proposed an excellent NR metric in the curvelet domain. Authors extracted the intermediate level image features using curvelet transform. The model has based on the hypothesis that these features evaluate regularities occurring in low-level NSS model. Features are extracted as log PDF curvelet coefficients and energy orientation distribution. The statistical relation between features and image degradations is then evaluated. This model is computationally efficient and demonstrates 86% accuracy for the TID2008 dataset.

W. Lu *et al.* [108] developed an NSS based metric for NRIQA using contourlet transform. Directional multiscale dependencies are present in every natural image which are responsive to the distortions. The joint histogram of contourlet coefficients evaluates joint information among the coefficients at different scales, spaces, and directions. Image dependent threshold is computed to separate the joint space into two areas. Quality score is predicted by mapping the pooled coefficients with the most significant region.

G. Wang *et al.* [109] proposed an NSS based NR metric. Authors applied non-sampled shearlet transform to an image and statistical features, namely coefficient

distribution, energy distribution, and structural correlation across orientation and scale, are extracted. Quality is predicted using SVM.

L. Li *et al.* [110], in a novel approach, proposed an NRIQA using a gradient induced dictionary. The dictionary is built by computing gradient of natural images by applying K-mean clustering. Thereafter, features of images are extracted with the dictionary based on Euclidean-Norm coding and max pooling. SVM predicts the quality score. This approach is comparatively lesser complex. T. Feng *et al.* [111] proposed a framework using sparse feature extraction from sparse matrix representation which is evaluated using sparse coding method. This algorithm works in four stages – extractions of image patches from salient regions, learning the over-complete dictionary with unlabelled learning images, sparse matrix computation, and sparse feature extractions.

M. Qin *et al.* [112] developed a metric using hybrid NSS features. Authors constructed a codebook by extracting image patches from the original image using K-Means algorithm. The model combines the statistics of an image patch coefficients in codebook space with the statistical characteristics of local mean subtracted and contrast normalised coefficients in the spatial domain. After handcrafting statistical features, the two-stage framework is developed for image quality assessment. Distortions are classified first, followed by the quality score computation. As compared to other two-stage prediction models, the proposed model shows good results.

State-of-the-art NSS based frameworks first classify the distortion based on the group of features before using them for quality estimation. However, different distortions have different effects on certain sets of features. I. Nizami *et al.* [113] presented a three-step framework to address the issue. Authors extracted NSS based DIVINE features. SVM classifier determined the type of distortion using these features. Based on Spearman Rank Order Correlation Coefficient (SROCC) score, features are selected for the quality prediction. SVM regression model is used for quality prediction.

Table 2.3 briefly summarizes the representative research work in the area of NSS based NRIQA.

Table 2.3: Summary of the survey of the literature on the development in NSS based approach

SN.	Author	Features	Prediction Model	Remarks
1.	Mittal <i>et al.</i> [95]	Mean Subtracted Contrast Normalized (MSCN) coefficient statistics	SVR	Database independent, Comparatively less robust
2.	Mittal <i>et al.</i> [96]	Quality aware statistical features	SVM	Database independent, Time consuming
3.	A. Moorthy <i>et al.</i> [97]	mean, variance and shape parameters of different sub-bands in the wavelet domain	SVM, SVR	Distortion generic, Doesn't work well for JPEG compression
4.	Y. Zhang <i>et al.</i> [98]	Complex generalized Gaussian distribution, GGD and wrapped Cauchy distribution shape of the distribution	SVM	Database independent, Time consuming
5.	Y. Jhang <i>et al.</i> [99]	Log Gabor frequency domain, log derivative statistics spatial domain	Statistical model	Database independent, efficient, None
6.	D. Zhang <i>et al.</i> [100]	Independent Component Analysis	KLD	Use of ICA, Deals only with white noise and Gaussian blur
7.	H. Sheikh <i>et al.</i> [101]	Wavelet coefficients	Weighted average	Simple and efficient, Limited to JPEG2000
8.	M. Cheng <i>et al.</i> [102]	Gradient histogram	SVM	Simple and efficient, Distortion specific
9.	M. Saad <i>et al.</i> [103]	Coefficient of frequency variation, energy sub-band measure ratio, and orientation for DCT	Average of block based features	Simple to implement, Time consuming
10.	Y. Fang <i>et al.</i> [104]	Moments and entropy	SVR	Simple to implement, Distortion specific
11.	L. Liu <i>et al.</i> [107]	Log PDF curvelet coefficients and energy orientation distribution	Statistical model	Computationally efficient, Time consuming
12.	W. Lu <i>et al.</i> [108]	Joint histogram of contourlet coefficients	Mathematical model	Good accuracy, Image dependent threshold
13.	G. Wang <i>et al.</i> [109]	Non-sampled shearlet transform	SVM	Database independent
14.	L. Li <i>et al.</i> [110]	Gradient induced dictionary	SVM	Less complex, Time consuming

This section summarises the literature published in NSS based NRIQA approach. Initially, the development of the metrics was focused on usage of heavy transforms and distribution models. It makes them resource demanding and time consuming. Quality is estimated using different distance norms. Lately, the researchers started

implementing the distortion generic NRIQA metrics using features independent of the transforms. Nowadays, machine learning based quality predictors are also being used extensively.

Literature survey illustrates that initial research in the NRIQA domain was focused on the distortion specific metrics to address the specific distortions such as blockiness, ringing, blur and noise individually. Most of the distortion specific metrics use mathematical models which do not always establish a complex and nonlinear relationship between subjective score, extracted features, and objective score. Researchers explored the feature extraction and machine learning approach to overcome this drawback. However, most of the NRIQA metrics, developed using machine learning approach, are also distortion specific. Researchers also explored the use of optimised predictors. Lately, researchers started developing the generalised metrics using NSS features. It is observed from the literature survey that NSS based metrics use distributions and transforms heavily for the quality assessment. NSS features are handcrafted from the statistics followed by the natural images and assumed that other images also follow similar statistics.

Study of the literature shows the scope for designing the NRIQA metric which should assess the quality of an image for different types of distortions and independent of image processing applications. The combination of the above discussed approaches for NRIQA can lead to the designing of such framework.

2.3 Research Gaps

Literature survey shows that development of IQA methods began with focusing on FR methods. Availability of original reference image is required for these FR methods. However, in most of the real-world applications, original images often cannot be made available for the assessment or simply don't exist. No reference IQA techniques prove useful in such situations. Researchers have proposed various approaches to develop the robust NR metrics.

Until recently, the research work published in NRIQA domain is mostly related to the distortion specific metrics. Researchers have proposed various methods specific to the distortions. From the literature, it is observed that the growth of NSS based

generalized no reference image quality assessment methods have been accelerated in recent times. Use of transforms increases the time complexity of these metrics.

Machine learning based quality prediction model are also proposed by many researchers. Researchers suggested continued experimentation for the improvement of the methods on various counts such as simplicity in implementation, computationally less demanding, less time complexity, and ability to deal with different types of distortions and multiple distortions fused in an image as well.

Literature survey helped in better understanding of research being carried out presently in the area of image quality assessment. These observations helped in identifying the research gaps and defining the problem statement for this doctoral study. The research gaps observed are discussed below:

- Most of the existing quality assessment frameworks discussed in the literature are distortions specific. These methods, being distortion specific, do not address other types of distortions. Thus, a need to define a framework that can evaluate image quality irrespective of the distortion is identified.
- As reported in the literature, generalized natural scene statistics based NRIQA methods provide a better alternative to assess the image quality. However, as the statistical characteristics of natural images are used to establish NSS based methods, these methods can hardly give full consideration to the human visual characteristics. Also, the use of transforms and distributions make the NSS based metrics complex and computationally demanding. NSS based NRIQA metrics require improvement in this aspect.
- The parametric or non-parametric features are often handled individually while defining the quality metric. However, both these kinds of features play an important role in defining the quality metric. Use of combinations of different features has potential in the development of the generic quality assessment metric.
- Researchers have used different types of classifiers/predictors for quality estimation. Neural network based classifiers are used popularly. Adaptive learning, generalization capacity, and self organization are the advantages of ANN. However, understanding of the functionality of ANN is difficult as it hides the relation between inputs and outputs in the weights of the neurons

behind its hidden layers. Also, the number of hidden layers and the number of neurons in each hidden layer are determined randomly based on the trial and error method. Fuzzy or neurofuzzy classifiers can be used to overcome this limitation of ANN.

The above discussed research gaps points to an opportunity for further research in the area of NRIQA. Literature survey helped in identifying the need to design the framework capable of addressing different distortions for quality estimation of images. It also helped in realizing the potential of machine learning approach for developing the lightweight and efficient framework. Literature survey leads to understand the use of a combination of different features in overall quality estimation of digital images and its significance in designing the framework. Thus, the scope is identified to design an optimum framework for NRIQA by defining the optimized set of features and the efficient classifier.

Chapter 3.

Methodology

Image quality assessment has an important role to play in the field of image processing. Distortions are introduced in the images during various image processing tasks and applications. Objective IQA algorithms are developed to estimate these distortions. Various IQA algorithms have been developed for estimating the image quality by the researchers over the last few decades. IQA development approaches are mainly classified in three categories as FRIQA, RRIQA, and NRIQA. NRIQA metrics attempt to estimate image quality in absence of the pristine or original image. NRIQA approaches are usually categorized as a distortion specific, feature extraction and machine learning, and NSS based. Distortion specific and machine learning approaches focus on individual or specific distortion for quality estimation of an image. On the other hand, NSS based metrics are computationally demanding due to heavy use of the transforms. This study proposes a combination of metrics irrespective of the approach they follow.

The need of NRIQA framework for estimating the image quality irrespective of the distortion has been identified through the literature survey. Machine learning based and NSS based are the two approaches usually followed to develop the distortion generic NRIQA metrics. A machine learning based framework is proposed in this study for overall image quality estimation. It uses combination of the features and learns from the combined features.

This chapter is organized in 6 sections beginning with the motivation discussed in section 3.1, followed by the problem statement in section 3.2. Section 3.3 describes the methodology for the proposed framework. Image datasets used for the experimentations performed throughout the study are discussed in section 3.4. While section 3.4.1 discusses about the created dataset, subsequent sections present the detailed description about standard image datasets LIVE, TID2008, and TID2013. Subsection 3.4.6 presents the preparation of training and testing image datasets. Performance evaluation of NRIQA metrics is discussed in section 3.5 and working of the ANN classifier is presented in section 3.6.

3.1 Motivation

While working with a group of experts in compression, it has been noticed that the performance of the compression scheme is evaluated by assessing the quality of decompressed images using full reference image quality measures. Though availability of a reference image is feasible in such experiments conducted in the laboratory, in real-world image processing applications images are usually acquired, processed, stored, and transmitted without the reference images. Sharing of images over the social media applications is one such example. These applications transmit images over the internet without any reference image. The quality determination of the received image is highly essential for better user experience in such applications. This defines the necessity of NRIQA in image processing applications.

Further, it is a well accepted principle that the objective quality assessment of an image should be consistent with that of HVS. The example of “stretched contrasted” image can explain the importance of this hypothesis. Images with stretched contrast during the pre-processing for better edge detection suffer the loss of overall image quality. The objective image quality evaluation metrics can rightly identify and estimate the degradation using various algorithms. However, quality of the same image may appear as better or improved to the human eye. This scenario suggests the importance of the development of NRIQA metrics that are consistent with HVS.

Most of the NRIQA metrics developed so far are distortion specific. However, an image may acquire different distortions when processed by different image processing applications. The visual appearance is different for the different distorted variants of a pristine or original image though they are from the same content and have the same MSE [114]. Also, rarely an image is degraded by only one distortion. Quality of such an image cannot be estimated by simply assessing it for any one specific distortion or another. Both of the above scenarios suggest a scope to design image quality assessment metric that is independent of the distortions.

The literature survey shows that the research in the area of image quality assessment has achieved sustainable growth in this decade. The authors [115, 116, 117] suggested that there is a scope to improve the existing NRIQA algorithms for better computational efficiency, adaptability, and accuracy. Also, these metrics need to be designed so that they can easily be implemented in real-time applications.

The above observations suggest an opportunity for developing an NRIQA metric that follows a generic approach for better adaptability, for the different distortions. A framework capable of quantifying the different distortions by accommodating different NRIQA metrics is thus considered for the present study. Hence, an optimum framework based on a generic statistical model is proposed for NRIQA.

3.2 Problem Definition

The literature illustrates that the NR quality methods are developed using parametric or non-parametric features and are effective only in case of one or two degradations. These methods are used for particular image type, such as JPEG, by applying limited types of classifiers available. The research gaps identified from the study of the literature, as discussed in chapter 2, helped in articulating the problem definition for this study -

The proposed research work aims towards the generic approach of combining parametric and other level-two features optimized with Principal Component Analysis (PCA) using fuzzy classifier or equivalent. The proposed framework is supposed to assess the quality score of gray images with high accuracy and be applicable to different types of distortions as well. In short, the proposed problem is to design the optimum framework for no reference image quality assessment based on generic statistical modeling.

Following objectives are defined to provide a solution to the above stated problem.

1. To study, implement and comparison of the various types of features.
2. Implementation of NSS concept and combine the extracted features with the other features in order to improve accuracy.
3. Dimensionality reduction and dataset management using PCA.
4. Design and implementation of the fuzzy or neurofuzzy classifier.
5. Design of diagnostic type of feedback mechanism for realization of system prediction.

3.3 Research Methodology

The framework is proposed for the quality estimation of gray images irrespective of the distortions they suffer. It is proposed to build the feature repository of the extracted relevant features from the training images. The combined feature vector residing in the feature repository is optimized by applying PCA. The machine learning based classifier is trained on this optimized feature vector for quality estimation of test images.

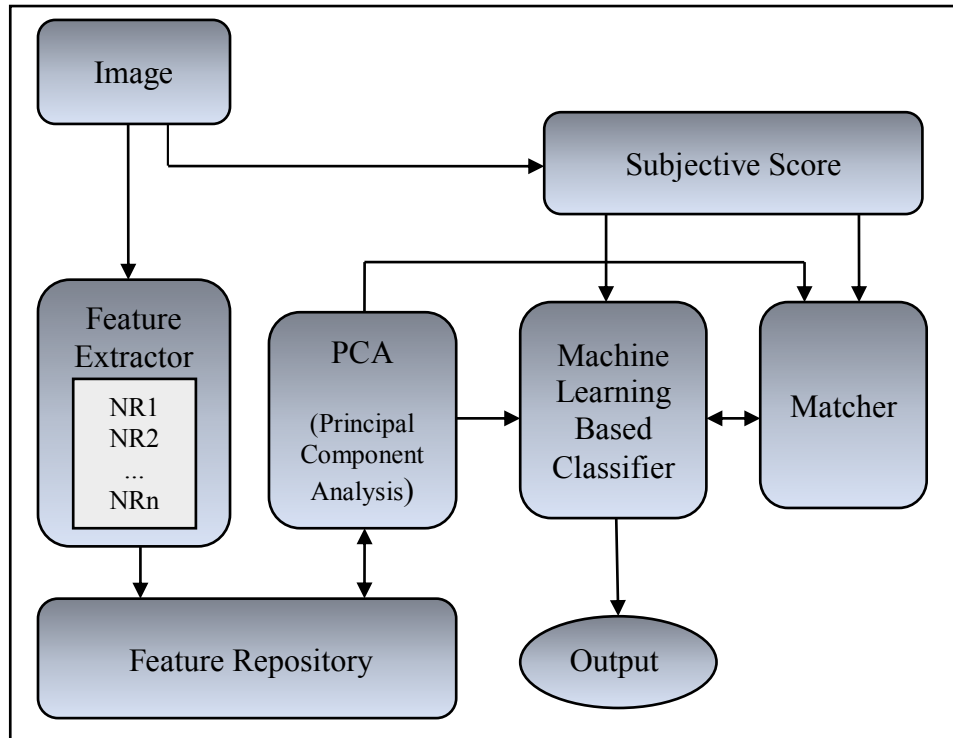


Figure 3.1: Schematic representation of the proposed framework

The methodology of the proposed framework is depicted through the schematic representation in the above figure 3.1. The same is described below.

- Various types of features such as statistical features, parametric and other level-two features, and some existing features are implemented are to address different distortions.
- NSS based features are implemented and combined with other extracted features and stored in the feature repository.
- The combination of features generates a huge feature data. Hence, PCA is applied to obtain significant feature vector. PCA first understands the

distribution of objective features that best characterizes the information in the image data.

- The optimized classifier is an important part of this framework. The neurofuzzy and neuro-wavelet based classifiers are trained using the proposed features. The trained classifiers associate the selected features from the test images with the corresponding predictions of quality ratings.
- The feedback type probabilistic neural network is designed for updating the weight matrix in the probabilistic sense. The feedback continued until the classifier indicates that the image quality is of a good class.

3.4 Image Datasets used for The Experimentation

The standardized image datasets are used by the researchers to compare and benchmark the performance of objective IQA metrics. The experimentations presented in this thesis are performed on widely used image datasets LIVE, TID2008, and TID2013. However, the initial experimentation, the analysis of the existing metric proposed by Z. Wang *et al.* [43], is carried out on a dataset created for the experimentation as described below.

3.4.1 Image Dataset Created and Used for Analysis of Existing NR Metrics

The study began with the analysis of block based feature metric proposed by Z. Wang *et al.* [43]. This study is discussed in chapter 4, section 4.1.1. The authors have created their own image dataset for the experimentation. Therefore, a new dataset is created for the performance analysis of the metric proposed by the authors. This image dataset comprises 60 reference images sourced from “ComputerVisionresearchgroup.com” and LIVE dataset. These pristine images were distorted with 50% blur. The images are compressed with quality factor 10, 50, and 70 using IrfanView© software to create four sets of distorted images. Total 20 subjects, 10 experts, and 10 non-experts, were chosen for the subjective quality evaluation. Total 300 images, including the reference images, were presented to the subjects. Subjects were asked to score their opinion of quality for each image in the range of 1 to 10; 1 for the reference image and higher with the increasingly degrading

image. MOS is obtained by averaging the subjective quality score. Experimentation was carried out in a real-life environment.

The sample images used for the experimentation are shown below in figure 3.2.

Preparation of Training and Testing Datasets from Created Image Dataset:

Two sets of different images are created for training and testing to validate the model. The set of images created for training contains 200 images including 40 pristine JPEG images, their three variants distorted using different quality factors, and a blur variant. Whereas, the set created for testing contains 100 images which include 20 pristine images and its distorted variants.



Figure 3.2: Sample images from the created dataset

3.4.2 Laboratory of Image and Video Engineering (LIVE)

A human being is considered as the ultimate receiver and assessor of perceived images. The IQA model has to predict image quality consistent with the human opinion. A standard set of images pre-assessed by the human subjects is required for the standardization of the performance evaluation of objective image quality assessment metric. These objective metrics are supposed to be evaluated on standard set of images.

International Telecommunication Union (ITU) has laid down the guidelines to create such image datasets [118]. Extensive experimentation is conducted by LIVE as per

these guidelines in collaboration with the Department of Psychology at the University of Texas to create the image dataset called as LIVE.

LIVE dataset consists of 29 original low-frequency and high-frequency images which are the pictures of human beings, animals, man-made objects, landscape scenes, wide-angle and close-up shots, etc. [119]. It has five types of distortions such as JPEG, JPEG2000, Gaussian blur (Gblur), White Noise (WN), and Fast Fading (FF), with five levels of distortions along with the pristine images. Dataset is comprised of a total number of 779 distorted images having DMOS in the range of 0 to 80. The single-stimulus methodology was used for the experimentation to record subjective quality score [120].

Pristine images with varied image contents and their distorted variants degraded with different distortions of different strengths were evaluated in the same session by a large number of human subjects. Overall seven sessions were conducted including two sessions each for JPEG and JPEG2000, and one session each for Gblur, WN, and FF. The average number of participating subjects for each session was 23, and each subject was allowed to participate in one session only. The selected subjects had no background of studying in the area of IQA. However, they were trained to assess the image quality using certain images. Images used to train human subjects were different from those used for the subjective evaluation. Subjects were presented with images in the random order on the screens, and the order of randomization for each subject was different. They were asked to assign the score in the range of 0 to 100 to the image based on their assessment of image quality. The reference image is assigned the zero score, and higher score recorded for the increasingly degraded image. Subjects recorded their opinion about image quality using a quality scale [120]. CRT monitors with the resolution of 1024×720p were used for this experiment.

The LIVE dataset is made freely available to the researchers to evaluate the performance IQA metrics. DMOS values evaluated during subjective analysis of all reference and distorted images are also provided by the LIVE team.

The collage of the sample images from the LIVE dataset is shown in figure 3.3(a).



Figure 3.3(a): Sample images from LIVE dataset.

3.4.3 Tampere Image Database 2008 (TID2008)

Nikolay Ponomarenko *et al.* [121, 122] developed another image dataset named as TID2008 and made freely available to the researcher community in 2008. It comprised total 1700 distorted images for 25 reference colour images sourced from Kodak database ([url-“http://r0k.us/graphics/kodak/”](http://r0k.us/graphics/kodak/)) and cropped to the same fixed size of 512×384 p. There are 17 types of distortions provided in this dataset, namely additive Gaussian noise, additive noise in colour components, JPEG, JPEG2000, masked noise, spatially correlated noise, impulse noise, high frequency noise, quantization noise, image de-noising, Gblur, contrast change, JPEG transmission errors, JPEG2000 transmission errors, non-eccentricity pattern noise, mean shift, local block-wise distortions of different intensities with four levels of distortions. PSNR for the four different level distortions are 30, 27, 24, and 21dB. Images in this database are in uncompressed bitmap file format. Total 838 experiments were carried out in three countries; Ukraine, Finland, and Italy. The tri-stimulus methodology was used to record a subjective quality score. The subjects were presented with a pristine image and its two distorted variants simultaneously to choose better one from the distorted images. Comparisons were designed so that each distorted image participated in nine comparisons. Each chosen image scored one point for each comparison, and the scored points were summed up.

The scores assigned by the subjects to each tested image were averaged after eliminating the abnormal ones to obtain the MOS values in the range from 0 to 9.

Higher MOS value implies better visual quality assessed by the subjects. The authors provided these MOS values with the TID2008 dataset. This dataset is used widely in performance evaluation of objective metric by comparing predicted score with MOS [121, 122]. The collage of sample images from the TID2008 dataset is presented below in figure 3.3(b).



Figure 3.3(b): Sample images from TID2008 dataset

3.4.4 Tampere Image Database 2013 (TID2013)

Authors *et al.* [123, 124] extended TID2008 dataset to create a new image dataset called TID2013 by introducing 7 new types of distortions, viz. change of colour saturation, Multiplicative Gaussian noise, Comfort noise, Lossy compression of noisy images, Image colour quantization with dither, Chromatic aberrations, Sparse sampling and reconstruction. Thus, it takes a total number of distortion types to 24 with five levels of distortions for each of them. It comprises of total 3000 distorted images for 25 pristine images. PSNR for the new fifth level of distortion is about 33dB. Total 971 experiments were carried out. This new dataset made freely available in 2013.

TID2013 image dataset is created with the similar methodology that was used for creating TID2008 dataset barring few changes such as; shorter time is allotted to evaluate each test image, and comparisons were designed so that each distorted image had to go through an equal number of quality assessment tests.

3.4.5 Summarized Comparison of LIVE, TID2008, and TID2013

Observation conditions such as monitor display size and resolution, display type, lighting, observer's distance and angle, etc. matters in the subjective quality evaluation of images [125]. LIVE and TID differ on these criteria in their subjective quality evaluation methodology. While LIVE followed well-controlled experimental conditions and strict compliance with ITU recommendations, observation conditions were variable in TID [123, 125]. The summary of the characteristics of the above discussed datasets is presented in table 3.1.

Table 3.1: Summarised characteristics of image datasets LIVE, TID2008 and TID2013.

Datasets	Characteristics						
	Reference Images	Distortion Types	Distorted Images	No. of Experiments	Subjective Scores	Subjective Score Evaluation	
						Methodology	Scale
LIVE	29	5	779	161	DMOS	Single-stimulus	0 to 100
TID2008	25	17	1700	838	MOS	Tri-stimulus	0 to 9
TID2013	25	24	3000	971	MOS	Tri-stimulus	0 to 9

3.4.6 Preparation of Training and Testing Datasets from LIVE, TID2008, and TID2013 Image Datasets

Separate training and testing image datasets are prepared for each standard dataset that is used for the experimentations performed for the study. The images from each standard dataset are divided into training and testing image sets such that two sets comprise different images. Following table 3.2 describes the total reference images contribute in preparation of training and testing datasets along with the number of distortion variants and distortion levels. Though the images included in the testing datasets are different from the training datasets, they share some common images. Three images from LIVE, and five each from TID2008 and TID2013 are shared between training and testing datasets. Five variants, namely JPEG, JPEG2000, Gblur, WN, and FF distortion variants from LIVE are used for the experimentations. TID2008 and TID2013 both dataset do not have FF distortion variant, hence remaining four are considered. These training and testing datasets are used for the experimentations carried out throughout this research study.

Table 3.2: Preparation of Training and testing datasets for LIVE, TID2008, and TID2013

Dataset	Training Images				Testing Images			
	Reference images	No of Distortion types	Levels	Total	Reference images	No of Distortion types	Levels	Total
LIVE	20	5	5	600	12	5	5	360
TID2008	18	4	4	360	12	4	4	240
TID2013	18	4	5	432	12	4	5	288

3.5 Performance Evaluation

Video Quality Experts Group (VQEG) has defined the procedure to evaluate the performance of IQA metrics [126]. The performance of an IQA metrics is evaluated by comparing the predicted score with the subjective score which represents the perceived quality. According to the guidelines set by VQEG, the performance of the objective methods may be analysed by evaluating the characteristics of their capabilities to predict the subjective evaluation [117]. There are two characteristics as follows -

Prediction Accuracy – It is the measure of the ability of a metric to predict subjective quality score precisely by maintaining error level at the lowest.

Prediction Monotonicity – It is the measure of a degree to which the predicted score correlate with the subjective score.

Literature illustrates that Root Mean Square Error (RMSE) and Pearson’s Linear Correlation Coefficient (PLCC) are extensively used to calculate the accuracy of the model. Spearman Rank Order Correlation Coefficient (SROCC) computes the monotonicity.

PLCC and SROCC between two datasets p and q are given in equation 3.1 and equation 3.2 respectively. RMSE is computed as shown in equation 3.3.

Suppose, S be the total number of data samples in the datasets p and q , then $\bar{p}$ represents mean of p and $\bar{q}$ represents mean of q .

$$R_{1pq} = \frac{\sum_i p_i q_i - (\sum_i p_i)(\sum_i q_i)}{\sqrt{S \sum_i p_i^2 - (\sum_i p_i)^2} \sqrt{S \sum_i q_i^2 - (\sum_i q_i)^2}} \quad (3.1)$$

$$R_{2pq} = \frac{\sum_i (p_i - \vec{p})(q_i - \vec{q})}{\sqrt{\sum_i (p_i - \vec{p})^2 \sum_i (q_i - \vec{q})^2}} \quad (3.2)$$

$$rmse = \sqrt{\frac{\sum_i (q_i - p_i)^2}{S}} \quad (3.3)$$

3.6 Study of Artificial Neural Network Classifier

ANN classifier is a type of non-parametric algorithms which relies on a three-layered neuro network machine. This network learns to classify the input vector $X_1, X_2, X_3, \dots, X_q$ into output classes $Y_1, Y_2, Y_3, \dots, Y_n$ by adjusting the weights of the layers iteratively. The values $V_1, V_2, V_3, \dots, V_n$ are linear combinations of the output from the hidden layers and are weighted by a threshold function to produce the classified pattern. The threshold function is usually a “sigmoid” function. The following figure 3.4 depicts the basic architecture of ANN.

ANN is a computational framework motivated by understanding of the process of acquiring knowledge by the network of biological neurons present in human brain. The fundamental function of an artificial neuron is to transform the weighted sum of their inputs using a nonlinear hard limiting or soft squashing functions. Artificial neurons are organized in interrelated layers to build ANN. Generally, initial weight values of the interconnections are allocated arbitrarily.

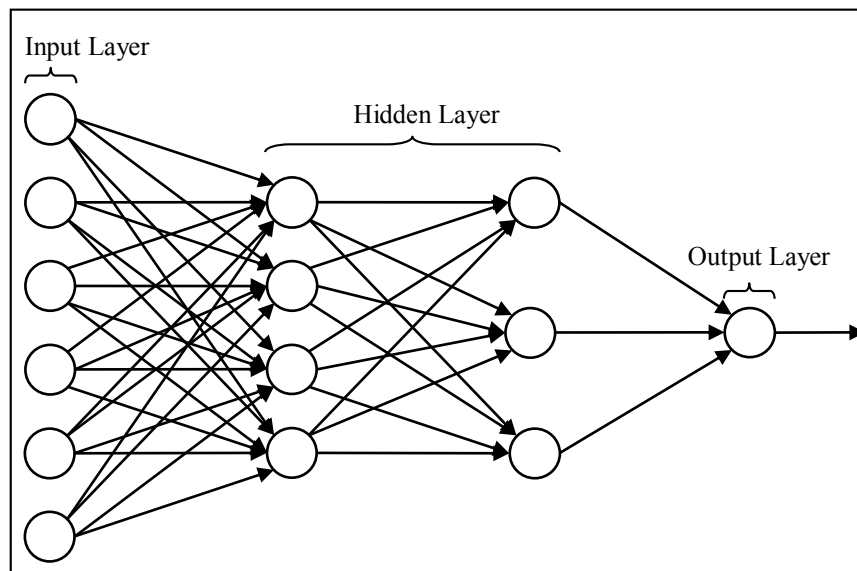


Figure 3.4: The basic architecture of ANN

ANN functions in two phases, namely training and recall. The learning rules are classified as supervised, unsupervised, and reinforcement and are applied in the training phase. In the supervised learning, the weight of each connection is altered to reduce the error between the expected and the obtained output during the training. On the other hand, in unsupervised or competitive learning, weights are adjusted to maximize the differences between the output categories. In the reinforcement learning, functional relationship of input-output is learned through the continued communication with the environment so as to gear up the performance scalar index. In recall step, ANN produces an output for the problem using new input data, called as test data.

ANN is evolved around the concept of neurons, interconnections, and transfer functions. One of the popular ANN architecture is multilayer feedforward back propagation network. This efficient and simple to learn model is used for composite multilayered networks. A huge group of networks generated by back propagation architecture are categorized using different topologies and training algorithms. The strength of this architecture is its capability to generate non-linear solutions to the poorly described problem.

The architecture of the back propagation network has an input, an output, and minimum single hidden layer. Hypothetically, there is no restriction on the number of hidden layers used in the network. However, generally, one or two are used. Every preceding layer is entirely connected to the succeeding layers. The input and output layers show the flow of data in the recall process. In the recall phase, the test data is fed to the trained network to obtain a result from ANN. The back propagation algorithm is applied in the learning phase with the training dataset.

In the majority of learning phases, network calculates the differentiation between obtained output and the expected output. The unprocessed error is changed through the error function to match a particular network structure. This error is directly utilized by most of the architectures. However several networks square the error, many cube the error, and many paradigms change the raw error to fit according to the function. The error is propagated as a current error to the learning function of further neurons. This error is usually transmitted backwards to the preceding layer. The back circulated value can be the current error, scaled value, usually the derivative of the transfer function or a few other expected outputs as per the network

type. Usually, the back propagated value, which is scaled by the learning method, is multiplied by the weights at the incoming connection to change them before the subsequent learning iteration. In the present study, feedforward ANN has an input layer with the number of neurons equal to the size of feature vector followed by the hidden layer with nine neurons, and an output layer with one neuron for quality prediction. Neurons in the hidden layer uses the log-sigmoid transfer function whereas, neuron in the output layer uses linear transfer function. Levenberg-Marquardt training algorithm is used to train the network. The non linear transfer function of the neurons in the hidden layer allows the network to acquire the knowledge about the correlation between the input and the output [127, 128].

However, this architecture is not without drawbacks. The back propagation network requires extensive supervised training using many input-output exemplars. Moreover, the inner mapping processes are not well realized, and do not assure that the system will lead to the satisfactory solution. Sometimes, the training gets caught in the local minima, and the best possible solution could not be achieved. This happens when the network system computes an error which is lesser than the nearby possibilities, and thus it does not reach the least probable value.

Chapter 4.

No Reference Image Quality Assessment using Textural Features

Study of the literature has illustrated that the researchers have explored different approaches namely distortion specific, machine learning, and NSS based, to develop NRIQA metrics for image quality estimation. Initially, researches were focused on the development of the distortion specific metrics. The mathematical models were used for the quality prediction. Generally, these mathematical model perform well for predicting the image quality. NSS based approach is also extensively explored by the researchers for developing the NSS metrics. Unlike distortion specific metrics, NSS metrics are distortion generic. These NSS based NRIQA metrics use distributions and transforms heavily and hence complex and computationally demanding [116, 117].

No single handcrafted image feature can convey all the information required to assess image quality for the different classes of distortions [114]. The framework proposed in this thesis is aimed to estimate the image quality for different distortions. A set comprising a number of different features is an essential part of the proposed NRIQA framework. Different features are thus required to be extracted and evaluated using the quality prediction models before combining or merging them into a larger set of features. The experimentations discussed in this chapter are performed to establish the ground for the proposed framework. Machine learning approach is followed for the experimentations in this and other contributory chapters across the thesis.

The study begins with an analysis of the existing metric. The NRIQA metric for most common JPEG distortion proposed by Z. Wang *et al.* [43], is considered for this study. The experimentation is further extended by combining blocking parameters with first and second order statistical parameters. An image can also be affected by a distortion other than JPEG. Therefore, a generic approach is proposed to quantify the different types of distortions by combining the blocking parameters with second ordered histogram features, and frequency based features as proposed

by P. Gastaldo *et al.* [76], and DCT based NSS parameters proposed by M. Saad *et al.* [103]. It has been observed that the performance of this metric is average for fast fading images. It is required to define more features that are sensitive to the fast fading distortion to address this performance drawback of the above metric.

The texture is a complex pattern consists of entities or subpatterns, i.e. repetition of patterns, which have features such as colour, size, brightness, etc. [129]. It helps to differentiate objects from their surroundings. The distortions deteriorate the texture details and degrade the image quality. The change in the texture of an image can be quantified using textural features. Therefore, the textural features are considered for the experimentation [130].

The study of the existing block based metric and the implementation of a textural feature metric is presented in this chapter and is broadly categorized in three sections.

Analysis of existing block based NRIQA metric proposed by Z. Wang *et al.* [43] is discussed in section 4.1. This metric is combined with other existing metrics proposed by P. Gastaldo *et al.* [76] and M. Saad *et al.* [103] for the further experimentations which is discussed in the subsections. Section 4.2 discusses the design of the textural feature metric, methodology, and results. Section 4.3 summarizes the chapter

4.1 Analysis of Existing NR Metrics

Study of the existing NR metrics is essential to understand the working of NR metrics. The NR metric proposed by Z. Wang *et al.* [43] is considered as a pioneer work in the area of NRIQA. Therefore this metric is considered for the study. The authors validated the proposed NR metric on the created image dataset. Therefore, a new dataset is created for the analysis of the existing metric proposed by the authors. Preparation of the image dataset is discussed in chapter 3, section 3.4.1.

4.1.1 Study of The Existing Block based NR Metric

Z. Wang *et al.* [43] proposed the block based NR metric in their pioneer work. Authors extracted the average absolute difference between the in-block image sample and zero-crossing rate features for a grayscale JPEG image. Authors used

these features for the quality prediction using nonlinear regression model. The predicted model and the features used in the experimentation are discussed below.

4.1.1.1 Block based features

Following expression represents the image.

$$s(x, y) \text{ for } x \in [1, X] \text{ and } y \in [1, Y] \quad (4.1)$$

A differencing signal along each horizontal line is calculated as

$$p_h(x, y) = s(x, y+1) - s(x, y), y \in [1, Y-1] \quad (4.2)$$

Following features are evaluated based on the pixel variations across rows and columns.

The average difference across block boundaries is calculated to obtain blockiness as,

$$bl_h = \frac{1}{X(\lfloor Y/8 \rfloor - 1)} \sum_{i=1}^X \sum_{j=1}^{\lfloor Y/8 \rfloor - 1} |p_h(i, 8j)| \quad (4.3)$$

Image signal activity is calculated as two factors. The first factor, the average absolute difference between in-block image samples is calculated as

$$avg_h = \frac{1}{7} \left[\frac{8}{X(Y-1)} \sum_{i=1}^X \sum_{j=1}^{Y-1} |p_h(i, j)| - bl_h \right] \quad (4.4)$$

The other factor, zero crossing rate (zc) is calculated as for $y \in [1, Y-2]$,

$$zc_h(x, y) = \begin{cases} 1 & \text{horizontal } zc \text{ at } p_h(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (4.5)$$

zc rate is estimated horizontally as

$$zc_h = \frac{1}{X(Y-2)} \sum_{i=1}^X \sum_{j=1}^{Y-2} zc_h(x, y) \quad (4.6)$$

Features computed row and column wise are averaged as below

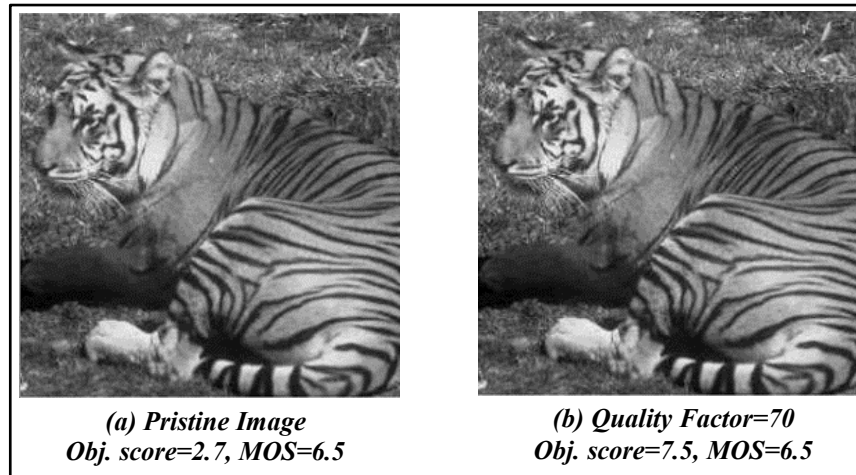
$$bl = \frac{bl_h + bl_v}{2}, \quad avg = \frac{avg_h + avg_v}{2}, \quad zc = \frac{zc_h + zc_v}{2} \quad (4.7)$$

The score is calculated as

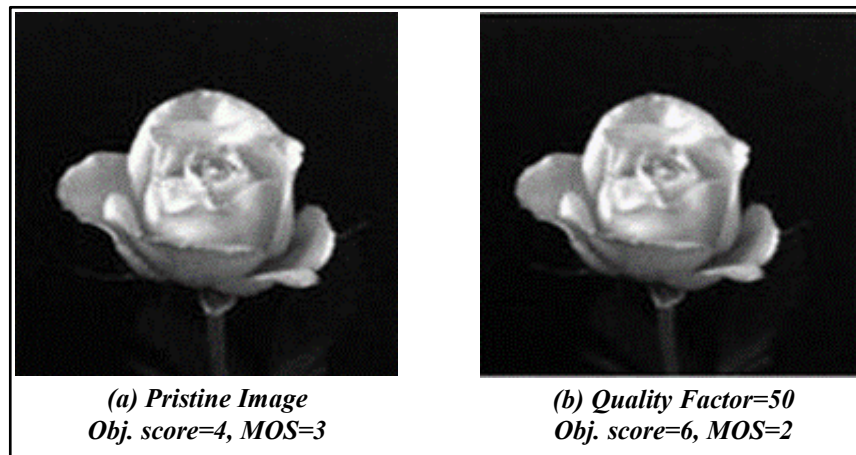
$$S = \alpha + \beta bl^{\gamma_1} avg^{\gamma_2} zc^{\gamma_3} \quad (4.8)$$

where $\alpha = -245.9$, $\beta = 261.9$, $\gamma_1 = -0.0240$, $\gamma_2 = 0.0160$, and $\gamma_3 = 0.0064$ are model parameters

The study of the NR metric shows that unlike the subjective evaluation, the objective scores computed by the above mathematical model are higher for the distorted images than the pristine images as depicted in figure 4.1.



(a) *Tiger.jpg*



(b) *Rose.jpg*

Figure 4.1: Examples of higher objective score computed by a mathematical model for distorted images than the subjective score of pristine images.

The study explained that the model parameters required for the mathematical model to predict the image quality are estimated using subjective scores as a training dataset. However, these parameters are same for all the test images. Therefore, this mathematical model cannot always establish the correlation between extracted features, subjective score, and predicted score, which may lead to the incorrect prediction. The relationship between subjective score and extracted features is

nonlinear in nature and composite for computation. The machine learning approach is proposed to address this problem. The ANN is employed for the quality prediction. It is trained on the features extracted from training dataset. Preparation of training and testing image datasets for this experimentation is discussed in chapter 3, section 3.4.1. Figure 4.2 depicts the schematic representation of the methodology.

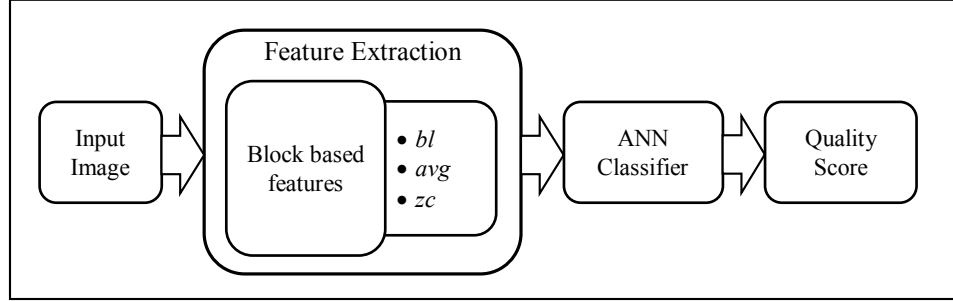


Figure 4.2: Schematic representation of methodology for existing block based feature metric [43] using ANN

The back propagation neural network is used for the experimentation to predict the quality score. The three-layered feed-forward ANN is employed for the experimentation using Levenberg-Marquardt training algorithm. The input layer of ANN has three neurons in input layer; nine in hidden layer, and one in output layer. The performance of ANN for few sample images is given below in table 4.1. The table presents the subjective score (MOS) as a target score, output by the ANN, and error between target and obtained score. Regression plot for the subjective scores as a target against the objective scores as output is depicted in figure 4.3. Regression coefficient between MOS and the obtained score is 0.96. The results obtained from the experiment shows that the machine learning approach performed well in comparison to the mathematical model.

Table 4.1: Performance evaluation of block based feature metric [43]

Sr. No.	Image	Target (MOS)	Output (Score)	Error
1.	Airplane.jpg			
i.	Original image	9.12	9.46	-0.34
ii.	70 Quality Factor	8	8.28	-0.28
iii.	50 Quality Factor	7.98	8.11	-0.13
iv.	10 Quality Factor	3.45	2.97	0.48
v.	50 Blur Factor	9.01	9.54	-0.53

Sr. No.	Image	Target (MOS)	Output (Score)	Error
2.	Fishingboat.jpg			
i.	Original image	9.43	9.14	0.29
ii.	70 Quality Factor	8.57	8.66	-0.09
iii.	50 Quality Factor	7.83	7.77	0.06
iv.	10 Quality Factor	3.52	2.34	1.18
v.	50 Blur Factor	9.46	9.27	0.19
3.	Fruits.jpg			
i.	Original image	8.45	8.92	-0.47
ii.	70 Quality Factor	8.34	8.64	-0.30
iii.	50 Quality Factor	6.59	6.63	-0.04
iv.	10 Quality Factor	1.67	1.67	0.00
v.	50 Blur Factor	8.45	8.91	-0.46
4.	Tiffany.jpg			
i.	Original image	9.54	9.30	0.24
ii.	70 Quality Factor	8.77	8.63	0.14
iii.	50 Quality Factor	7.76	7.75	0.01
iv.	10 Quality Factor	3.03	2.24	0.79
v.	50 Blur Factor	9.45	9.25	0.20

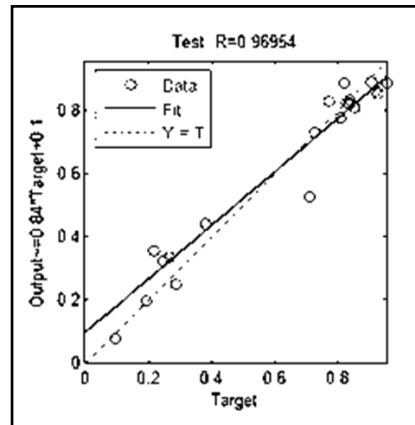


Figure 4.3: Regression plot of ANN using block based features [43]. MOS (target) are plotted against predicted scores (output)

4.1.1.2 Block based and Statistical Features

Compression process affects the brightness, contrast, local amplitude, and texture of an image which degrades the image quality by introducing blurring, ringing, and blocking artifacts into the image. Even though HVS cannot detect them, these artifacts are present in an image even if its quality is 70% or higher. Such

degradation to the image can be evaluated using the features in the frequency domain. Therefore, the block based features are combined with first and second order statistical features like kurtosis, skewness, SD, mean, entropy, and variance in frequency domain. The images from the created dataset are transformed into frequency domain by applying DCT transform before they are used for the feature extraction.

The back-propagation neural network is trained on this combined feature vector with MOS as a target [128]. The architecture of the ANN employed for this experimentation is similar to the one used in the previous experimentation except that it has nine neurons in the input layer. The block diagram of the methodology is described in figure 4.4.

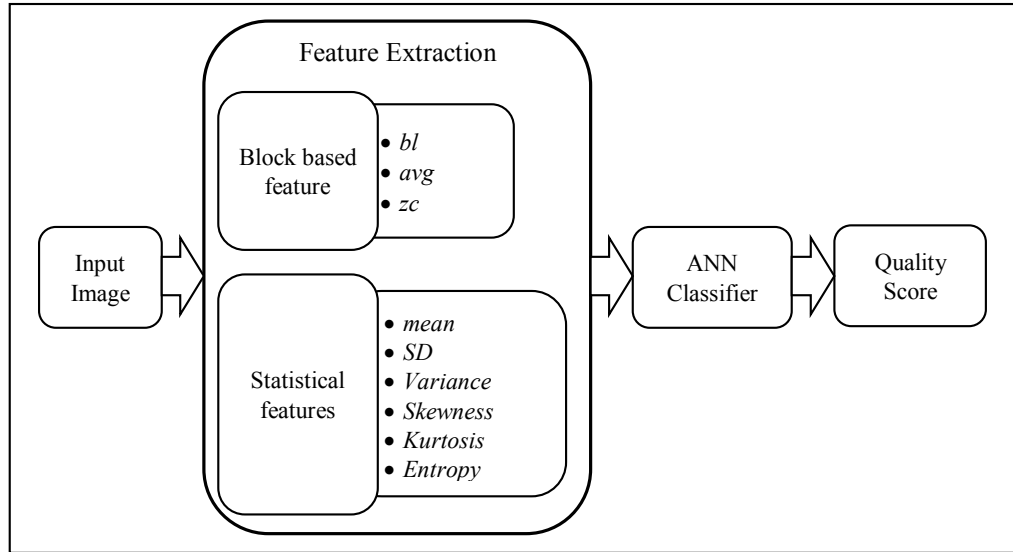


Figure 4.4: Schematic representation of methodology for a combination of block based features [43] and statistical features using ANN

The performance of the neural network for few of the test images is given in table 4.2. In the table, MOS is presented as a target score, output score by the ANN, and error between the target and obtained score. Regression plot for the target score (MOS) against the objective scores (output) is depicted in figure 4.5. Regression coefficient between MOS and the obtained score is 0.95. The result shows a good correlation between subjective and objective score. The regression coefficient calculated for this experiment is slightly lesser than the previous experiment. However, the fit of the regression plot is better.

Table 4.2: Performance evaluation of combination of block based [43] and statistical features

Sr. No.	Image	Target	Output (score)	Error
1.	Airplane.jpg			
i.	Original image	9.85	9.36	0.48
ii.	70 Quality Factor	8.77	8.48	0.29
iii.	50 Quality Factor	7.92	8.15	-0.22
iv.	10 Quality Factor	3.45	4.30	-0.75
v.	50 Blur Factor	7.47	8.32	-0.85
2.	Fishingboat.jpg			
i.	Original image	9.91	9.49	0.42
ii.	70 Quality Factor	8.18	8.21	-0.03
iii.	50 Quality Factor	6.92	7.46	-0.54
iv.	10 Quality Factor	1.91	3.01	-1.15
v.	50 Blur Factor	8.23	7.91	0.91
3.	Fruits.jpg			
i.	Original image	8.85	8.12	0.72
ii.	70 Quality Factor	8.96	8.18	0.81
iii.	50 Quality Factor	8.11	7.12	0.93
iv.	10 Quality Factor	3.87	2.67	1.20
v.	50 Blur Factor	8.76	8.42	0.34
4.	Tiffany.jpg			
i.	Original image	9.87	8.79	1.07
ii.	70 Quality Factor	8.74	8.45	0.28
iii.	50 Quality Factor	7.90	7.95	-0.05
iv.	10 Quality Factor	4.04	3.96	0.07
v.	50 Blur Factor	9.17	8.99	0.18

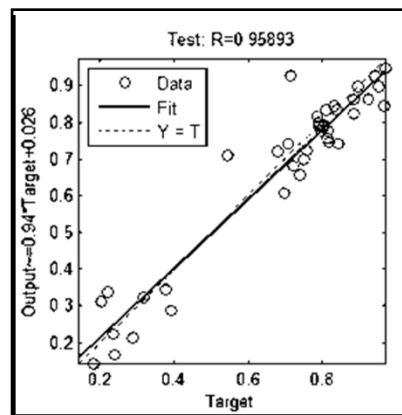


Figure 4.5: Regression plot of ANN using a combination of block based [43] and statistical features. MOS (target) are plotted against predicted scores (output)

Though the above metrics show good performance for JPEG and blurred images; there are many applications which introduce various other types of distortions. So, the development of the framework capable of quantifying different types of distortions becomes the necessity to provide the solution to this problem.

4.1.1.3 Combination of Block based, Statistical, and NSS based Features

Second ordered histogram features are obtained from co-occurrence metric. Frequency based features are also extracted [76]. The authors [76] pre-processed the images using contrast stretching before feature extraction. However, images are not pre-processed for the present experimentation. DCT based NSS features namely coefficients of frequency variation, energy sub-band ratio, and orientation are computed [103]. These features are combined with the block based features average difference across the block boundaries, the absolute difference between in-block image samples, and zero crossing rate, discussed in section 4.1.1.1. The statistical features and NSS based features are discussed below.

1. Statistical Features:

P. Gastaldo *et al.* [76] proposed a metric where features are computed as second-order histogram features obtained from co-occurrence metric and frequency based features derived from the energy of frequency content.

a. Second-order histogram features

$$co_autoc = \sum_{g_i, g_j} g_i g_j c_q(g_i, g_j, r, \omega) \quad (4.9)$$

$$co_invd = \sum_{g_i, g_j} \frac{c_q(g_i, g_j, r, \omega)}{1 + (g_i - g_j)^2} \quad (4.10)$$

$$co_cov = \sum_{g_i, g_j} (g_i - \mu_i)(g_j - \mu_j) c_q(g_i, g_j, r, \omega) \quad (4.11)$$

$$co_energy = \sum_{g_i, g_j} [c_q(g_i, g_j, r, \omega)]^2 \quad (4.12)$$

where $c_q(g_i, g_j, r, \omega)$ denotes the co-occurrence matrix associated with the probability distribution of pairs of pixels with gray levels g_i and g_j respectively, and are separated by r radial units at an angle ω to the horizontal axis.

b. Frequency based features

$$f_dcEn = B_p[0,0] / \sum_{x,y} B_p[x,y] \quad (4.13)$$

$$f_horEn = \sum_y B_p[0,y] / \sum_{x,y} B_p[x,y] \quad (4.14)$$

$$f_verEn = \sum_x B_p[x,0] / \sum_{x,y} B_p[x,y] \quad (4.15)$$

$$f_diagEn = \sum_{\substack{x,y \\ x=y}} B_p[x,y] / \sum_{x,y} B_p[x,y] \quad (4.16)$$

where $B_p[x, y]$ is the DCT component at the angular frequencies m, n .

2. Natural Scene Statistics based Features:

M. Saad *et al.* [103] proposed DCT based NSS features. These DCT based features are computed block wise. The image is divided into 5×5 blocks. Features are then combined by averaging the values.

a. Coefficient of Frequency Variation

Let Y be a random variable which describes the DCT coefficients. The coefficient of frequency variation feature is

$$\psi = \frac{\sigma|Y|}{\mu|Y|} \quad (4.17)$$

where μ and σ represent mean and standard deviation respectively.

b. Energy Sub-band Ratio Measure

DCT coefficients are represented increasingly in higher radial spatial function from the top-left corner to bottom-right corner. Three energy bands are used for feature calculation. The Average Energy (AE) is square of deviation. In frequency band x , it is defined as

$$AE_x = \sigma_x \quad (4.18)$$

The ratio is defined as the difference between the average energy in frequency band x and average energy up to frequency band x to the sum of these measures.

$$RE_x = \frac{\left| AE_x - \frac{1}{n-1} \sum_{i \in x} AE_x \right|}{AE_x + \frac{1}{x-1} \sum_{i \in x} AE_x} \quad (4.19)$$

c. Orientation

As an effect of distortion, local orientation energy changes abnormally. This directional information is represented as block DCT coefficients along with three orientations.

The training and testing image datasets prepared from standard image datasets LIVE, TID2008, and TID2013 as discussed in chapter 3, section 3.4.6 are used for the experimentation.

A combined feature vector is prepared from the features extracted from the training dataset. The network is trained on these feature. The methodology for this experimentation is shown in figure 4.6.

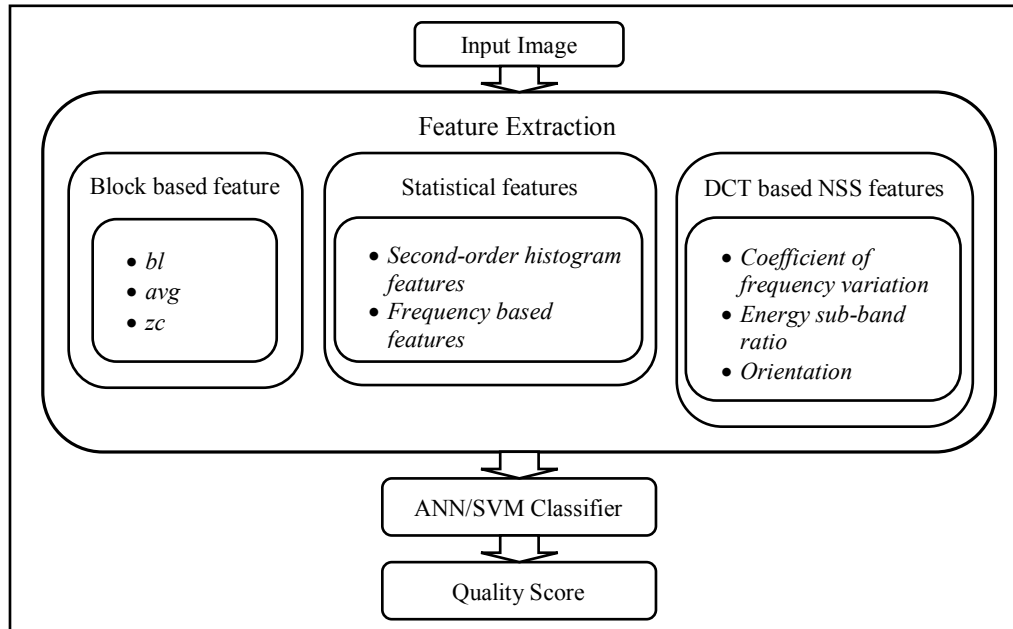


Figure 4.6: Schematic representation of methodology for combination of the block based [43], statistical [76], and DCT based NSS feature [103] using ANN and SVM

The analysis shows that these features demonstrate nonlinear behaviour. Due to the nonlinear behaviour of the features, machine learning is considered a better solution for quality prediction. ANN and SVM classifiers are used for quality prediction. The ANN employed for this experimentation has sixteen neurons in its input layer.

The architecture of SVM is implemented with 16 input nodes, a hidden layer with radial basis kernel function, and 4 output nodes. ANN and SVM, being the supervised classifiers need targets in the training phase. DMOS are used as targets for LIVE, whereas MOS are used as targets for TID2008 and TID2013. Unlike ANN, SVM requires classes as targets. Therefore, classes are created by distributing the subjective scores in separate groups. Each class represents a different range of subjective scores as shown in table 4.3.

Table 4.3: DMOS and MOS values as SVM classes

SVM	DMOS Values (LIVE)	MOS Values (TIDs)
Class 1	0 to 25	$10 \geq n \geq 8$
Class 2	26 to 50	$8 > n \geq 6$
Class 3	51 to 75	$6 > n \geq 3$
Class 4	76 and more	$3 \geq n$

Table 4.4(a)-(c) depict the performance comparison of the classifiers for LIVE, TID2008, and TID2013 image datasets respectively measured using Linear Correlation Coefficient (LCC) and SROCC. This combined metric shows good performance independent of the datasets. However, it underperforms for fast fading as shown in table 4.4(a).

ANN shows an edge in the performance over SVM. The regression plots of ANN using LIVE, TID2008, and TID2013 datasets are depicted in figure 4.7. The plots are presented for the training and testing for the respective image datasets. The regression plots show good correlation between subjective and objective scores for all the datasets.

Table 4.4(a): Performance evaluation of combined feature metric using ANN and SVM on LIVE dataset

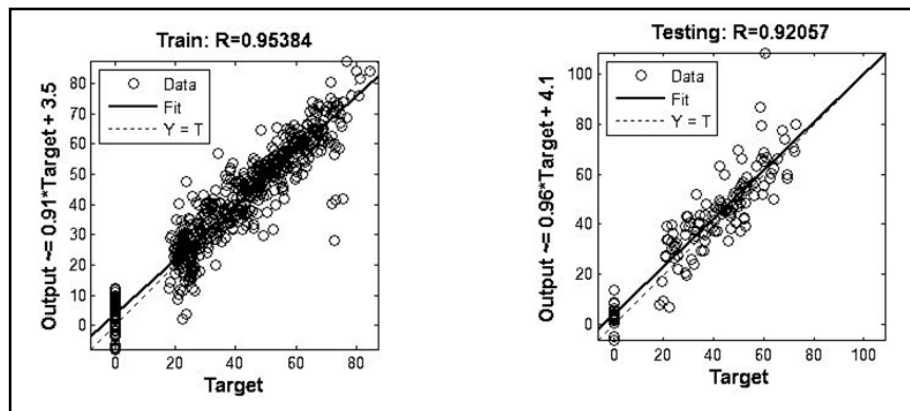
Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.94	0.94	0.92
JPEG2k	0.91	0.92	0.91
Gaussian blur	0.97	0.95	0.93
White Noise	0.97	0.97	0.95
Fast Fading	0.88	0.89	0.87
Overall	0.92	0.91	0.89

Table 4.4(b): Performance evaluation of combined feature metric using ANN and SVM on TID2008 dataset

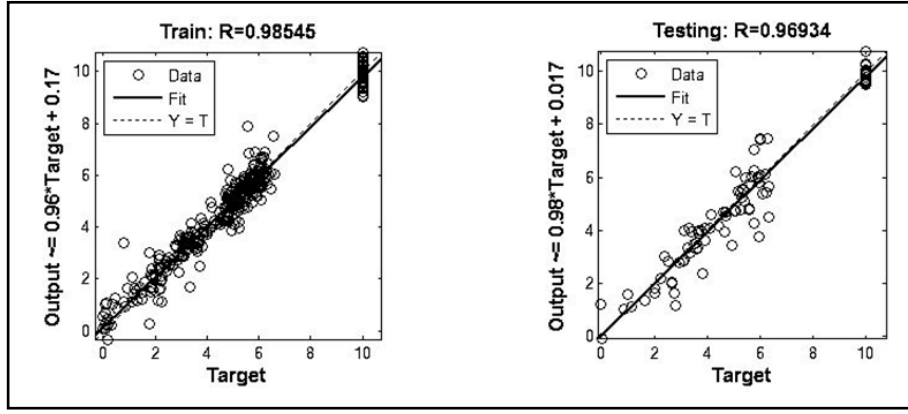
Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.96	0.97	0.94
JPEG2k	0.91	0.87	0.92
Gaussian blur	0.94	0.92	0.90
White Noise	0.96	0.94	0.95
Overall	0.96	0.94	0.90

Table 4.4(c): Performance evaluation of combined feature metric using ANN and SVM on TID2013 dataset

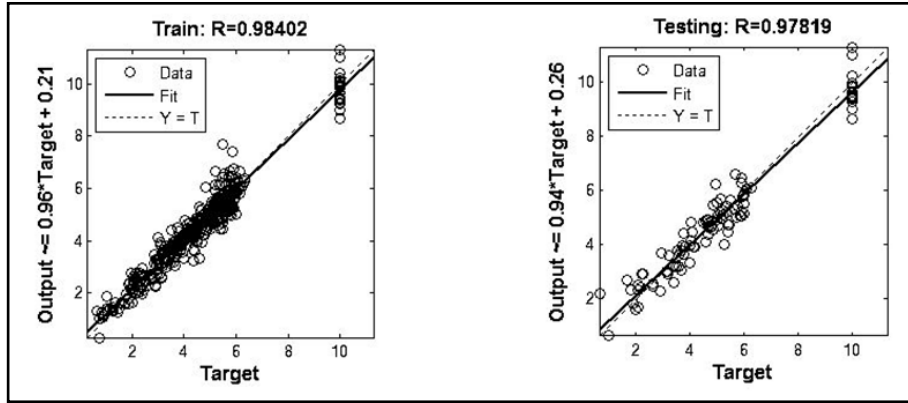
Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.98	0.96	0.94
JPEG2k	0.96	0.93	0.92
Gaussian blur	0.96	0.94	0.90
White Noise	0.96	0.97	0.95
Overall	0.97	0.95	0.90



(a) LIVE: DMOS (target) vs predicted score (output)



(b) TID2008: MOS (target) vs predicted score (output)



(c) TID2013: MOS (target) vs predicted score (output)

Figure 4.7: Regression plots of ANN for training and testing datasets using combined feature metric

The average performance of the proposed metric for the fast fading distortion suggests that an efficient classifier is required to be developed. Also, more features relevant to the fast fading distortion are required to be extracted to achieve better accuracy.

The performance of the proposed scheme is benchmarked with the performance of the original individual schemes, and shown in table 4.5. Results clearly show that the framework outperforms the original schemes.

Table 4.5: Comparison of the proposed metric with state-of-the-art metrics for LIVE dataset

Scheme	LCC
Proposed	0.92
Block based [43]	0.75
CBPN [76]	0.86
BLIINDS-II [103]	0.82

4.2 Design of Textural Feature Metric

The texture is an important property of an image which differentiates one image from another [131, 132]. In computer vision, it is defined as a spatial interrelationship between neighbouring pixels which illustrates the change in the intensity values or gray levels of neighbouring pixels.

The texture information in an image is represented in the high-frequency domain. Distortion affects the texture details in the image. Loss of the texture details ultimately degrades the image quality as shown in figure 4.8. The original image along with JPEG compressed images with quality factor 50 and 10, and blurred images with 50% blur are shown in the figure. Texture details in the image are affected badly with the increase in compression factor as marked by the red circle in the figure. The similar effect is also observed in the case of blurred images.

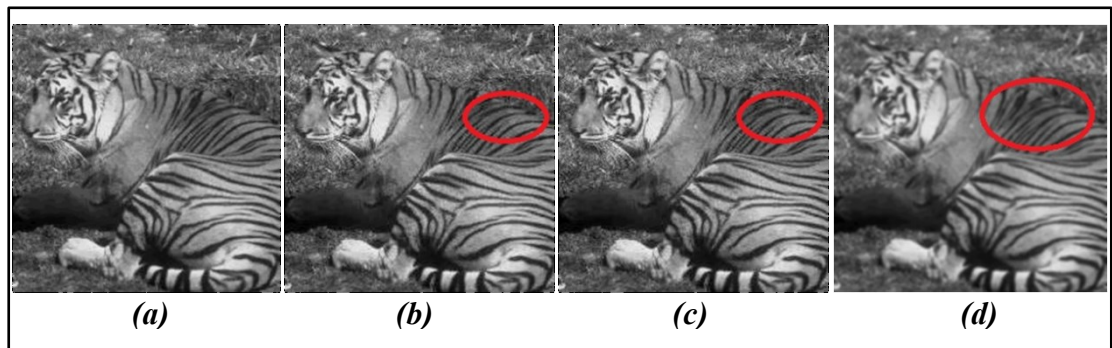


Figure 4.8: Loss of texture details in “Tiger.bmp”. (a)Pristine image, (b)compressed with quality factor 50, (c)quality factor 10, and (d)50% blur variants.

The image capturing devices such as point-and-shoot and mobile phone cameras use small image sensors. Edge sharpening algorithm and low pass filtering applied to reduce the noise during image acquisition process in such imaging devices result in loss of texture details. Sharpened window and shingles in the image are marked with a blue circle as shown in figure 4.9. On the other hand, texture in pine shrub in the same image is lost which is marked with red circle in the figure. The blur and noise affect the intensity values in an image and spatial relationship between neighbouring pixels which results in loss of texture details and leads to the degradation in image quality. Human visual system perceives this quality degradation effortlessly. Hence the need of texture as a feature to quantify degradation is justified.



Figure 4.9: Sample image captured by phone camera

4.2.1 Texture Analysis

Texture analysis is the process of measuring image texture using local or global qualities such as density, contrast, entropy, correlation, linearity, energy, etc. Methods for texture analysis are classified as structural, statistical, and model based.

In structural technique, the texture is represented as definite primitives called as micro texture, and its hierarchal structure for spatial arrangements is called as “macro texture”. Definition of primitives and positional rules are must be defined. Selection of primitives, as well as the probability of the primitive that is to be located at a particular position, is defined as the function of position or the neighbouring primitives to the position. The structural method represents a good figurative description, and it is mostly used for synthesis instead of analysis [133].

In the statistical approach, the texture is represented as distribution and relationship between gray levels present in an image. Discrimination rate is very high for the second order moments. The widely accepted second order statistic characteristics are obtained using co-occurrence matrix [133, 134].

Model-based approach determines the texture using stochastic and fractal models. These models try to interpret texture with stochastic and generative image models. The transform techniques like Fourier, Gabor, and Wavelet transforms describe an image in space. Characteristics of texture like size and frequency are directly related to the analysis of the coordinate system of space.

In the study, the statistical approach is used for the texture analysis. Researchers proposed first order and second order statistical parameters as textural features. First order statistical parameters such as kurtosis, standard deviation, mean, and skewness

evaluate characteristics of an individual pixel. However, it does not calculate spatial relationship with other pixels. One of the second order statistical features is extracted using Gray Level Co-occurrence Matrix (GLCM). GLCM evaluates spatial relationship between neighbouring pixels and extracts features such as contrast, homogeneity, correlation, entropy, etc. Although GLCM based features are widely accepted for texture analysis, they do not work well for isotropic images due to their heavy dependence on the directional information.

4.2.2 Feature Extraction

The texture energy filters use frequency or spatial filters to analyze the texture of an image and hence, are free from dependency on the directional information. It also defines the spatial relationship between many pixels. Laws' texture energy filters identify texture properties such as coarseness, density, uniformity, and roughness [135]. GLCM depends heavily on the magnitude of gray level differences but fall short to detect noise and change in the luminance. LBP can be used to quantify noise in an image. LBP codes are rotation invariant and find the relation between the pixels P on a circle with radius R around the central pixel. Hence, Laws' texture energy filters and LBP are implemented as textural features.

4.2.2.1 Laws' Energy Measures

The tiny convolution kernel is applied to an image to design Laws' energy measures. The textures are differentiated using two dimensional convolutional kernels which are calculated from the following 1-D kernels –

$$L_5 = [1 \ 4 \ 6 \ 4 \ 1]$$

$$E_5 = [-1 \ -2 \ 0 \ 2 \ 1]$$

$$S_5 = [-1 \ 0 \ 2 \ 0 \ -1]$$

$$R_5 = [1 \ -4 \ 6 \ -4 \ 1]$$

where **L**evels (L), **E**dge (E), **S**pot (S) and **R**ipple (R) represent average gray level, extracted edge, extracted spots and extracted ripple respectively.

Two-dimensional kernels are obtained using convolution of one dimensional vertical and horizontal kernel. For example, L_5E_5 kernel is computed as the convolution of vertical L_5 with horizontal E_5 and presented in terms of matrix multiplication as

$L_5 E_5 = L_5^T E_5$. The resulting kernel is convolved with an image as shown in the following figure 4.10.

For the texture analysis of sample image of size $M \times N$, each of 2D kernels of size 5×5 is applied to the image which results in a set of gray images. The dimension of each image is-

$$N - window_{size} + 1 \times M - window_{size} + 1$$

where window size is the size of the kernel, i.e. 5.

The outputs of the convolution form the basis for the texture analysis. The windowing operation is performed as shown in figure 4.11, where for every pixel in an image, different size window is applied around the pixel, and three descriptors are computed as follows:

Mean of the neighbour pixels,

$$\mu = \frac{\sum_s Neighbour_pixs}{S} \quad (4.20)$$

Mean of the absolute value of neighbour pixels,

$$abs_ \mu = \frac{\sum_s abs(Neighbour_pixs)}{S} \quad (4.21)$$

Standard Deviation of neighbour pixels,

$$\sigma = \sqrt{\frac{\sum_s (Neighbour_pixs - \mu)^2}{S}} \quad (4.22)$$

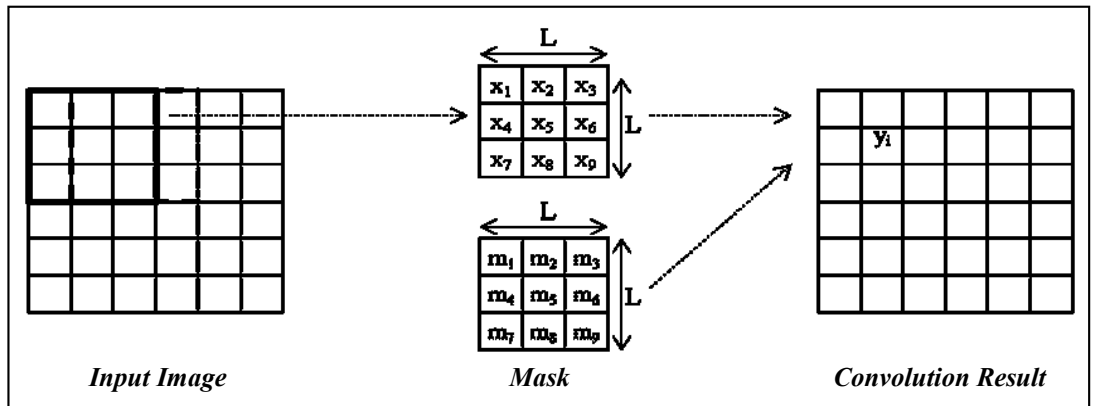


Figure 4.10: Mask Convolution, Pixels = N = L × L, $y_i = \sum_{i=1}^N x_i \times m_i$

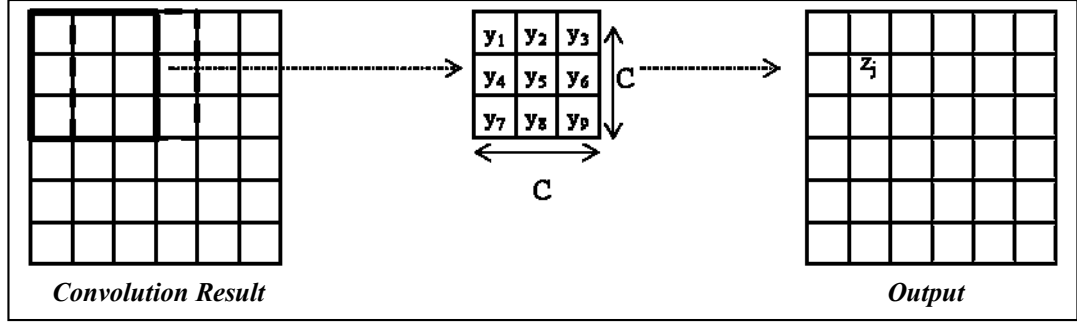


Figure 4.11: Statistic Evaluation, $\text{Pixels} = N = C \times C, z_i = f(y_1, y_2, \dots, y_n)$ where z_i denotes mean of neighbour pixels, mean of absolute values, and SD of neighbour pixels.

The windowing operation generates the new sets of images. By applying Min-max normalization technique, these images are normalized. Normalization process transforms N-dimensional image $I: \{Y \subseteq P^N\} \rightarrow \{\min, \dots, \max\}$ using intensity values in the range (min, max) to the new range (nmin, nmax), i.e. $I: \{Y \subseteq P^N\} \rightarrow \{n \min, \dots, n \max\}$. The complete procedure results into nine maps $L_5E_5, L_5S_5, L_5R_5, E_5E_5, E_5S_5, E_5R_5, S_5S_5, S_5R_5$ and R_5R_5 . L_5E_5 is applied to an image horizontally to detect edges and vertically to evaluate gray intensity. Other maps work in the similar manner [135].

Reference image “bikes.bmp” from LIVE dataset is depicted in figure 4.12 with its different distortions and respective subjective scores.

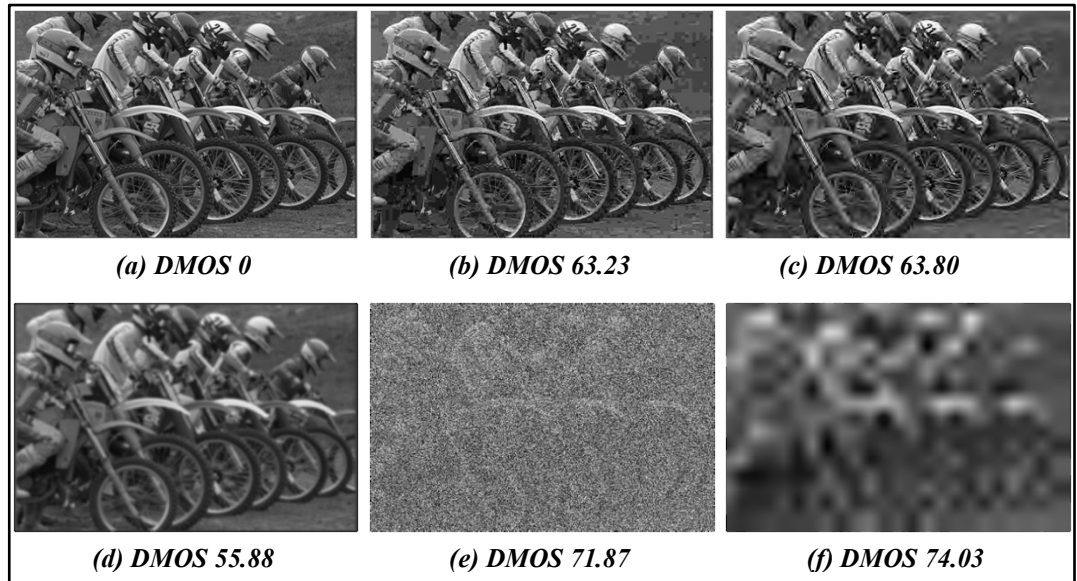
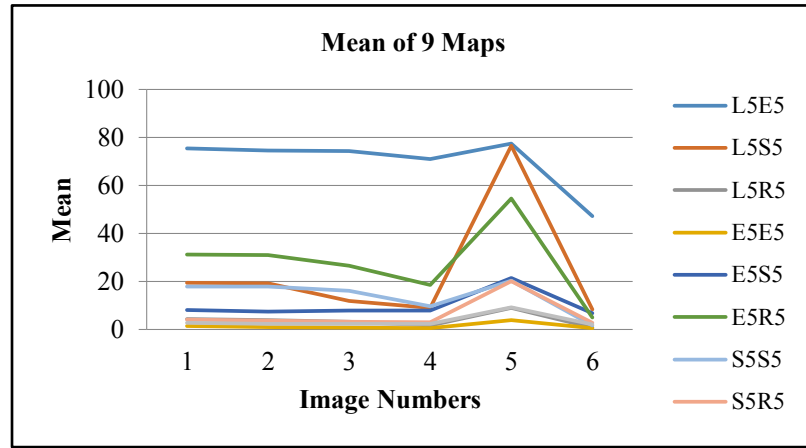
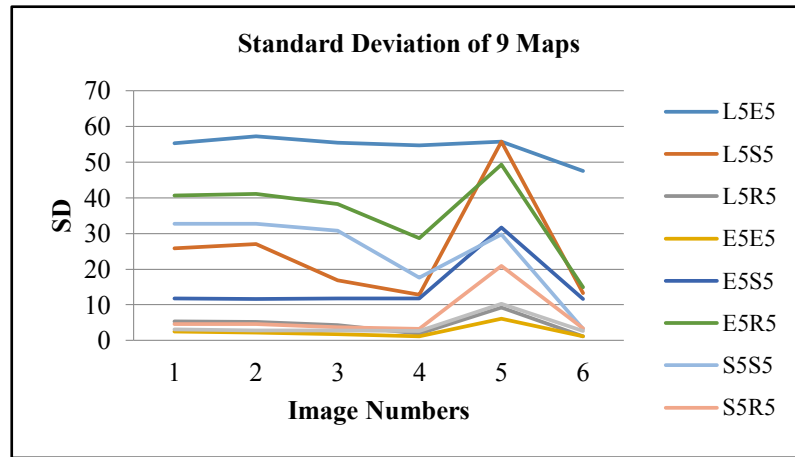


Figure 4.12: Reference image “bikes.bmp” from LIVE dataset and its different distortion variants - (a)Original, (b)JPEG, (c)JPEG2000, (d)Gblur, (e)WN, and (f) FF with their respective subjective scores

Mean and standard deviation values of nine energy maps for the original image “bikes.bmp” and its JPEG, JPEG2000, Gblur, WN, and FF variants are shown in following figure 4.13. From the figure, it is observed that the value of texture energy mask changes with the distortion. For example, the value of L_5S_5 is increased for white noise image due to the spots and loss of intensities. In another example, the value of L_5E_5 is decreased for Gaussian blur image because edges are smoothing down and intensity values are also decreased due to the blurring effect.



(a)



(b)

Figure 4.13: Plots showing variations for (a) mean and (b) standard deviation of nine energy maps for “bikes.bmp” and its distortion variants from LIVE dataset numbered as 1-original, 2-JPEG, 3-JPEG2000, 4- Gblur, 5-WN, and 6-FF

White noise is defined as a uniform distribution of uncorrelated random variables. Due to this nature of white noise, values show the high impact on all the maps. Loss of frequencies in the image due to the fast fading affects the texture and blurs the image. This blurring makes the image more difficult to identify. Hence the texture energy measure gives smaller feature values for all the nine maps.

4.2.2.2 Pattern Strength

Local binary pattern converts an image into a group or image of integer tags relating small-scale manifestation of an image. Subsequently, these tags or their statistics, predominantly the histogram, are applied for further image analysis. T. Ojala *et al.* [136] has defined the earlier version of LBP operator. The basis for this study is that the texture has two characteristics - pattern and its strength. This version works within a block of 3×3 pixels. Centre pixel value is deducted from 8 neighbouring pixels; the sign of the difference is considered for the thresholding. Summation of the difference is multiplied by powers of two to find out the centre pixel label. As it has 8 neighbourhood pixels, total $2^8 = 256$ different labels are obtained. The LBP is mathematically derived as follows:

A gray level image represented as $I_m(m, n)$. Let g_r indicates the gray level of a random pixel (m, n) , i.e. $g_r = I_m(m, n)$.

Further, let g_s be a gray value of sampling point in the evenly spaced circular vicinity of S sampling points and radius R around the point (m, n) .

$$g_s = I_m(m_s, n_s), \quad s = 0, \dots, S-1 \quad (4.23)$$

$$m_s = m + R \cos(2\pi s / S) \quad (4.24)$$

$$n_s = n - R \sin(2\pi s / S) \quad (4.25)$$

Suppose, the joint distribution of gray intensities of $S+1$ pixel is used to define the local texture of an image:

$$Z = z(g_r, g_0, g_1, \dots, g_{S-1}) \quad (4.26)$$

As a result of the exclusive loss of information, the value of the central pixel from the neighbourhood pixels can be deducted as,

$$Z = z(g_r, g_0 - g_r, g_1 - g_r, \dots, g_{S-1} - g_r) \quad (4.27)$$

Joint distribution is estimated assuming that the central pixel is statistically not dependant on the difference. So, the above equation is rewritten using factorization as,

$$Z \approx z(g_r) z(g_0 - g_r, g_1 - g_r, \dots, g_{S-1} - g_r) \quad (4.28)$$

The first factor $z(g_r)$ is defined as intensity allocation for an image $I_m(m, n)$. It does not hold useful information in the analysis of local textural pattern. Thus joint distribution becomes

$$z(g_0 - g_r, g_1 - g_r, \dots, g_{s-1} - g_r) \quad (4.29)$$

Further, the sign difference is computed for thresholding as

$$z(si(g_0 - g_r), si(g_1 - g_r), \dots, si(g_{s-1} - g_r)) \quad (4.30)$$

where $si(t)$ is the thresholding function.

$$si(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases} \quad (4.31)$$

The LBP operator uses this joint distribution. Initial LBP operator is computed by adding the threshold differences multiplying by powers of two. Therefore, $LBP_{S,R}$ is defined as

$$LBP_{S,R}(m_r, n_r) = \sum_{s=0}^{S-1} si(g_s - g_r) 2^s \quad (4.32)$$

The above discussed mathematical derivation of LBP, when applied to an image, generates a vector of 256 values. The dimension of this vector is reduced by calculating statistical features standard deviation, skewness, and kurtosis.

Let X be the vector, then skewness and kurtosis can be defined as

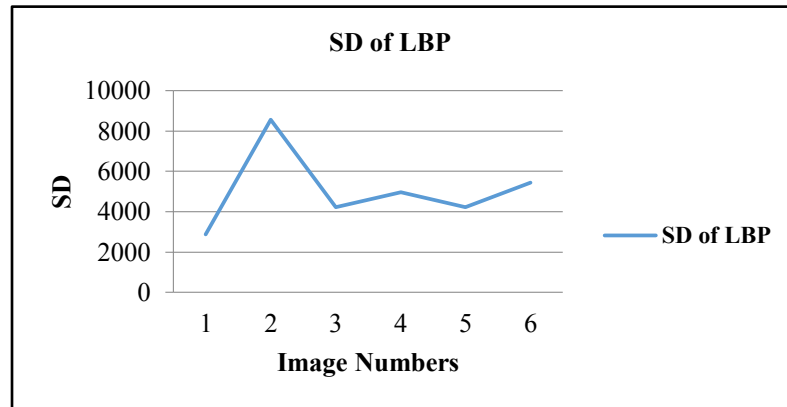
$$Skewness = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^3}{\sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}} \quad (4.33)$$

$$Kurtosis = \frac{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^4}{\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \right)^2} \quad (4.34)$$

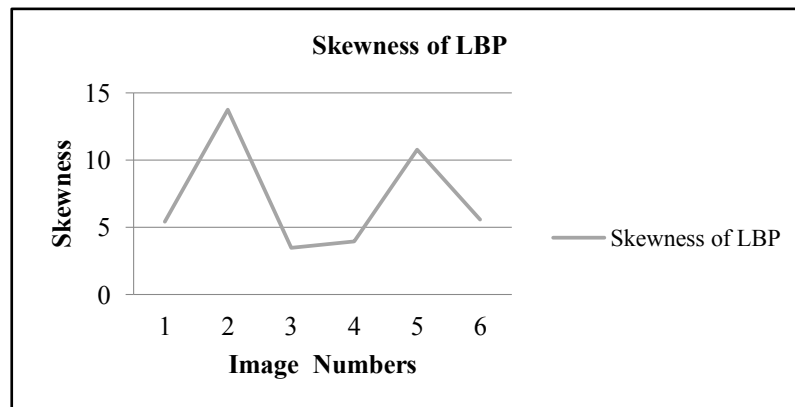
where $\bar{X}$ represents mean of sample points.

Plots of SD, skewness, and kurtosis of LBP for “bikes.bmp” and its distortion variants from LIVE dataset are shown below in figure 4.14. Plots show that type and level of degradations affect the values of statistical features. For example, pixels

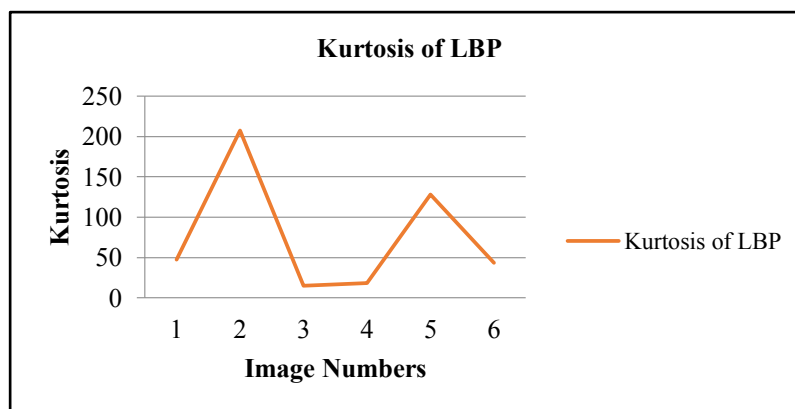
at the edges differ in intensity values from the neighbouring pixels due to effect of the blocking artifact. Therefore, JPEG image shows increase in the standard deviation.



(a) SD



(b) Skewness



(c) Kurtosis

Figure 4.14: Plots showing variations for (a)SD, (b)skewness, and (c)kurtosis of LBP for “bikes.bmp” and its distortion variants from LIVE dataset numbered as 1-original, 2-JPEG, 3-JPEG2000, 4-Gblur, 5-WN, and 6-FF

4.2.3 Methodology and Results

The feature vector is designed by evaluating statistical features mean and SD of Laws' of nine energy maps, which resulted in 18 values. The output of LBP operator is a vector of size 256. A feature vector of size 3 is obtained from this vector by calculating the statistical values SD, skewness, and kurtosis. Local texture based statistical features SD and entropy are also computed and finally feature vector of size 23 is formed. Local texture features are obtained by dividing an image into 9×9 overlapping blocks, and by computing SD and entropy for each block. The global descriptor is computed by averaging the local descriptors.

The nature of the features extracted from energy maps and LBP vectors is nonlinear which can be seen from figure 4.13 and figure 4.14. Since laws' energy measures and local binary pattern exhibits nonlinearity, it is difficult to use different fitness functions to describe the nonlinear relationship between these features and expected output (MOS/DMOS). Therefore, the machine learning approach is used for quality prediction. Schematic representation of methodology is depicted in the following figure 4.15.

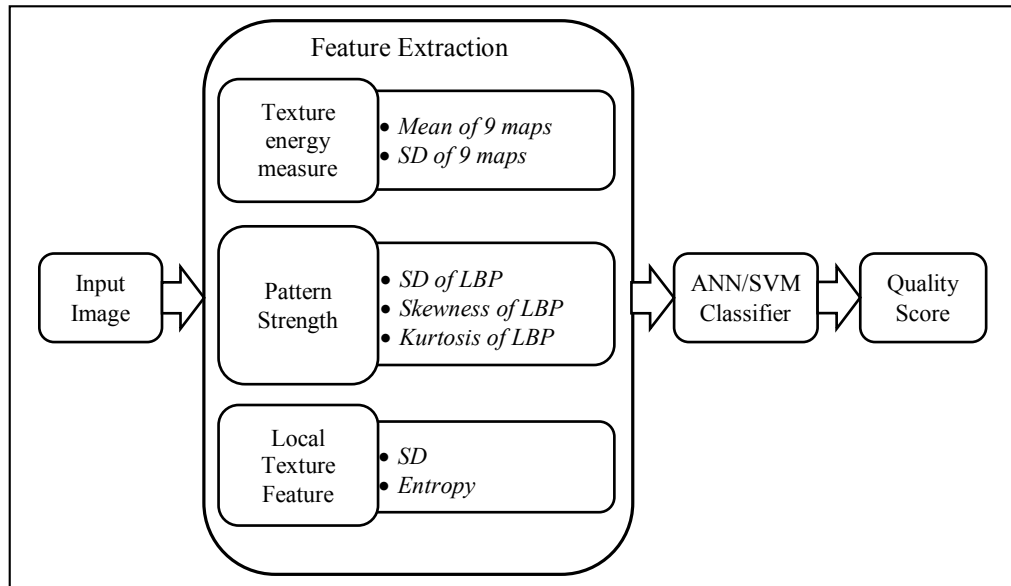


Figure 4.15: Schematic representation of methodology for textural feature metric

The results are tested with feedforward neural network using back propagation training algorithm and SVM. Model is tested on standard datasets LIVE, TID2008, and TID2013. Three layered feedforward neural network with Levenberg-Marquardt training algorithm is used for the classification. The architecture of ANN

has 23 input neurons, 9 hidden layer neurons and one output neuron. The log-sigmoid transfer function is used in hidden layer neurons, and the linear transfer function is used for output layer neuron. Training and testing datasets discussed in chapter 3, section 3.4.6 are used for the experimentation. ANN is trained on the features extracted from the training images and predicts the quality of test images. DMOS for LIVE, and MOS for TID2008 and TID2013 are given as targets to the network. LCC and SROCC between predicted score and the subjective score are presented in following tables 4.6(a)-(c) for LIVE, TID2008, and TID2013 datasets respectively. Regression plots of ANN using the textural feature metric for LIVE, TID2008, and TID2013 datasets are shown in figure 4.16. The plots are presented for both training and testing for the respective image datasets. These plots show good correlation between subjective scores and predicted scores independent of databases.

Performance of the feature metric is also evaluated with SVM classifier. It requires classes as a target. Hence, the classes are created from the subjective score as described in the above table 4.3. DMOS and MOS values corresponding to the images from LIVE and TIDs training datasets respectively are grouped into different classes and given as target to SVM. SVM classifier is trained on the features extracted from the training images. Trained classifier predicts the classes for test images. Results obtained from the SVM are presented in tables 4.6(a)-(c) for LIVE, TID2008, and TID2013 datasets respectively. It is observed that the proposed metric works well independent of the dataset. This texture based metric underperforms for JPEG and JPEG2000 because of the blocking artifact. However, it works well for Gaussian blur and white noise.

Table 4.6(a): Performance evaluation of textural feature metric using ANN and SVM on LIVE dataset

Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.88	0.87	0.85
JPEG2k	0.84	0.82	0.85
Gblur	0.96	0.95	0.95
WN	0.97	0.98	0.95
FF	0.90	0.88	0.80
Overall	0.89	0.90	0.90

Table 4.6(b): Performance evaluation of textural feature metric using ANN and SVM on TID2008 dataset

Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.87	0.88	0.85
JPEG2k	0.87	0.87	0.85
Gblur	0.94	0.91	0.92
WN	0.92	0.92	0.91
Overall	0.93	0.93	0.92

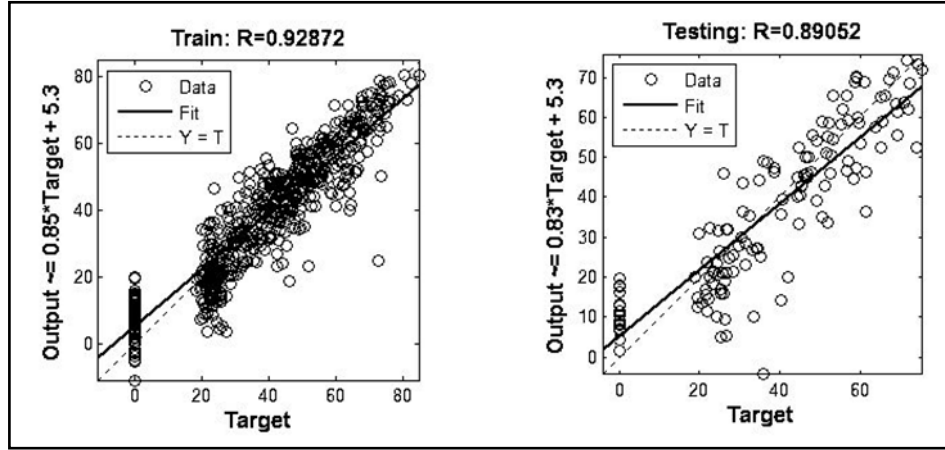
Table 4.6(c): Performance evaluation of textural feature metric using ANN and SVM on TID2013 dataset

Distortion Type	ANN		SVM
	LCC	SROCC	LCC
JPEG	0.88	0.87	0.85
JPEG2k	0.85	0.82	0.85
Gblur	0.93	0.94	0.91
WN	0.94	0.96	0.92
Overall	0.91	0.90	0.90

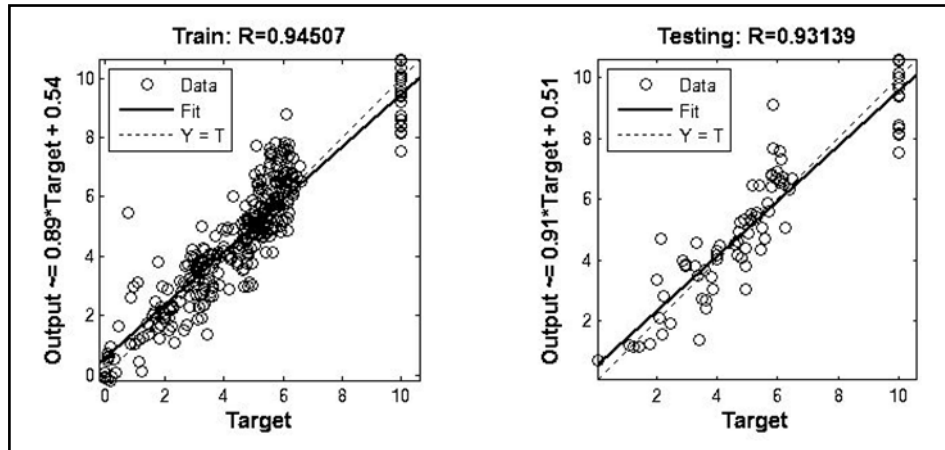
Comparison of the proposed metric with state-of-the-art NR metrics for LIVE dataset is shown in table 4.7. Performance of the proposed metric is comparable with the performance of state-of-the-art NR metrics.

Table 4.7: Performance comparison between proposed metric and state-of-the-art NR metrics on LIVE dataset

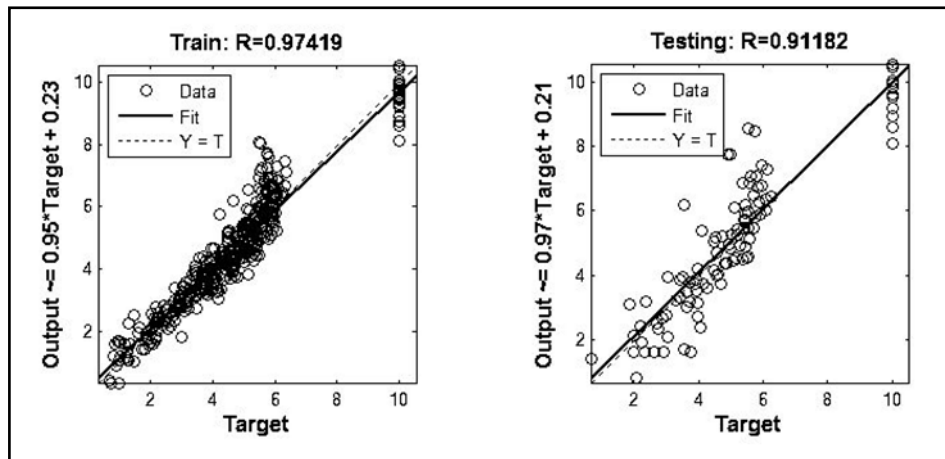
Sr. No.	Metrics	LIVE	
		LCC	SROCC
1.	DCT-based(NR-IQA) [19]	0.79	-
2.	BIQI [97]	0.82	-
3.	RCGA-ELM [74]	0.70	-
4.	CBPN [76]	0.86	-
5.	Local Ternary pattern [86]	0.94	0.94
6	Spatial & Spectral Entropies [80]	0.93	0.93
8	Texture Information bank [84]	0.97	0.96
10	Proposed metric	0.89	0.90



(a) *LIVE: DMOS as target vs predicted score as output*



(b) *TID2008: MOS as target vs predicted score as output*



(c) *TID2013: MOS as target vs predicted score as output*

Figure 4.16: Regression plots of ANN for training and testing datasets using textural feature metric

Regression plots of ANN for LIVE, TID2008 and TID2013 shown in figure 4.16 exhibit a good correlation between subjective score and predicted score with $R=0.89$,

$R=0.93$, and $R=0.91$ respectively. Results indicate that the proposed metric effectively measures the distortions for overall quality prediction.

4.3 Summary

The features average absolute difference between in-block image sample and zero-crossing rate used in the existing block based NR metric [43] are implemented and studied. The quality is predicted using a nonlinear regression model. The model uses parameters estimated from the subjective scores as training data. These model parameters are same for all test images. Also, the nature of the relationship between subjective score and extracted features is nonlinear and composite for the computation. The mathematical model thus cannot always establish this correlation and leads to the incorrect quality prediction. Hence, the machine learning approach is proposed for quality prediction. The experimentation is further extended by combining blocking parameters with first and second order statistical parameters kurtosis, skewness, SD, mean, entropy, and variance in frequency domain.

ANN is used in both the experiments for quality prediction. Both the experiments exhibit good performance for JPEG and blurred images. The proposed framework is expected to work with other distortion types also. Therefore, a combination of different features is proposed to extend the experimentation further. Accordingly, block based, statistical, and DCT based NSS features from the existing metrics are combined and tested. Quality is predicted by ANN and SVM to evaluate the performance of the combined metric. This framework shows good results except for the fast fading images. To address the performance issue for fast fading images, more features sensitive to the fast fading are required to be combined or added to the existing features set.

The texture is one of the important properties of an image. The texture energy measures evaluate energy present in the image. Whereas pattern strength addresses blockiness and noise. Therefore, textural feature metric is implemented by combining both these features. The ANN and SVM are employed for quality prediction. This metric performs well for white noise and Gblur, whereas its performance is slightly lower for fast fading. It also has been observed that the metric underperformed for JPEG and JPEG2000. Therefore, it is proposed to implement the block based features in further experimentations.

Chapter 5.

Design of NRIQA Metrics using Block based and Shape Adaptive Discrete Wavelet Features

JPEG is one of the most common distortion acquired by the images during the image compression process. JPEG images suffer from blocking artifact due to the block based coding method used in JPEG compression technique. JPEG coder first divides the source image into the blocks of 8×8 pixels and then transforms these blocks by applying DCT from left to right and top to bottom. The transformed blocks are quantized blockwise by discarding the less significant frequencies. As a result of blockwise quantization, coefficients of the adjacent blocks are quantized differently causing the discontinuities at the block boundaries. It ultimately introduces the blocking artifact in the regenerated JPEG image [137, 138]. The edge sensitive features can reflect these changes at the block boundaries in the image. The NRIQA methods can utilize this information for the quality prediction. Therefore, an NRIQA metric is proposed to evaluate blocking artifact using edge sensitive features namely Image Gradient (IG), Spatial Frequency Measure (SFM), and Spectral Activity Measure (SAM).

Images are analysed globally using certain features, or local features are extracted and pooled as a global descriptor. Due to the block based compression, edges of the blocks show discontinuity. Such degradation hampers the contours of an object present in an image, and ultimately changes the shape of the object. The change in the contour, if identified, can help in evaluating the degradation to the image. Shape Adaptive Discrete Wavelet Transform (SADWT) can be employed for identification of the shapes. SADWT is categorized as 1D, 2D and 3D [139, 140]. Since an image is represented in 2D, therefore 2D-SADWT is proposed for the present study.

Being one of the objectives of the present study, the neurofuzzy classifier is implemented for the experimentation. ANN is employed for the evaluation of block based feature and shape adaptive discrete wavelet feature as well. Though it worked well for block based feature, it failed to estimate the image quality using SADWT feature. NWT is thus explored for the experimentation.

This chapter presents the implementation of ANN, NFC, and NWT using the above discussed two metrics.

The chapter is organized into three sections. The first section 5.1 presents the implementation of block based features IG, SFM, and SAM. In subsection 5.1.1, implementation of NFC is discussed. Section 5.1.2 describes the methodology and discusses the results. Implementation of SADWT features is described in section 5.2. Design of NWT is discussed in section 5.2.2 and results are presented in section 5.2.3. Section 5.3 summarizes the study presented in this chapter.

5.1 Design of Block based Feature Metric

HVS uses edge information to acquire the knowledge of the shapes. Edges suffer heavily due to the blockiness in the JPEG distorted images. The image gradient feature provides quantifiable information about the edges and structures in an image [141]. The distortion in an image affects the image gradient magnitude. It makes image gradient feature a useful candidate for developing the NR metric for JPEG distortion.

Image gradient feature in the frequency domain is computed using the gradient of an image. The gradient of an image is a directional change in intensity values which are affected during the compression. Therefore, this feature is used to quantify the blockiness. The success of the approach proposed by W. Xue *et al.* [142], who employed statistics of the gradient magnitude and the Laplacian to predict image quality, encouraged to consider the image gradient feature for the experimentation.

An image with lots of details exhibits high SFM value. It signifies that the image has many components lie in the high-frequency region. Further, the small value of SAM signifies low predictability. Distortions introduced in the image changes these feature values [143].

1. Image Gradient (IG):

The image gradient is calculated as

$$\psi g = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{bmatrix} \quad (5.1)$$

where $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$ represent gradients in x-direction and y-direction respectively.

$$IG = 1.4826 \times MAD(\psi g) \quad (5.2)$$

$$IG = 1.4826 \times median_g[\|\psi g - median_f(\|\psi g\|)\|] \quad (5.3)$$

where MAD represents mean absolute deviation. Mean absolute deviation of zero-mean of normal distribution with unit variance is $0.6745 = 1/1.4826$. Hence is the constant 1.4826 in equation 5.2 and 5.3 [144].

2. Spatial Frequency Measure (SFM):

It is a measure of overall activity level in the digital image.

$$SFM = \sqrt{R^2 + C^2} \quad (5.4)$$

where row frequency, $R = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=2}^N (g(m, n) - g(m, n-1))^2}$

and column frequency, $C = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=2}^N (g(m, n) - g(m-1, n))^2}$

3. Spectral Activity Measure (SAM):

It is a measure of image predictability.

$$SAM = \frac{\frac{1}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} |F(j, k)|^2}{\left[\prod_{j=0}^{M-1} \prod_{k=0}^{N-1} |F(j, k)|^2 \right]^{\frac{1}{MN}}} \quad (5.5)$$

where $F(j, k)$ is Discrete Fourier Transform coefficients.

Changes in these block based feature values along with the subjective scores for the sample image “Plane.bmp” from LIVE dataset and its JPEG variants with different bit rates are depicted in following figure 5.1. It shows that with the increase in JPEG distortion, IG and SFM decrease whereas, SAM increases.

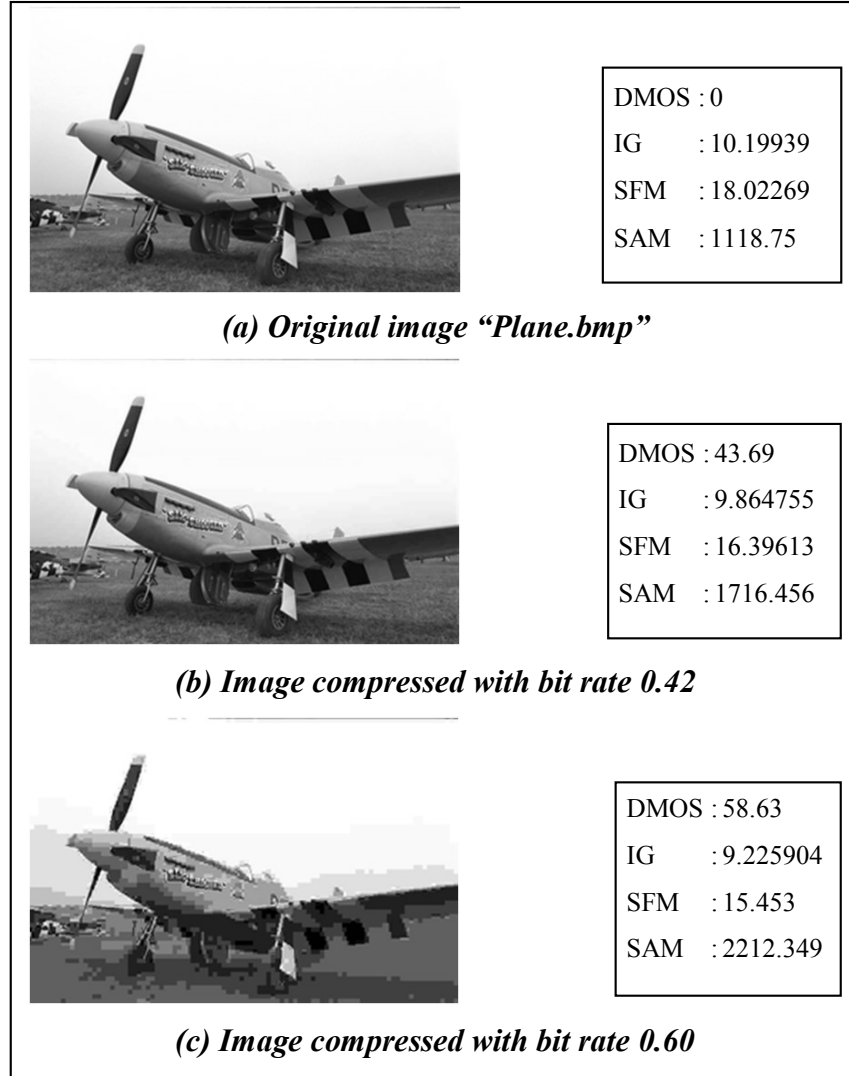


Figure 5.1: DMOS, IG, SFM, and SAM for reference image “Plane.bmp” and its compressed variants with different bit rates

The objective of the present study is to implement neurofuzzy classifier as a quality predictor for the framework. Therefore, neurofuzzy classifier is implemented and tested in the experimentation discussed in the subsequent section of this chapter.

5.1.1 Neurofuzzy Classifier

Adaptive learning, generalization capacity, and self organization are the advantages of ANN. However, understanding of the functionality of ANN is difficult as it hides the relation between inputs and outputs in the weights of the neurons behind its hidden layers [145, 146, 147]. On the other hand, the fuzzy system can represent knowledge in an interpretable manner but lacks the generalization capability. It also has difficulty in finding the appropriate membership values.

In the neurofuzzy system, fuzzy system and neural network are combined to develop an architecture which applies the fuzzy system to symbolize knowledge in an understandable way and learning capability of ANN to optimize its parameters.

Feature vector $y = \{y_1, y_2, y_3, \dots, y_N\}$ extracted from an image, is given as input to the NeuroFuzzy Classifier (NFC) as shown in figure 5.2. The classifier fuzzifies feature vector using Gaussian Membership Function (GMF) to its corresponding pattern values and assigns the degree of the memberships to each pattern to different classes.

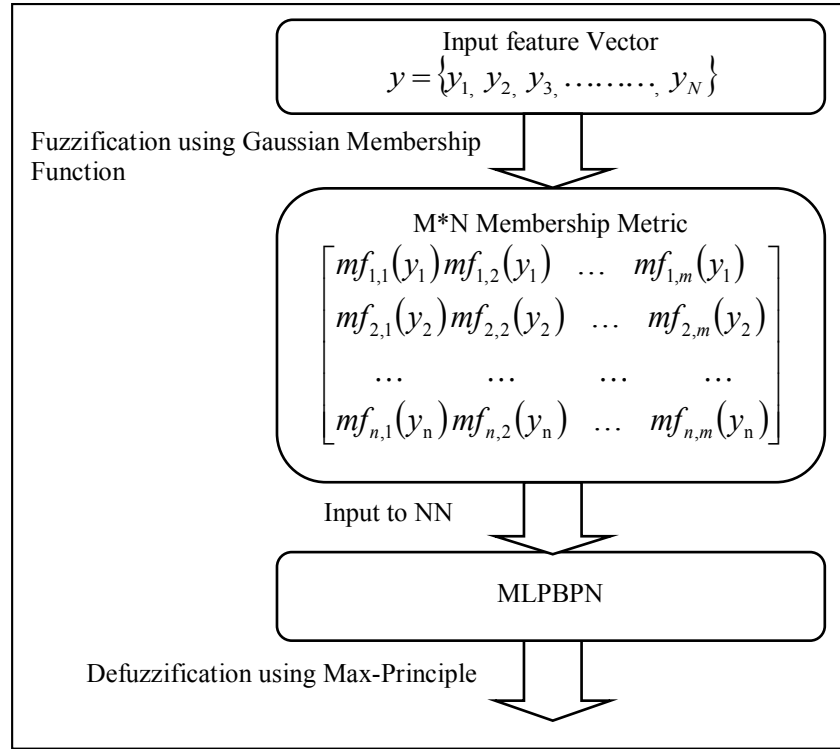


Figure 5.2: NeuroFuzzy Technique

The neurofuzzy classifier follows the three-step algorithm – 1. Fuzzification, 2. Classification, and 3. Defuzzification.

1. Fuzzification:

In Fuzzification, input vectors are used to translate into the membership values, and membership function fuzzifies the inputs. There are various membership functions such as bell-shaped Gaussian, trapezoidal, triangular, etc. In the present study, the Gaussian membership function is applied to produce a matrix of order $M \times N$, where M represents number of features and N indicates classes.

The Membership Function (MF) can be defined as,

$$mf_{i,j}(y_i) = \text{degree of membership of feature vector } i \text{ to class } j$$

where image feature $y = y_1, y_2, \dots, y_N$.

An input feature vector is fuzzified using Gaussian membership function which is represented as

$$mf(y : \sigma, c) = e^{\frac{-(y-c)^2}{2\sigma^2}} \quad (5.6)$$

where c indicates the centre and σ is the width of the MF.

A classic Gaussian membership function is depicted in figure 5.3. The classification model can be made flexible by obtaining different membership functions by changing the values of c and σ .

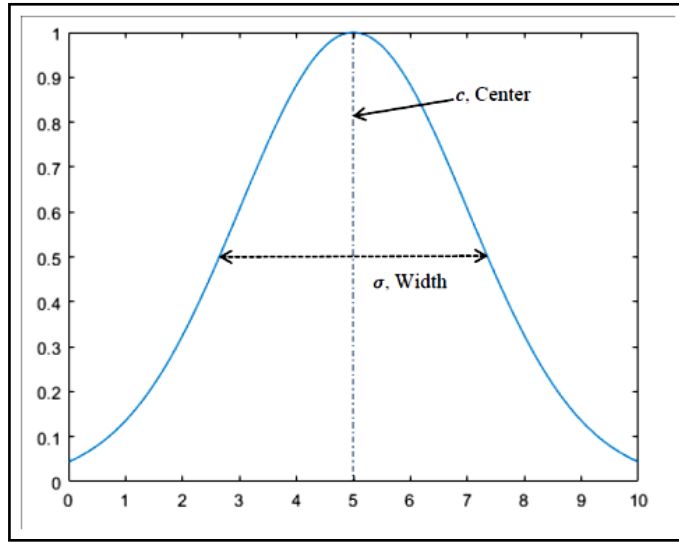


Figure 5.3: A typical Gaussian membership function

Membership matrix of feature vector y is calculated after fuzzification and can be represented as

$$MF(y) = \begin{bmatrix} mf_{1,1}(y_1) & mf_{1,2}(y_1) & \dots & mf_{1,M}(y_1) \\ mf_{2,1}(y_2) & mf_{2,2}(y_2) & \dots & mf_{2,M}(y_2) \\ \dots & \dots & \dots & \dots \\ mf_{N,1}(y_N) & mf_{N,2}(y_N) & \dots & mf_{N,M}(y_N) \end{bmatrix} \quad (5.7)$$

where $mf_{i,j}(y_i)$ represents the membership of i^{th} feature of input vector y .

2. Classification:

In the second step of neurofuzzy classifier, commonly used neural network “Multi-Layer Perceptron Back Propagation Network (MLPBPN)” is used. Training of MLPBPN is achieved through mapping of inputs to the expected output by applying nonlinear function. The mapping is done by transmitting inputs to output from one or more strongly joined hidden layers with computational nodes to the last output layer. A number of the hidden layers with their relative neurons, transfer function used in the hidden layer, and weights at each connection are the significant parameters which influence the performance of the neural network. Generally, MLPBPN comprises three layers namely input, output, and hidden layer. This network is feedforward type with back propagation training algorithm. Transpose of the membership matrix is given as an input to MLPBPN.

The net input to unit j of MLPBPN in the hidden or output layers is computed as a linear combination of inputs. The output of unit j is defined as:

$$Output_j = \frac{1}{1 + e^{-Net_j}} \quad (5.8)$$

where Net_j represents the net input of unit j .

Net input is computed as the addition of connection weights plus the output since the prior layer. The net input is defined as:

$$Net_j = \sum_i weight_{ij} \times output_i + bias_j \quad (5.9)$$

where $weight_{ij}$ represents the connection weight of the linkage between unit i in the prior layer to unit j , $Output_i$ represents the output of unit i from the prior layer, and $bias_j$ is the bias of the unit j .

The error is computed as the total sum of squared errors between the obtained output and the expected output as

$$E = \sum_j (target_j - output_j)^2 \quad (5.10)$$

To minimize overall error, the weights of the network are changed as,

$$weight_{update} = weight_{ij} - \eta \nabla weight_{ij} \quad (5.11)$$

Similarly, the bias is updated as,

$$bias_{update} = bias_j - \eta \nabla bias_j \quad (5.12)$$

3. Defuzzification:

The output obtained from the network is transposed and defuzzified using MAX technique. An input feature vector y is allocated to a specific class j through the maximum class membership value if and only if,

$$mf_j(y) \geq mf_i(y) \quad \forall i \in (1, \dots, M) \text{ and } i \neq j \quad (5.13)$$

where $mf_i(y)$ is an output value of i^{th} neuron in the output layer.

5.1.2 Methodology and Results

NRIQA metric is implemented using the features IG, SFM, and SAM to evaluate blocking artifact in JPEG images. JPEG images from the training and testing datasets discussed in chapter 3, section 3.4.6 are used for the experimentation. The classifier is trained on the features extracted from training datasets. The classifiers ANN and NFC are employed to evaluate block based feature metric. Schematic representation of the methodology is depicted in figure 5.4.

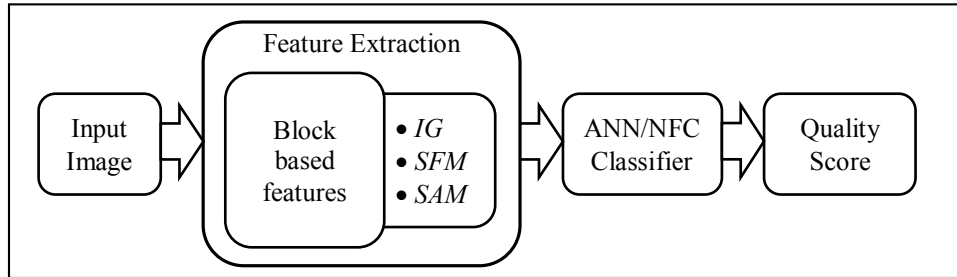


Figure 5.4: Schematic representation of methodology for block based feature metric

The architecture of this network is a three-layered feedforward with three input neurons in the input layer, nine in the hidden layer, and one in the output layer. The log-sigmoid transfer function is used for hidden layer neurons, and the linear transfer function is used for neuron of the output layer. The Levenberg-Marquardt training algorithm is used to train the ANN.

The features are also tested with neurofuzzy classifier where training data is divided into four clusters. These cluster definitions generate cluster labels. The centroid of the clusters is used for fuzzification. Each cluster represents a different range of

subjective scores as shown in table 5.1. This fuzzified matrix is given as input to MLPBPN for training. The trained network is used to find regression between target and output.

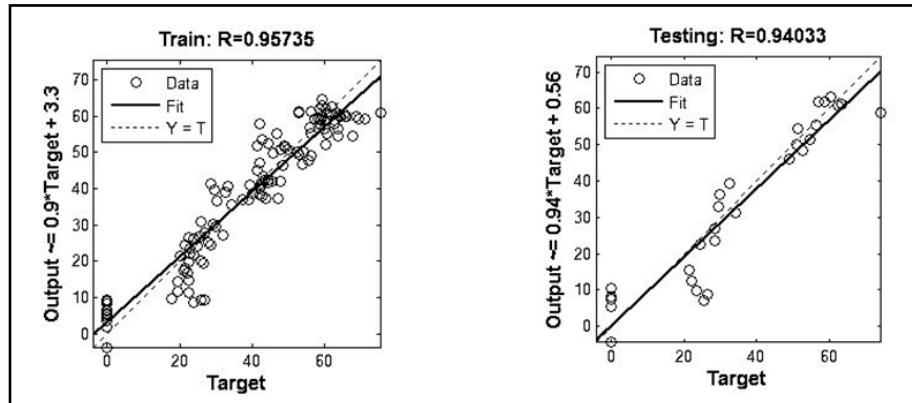
Table 5.1: Preparation of NFC clusters based on the subjective scores

NFC	Subjective Score (DMOS) for LIVE	Subjective Score (MOS) for TID
Cluster 1	0 to 20	$10 \geq n \geq 5$
Cluster 2	21 to 40	$5 > n \geq 3$
Cluster 3	41 to 60	$3 > n \geq 2$
Cluster 4	61 and more	$2 \geq n$

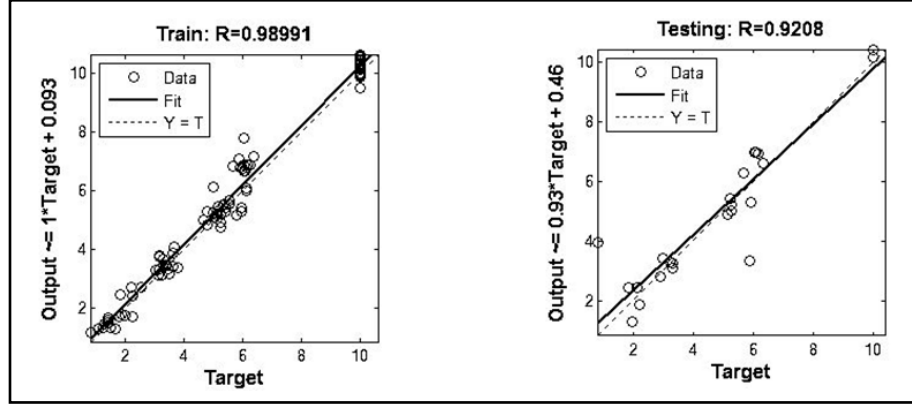
Performance of the metric is tested with ANN and NFC classifiers on LIVE, TID2008, and TID2013 datasets. Performance evaluation using LCC and SROCC for JPEG images is depicted in table 5.2 which shows good correlation between subjective score and objective score. ANN shows an edge over NFC. Regression plots of ANN for LIVE, TID2008, and TID2013 are shown in figure 5.5 which indicates good results with $R=0.94$, $R=0.92$ and $R=0.96$ respectively. The results show that the proposed metric works well for the blockiness distortion.

Table 5.2: Performance evaluation of block based features using ANN and NFC through LCC and SROCC for JPEG images from LIVE, TID2008, and TID2013 datasets

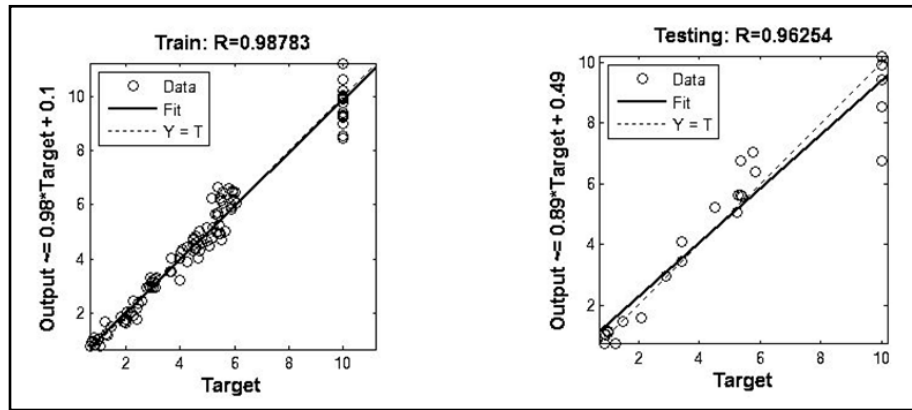
Classifier	LIVE		TID2008		TID2013	
	LCC	SROCC	LCC	SROCC	LCC	SROCC
ANN	0.94	0.92	0.92	0.92	0.96	0.95
NFC	0.91	0.88	0.90	0.89	0.91	0.92



(a) LIVE: DMOS (target) vs predicted score (output)



(b) TID2008: MOS (target) vs predicted score (output)



(c) TID2013: MOS (target) vs predicted score (output)

Figure 5.5: Regression Plots of ANN for training and testing datasets using block based features for JPEG images

The performance of the proposed metric for LIVE dataset is compared with state-of-the-art metrics. Performance comparison is presented below in table 5.3. It shows that performance of the proposed metric is comparable to the existing state-of-the-art metrics.

Table 5.3: Performance comparison of the proposed metric with state-of-the-art metrics on LIVE dataset

Sr. No.	Metrics	LCC	SROCC
1.	Z. Wang <i>et al.</i> [43]	0.75	-
2.	RCG-ELM [74]	0.70	-
3.	Suthaharan [44]	0.93	-
4.	R. Babu <i>et al.</i> [75]	0.96	0.94
5.	S. Lee <i>et al.</i> [45]	0.97	0.97
6.	S. Amiri <i>et al.</i> [47]	0.94	0.91
7.	Y. Zhan <i>et al.</i> [48]	0.97	0.96
8.	Proposed metric	0.94	0.92

5.2 Shape Adaptive Discrete Wavelet Transform Features

Loss of high frequencies caused by most of the distortions degrades the image quality. To model such distortions, majority of the existing algorithms analyse an image holistically and apply the transforms to entire image or by dividing it into blocks. However, the image may comprise different objects. Distortions that affect an image may also change the shape of these objects [139, 140]. Therefore, SADWT is considered to model the change in shapes of the objects present in the image. Also, the general wavelet transforms, due to their scaling function, analyse an image in three directions- horizontal, vertical, and diagonal, which restricts the level of analysis [148, 149]. On the other hand, SADWT tends to analyse more detailed edge information due to the conical nature of its scaling function. SADWT first analyses the shape and then applies the transform for extracting the features. Mean, SD, SVD, and shape features are computed from SADWT coefficients.

5.2.1 Design of Shape Adaptive Discrete Wavelet Transform Feature

The algorithm of Shape Adaptive Discrete Wavelet Transform:

SADWT effectively decomposes the object of any arbitrary shape into the multi resolution object. Algorithm for SADWT is described below [139].

1. Within the bounding box of the arbitrarily shaped object, use shape information to determine the first segment of pixels that belong to the object.
2. Scan the same row for subsequent pixels in the first segment.
3. Apply 1D-DWT to this section using orthogonal wavelets.
4. The wavelet coefficients generated by low-pass and high-pass filters are placed in the corresponding row in low-pass and high-pass bands respectively.
5. Repeat steps 3 and 4 for next section of successive pixels in the same row.
6. Repeat the same procedure (steps 2 - 5) for all subsequent rows one by one.
7. Follow the same procedure for all columns of low-pass and high-pass objects.
8. Repeat the same operations iteratively for low-pass band objects until decompositions level is reached.

This algorithm offers an approach to decompose arbitrarily shaped object in multi-resolution object effectively. It maintains object shape, locality and spatial correlation during the process. Arbitrary length wavelet decomposition applied in the above algorithm uses orthogonal wavelets.

Let $j(i)$, $k(i)$, $c(i)$, $d(i)$ be the impulse response of low-pass analysis, high-pass analysis, low-pass synthesis, and high-pass synthesis orthogonal wavelet filters with S channels and $i = 0, \dots, S-1$. The filter length S has even number. Subsequent equations define the relation among filters, wavelet analysis, and synthesis methods. The impulse responses of low-pass analysis and high-pass analysis filters with S channels have following relationships.

$$k(i) = (-1)^{S-1-i} j(S-1-i), \quad \text{for } i = 0, \dots, S-1 \quad (5.14)$$

Let low-pass synthesis filter be $c(i)$ then

$$c(i) = j(S-1-i), \quad \text{for } i = 0, \dots, S-1 \quad (5.15)$$

Let high-pass synthesis filter be $d(i)$ then

$$d(i) = k(S-1-i), \quad \text{for } i = 0, \dots, S-1 \quad (5.16)$$

Let $p(i)$ be the input signal. By filter-analysis method, the output of low-pass ($U(i)$) and high-pass ($V(i)$) filters before subsampling can be obtained as

$$U(i) = \sum_{j=0}^{S-1} p(i+j-S/2+1) c(S-1-j) \quad (5.17)$$

$$V(i) = \sum_{j=0}^{S-1} p(i+j-S/2+1) d(S-1-j) \quad (5.18)$$

The wavelet coefficients are obtained by subsampling of $U(i)$ and $V(i)$ by a factor of 2, as

$$CA(i) = U(2i-c) \quad (5.19)$$

$$CD(i) = V(2i-c) \quad (5.20)$$

where $CA(i)$ and $CD(i)$ depict low-pass and high-pass wavelet coefficients respectively; $c=0$ or $c=1$ depending upon whether the sub sampling is at even or odd position respectively.

For dimensionality reduction, statistical features mean, SD, and SVD are computed from the SADWT wavelet coefficients. Shape parameter of SADWT wavelet coefficients is also computed. Further, DWT which measures the energy present in the image and calculates the LH, HH, and HL coefficients, is applied to the image. Mean of DWT coefficients are calculated to obtain the features. The extracted feature vector of size 7 is evaluated for quality prediction by the classifier.

Figure 5.6 shows the original image “Monarch.bmp” from LIVE dataset with its different distortions along with their respective subjective score.

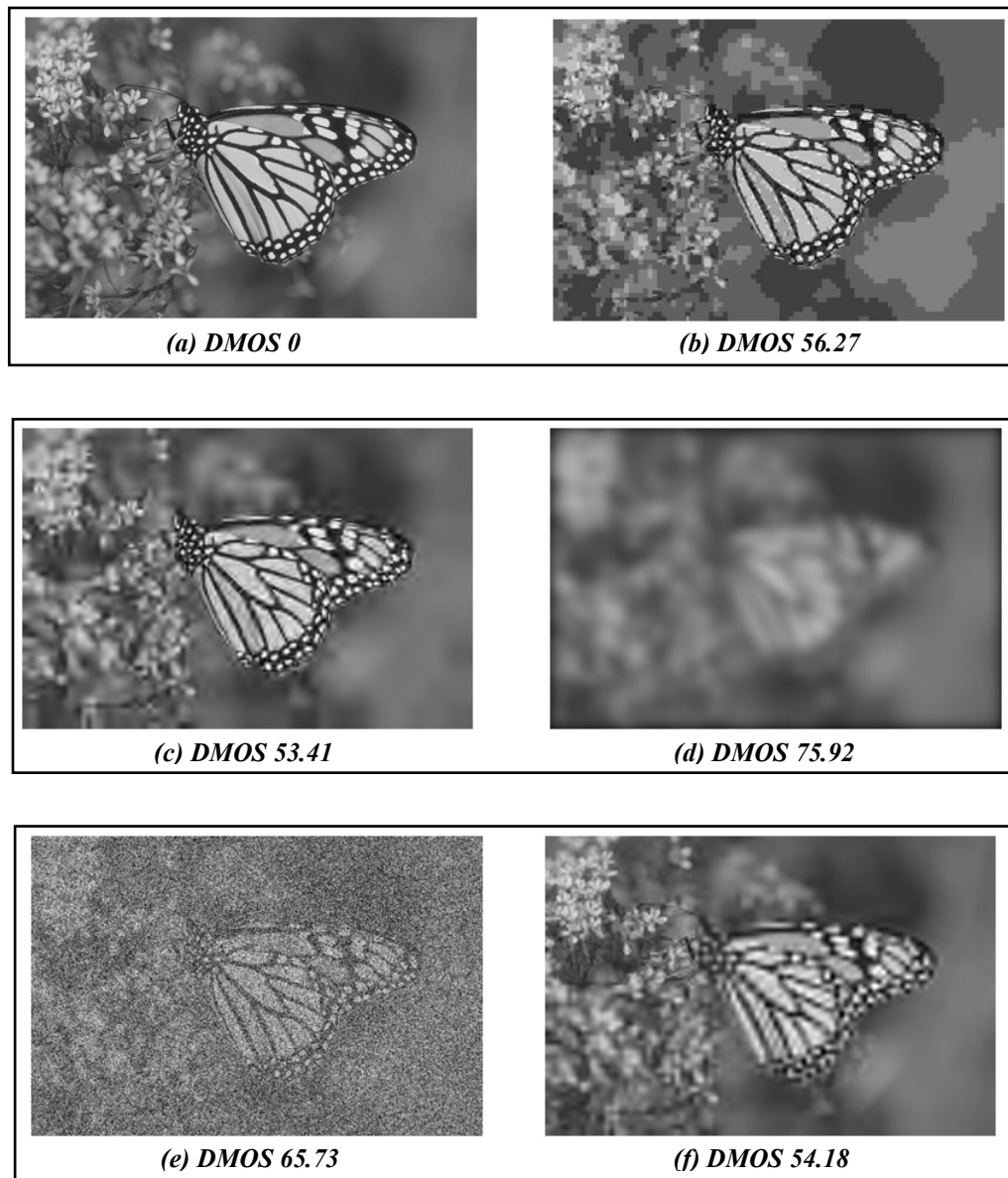


Figure 5.6: Reference image “Monarch.bmp” from LIVE dataset with different distortions (a) Original, (b) JPEG (c) JPEG2000, (d) Gblur, (e) WN, and (f) FF with their respective subjective scores

The following figure 5.7 presents a plot showing variations in mean, SD, SVD, and shape features computed from SADWT coefficients of “Monrach.bmp” and its five distorted variations. It is observed that the variation in the magnitude of feature values is not proportional to the amount of the degradations, and the variations with respect to the different distortion types do not follow linear relationship.

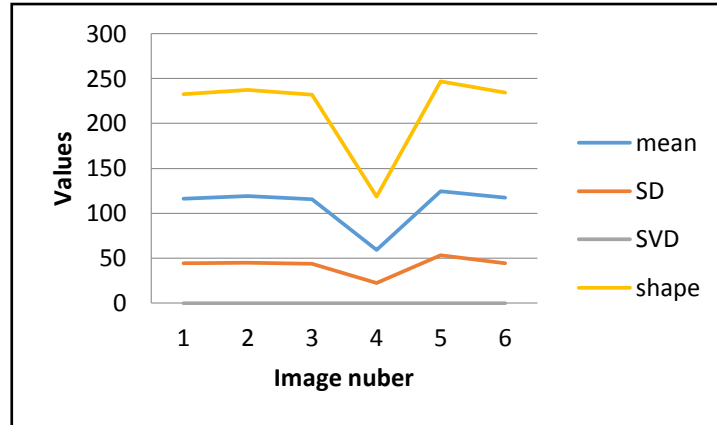


Figure 5.7: Plot showing variations in mean, SD, SVD, and shape features computed for SADWT coefficients of “Monrach.bmp” and its distorted variations

Due to the complex nature of the relationship between the features and subjective scores, ANN failed to handle the variations and remained untrained even though a sufficient number of images contributed in training. The performance of ANN depicted through regression plots shown in figure 5.8 demonstrate this fact. SADWT features of the training images from LIVE dataset are evaluated using ANN. Neuro-Wavelet Technique (NWT) discussed in the subsequent section is proposed to overcome this drawback.

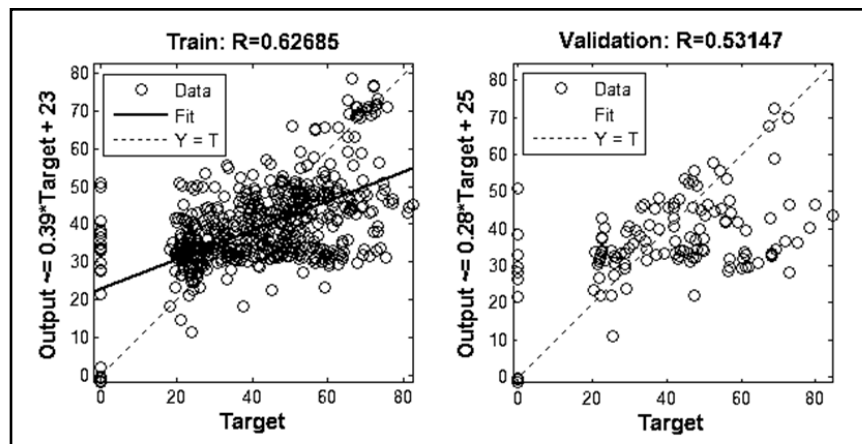


Figure 5.8: Regression plot for the undertrained ANN for LIVE dataset

5.2.2 Neuro-Wavelet Technique

The wavelet transform is widely used in the signal processing area to analyse the signal in frequency and time domain. [148, 149]. It is also used in the image processing domain to analyse low frequencies and high frequencies present in the image. The wavelet transform calculates the wavelet coefficients in various sub-bands and orientations. Multi resolution decomposition gives the insight of detailing about the distribution of pixels in the image. The wavelet transform is implemented through high-pass and low-pass filters which are applied to the image and sub-sampled into the approximations and detailed frequencies. Low-pass and high-pass filters are further applied to the low frequencies for multi-resolution investigation.

There are two types of wavelet transforms, namely Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The discrete wavelet transform is used in the present study which is defined as

$$\psi_{cs}(t) = 2^{\frac{-c}{2}} \psi(2^{-c}t - s) \quad (5.21)$$

where ψ represents mother wavelet, t and s represent the function of time and scale respectively, and $2^{-(c/2)}$ represents the normalized constant [149].

There are a number of mother wavelets such as Harr, Debauches (db), Biorthogonal, Symlet, etc. Debauches wavelet type 9 is used in the present study. A nonlinear variation in feature values causes the training problem to the neural network. The wavelet transform is applied to separate the low frequencies and the high frequencies. It helps the network to learn linearly through approximation which increases the accuracy.

Artificial Neural Network is a system which establishes a functional relationship between the nonlinear and linear correlation of feature vector as an input and the predicted score. Network architecture consists of an input layer with causative input neurons, and the output layer joined with the hidden layer. Nonlinear transfer functions in multiple layers of neurons help the network to learn the relationship between input and output variables. The mapping is achieved through an iterative process of reducing an error between expected (DMOS) and calculated output. The network is trained on the dataset using training algorithm to classify new dataset as the test data.

The discrete wavelet transform calculates the approximation and detailed coefficients when applied to the feature vector. The approximation coefficients are decomposed till the desired level to get more approximation and detailed coefficients to obtain the detailed variations in frequencies. Approximation coefficients and detailed coefficients are used for training two different neural networks. The inverse DWT is applied to the outputs of these networks to compute the quality score. Algorithm for the neuro-wavelet model is shown in figure 5.9.

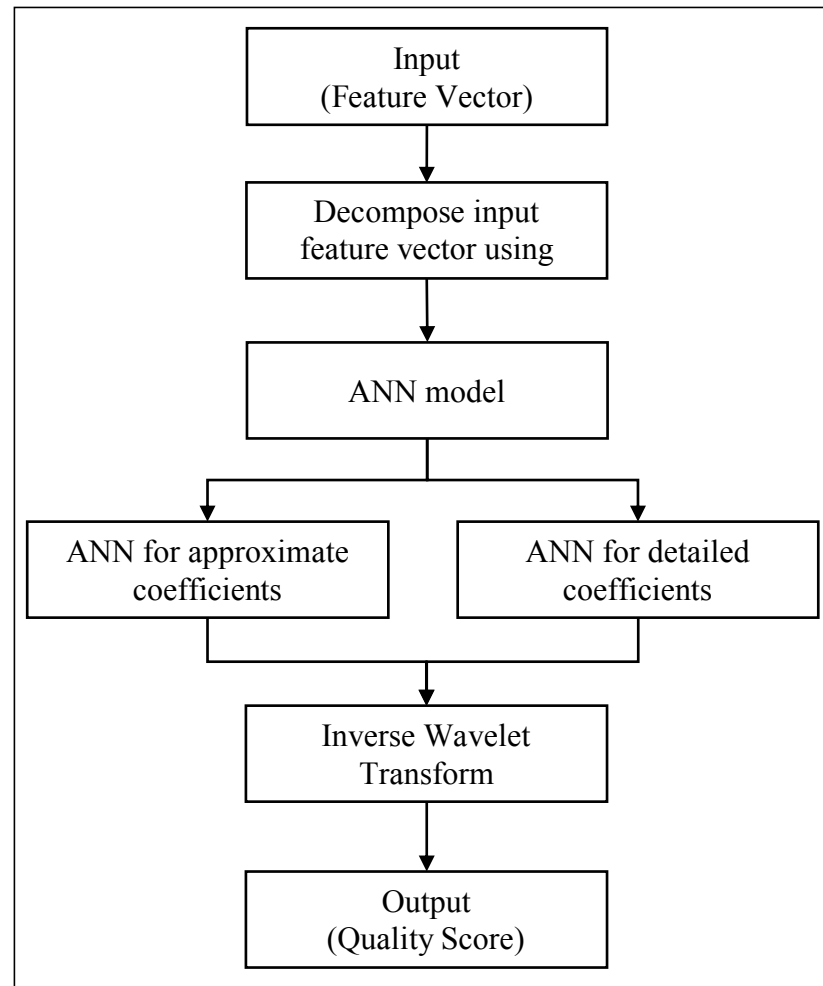


Figure 5.9: Algorithm of neuro-wavelet model

In the first level, image data is decomposed into approximate (CA1) and detailed (CD1) coefficients. The approximate (CA1), in the second level, is further divided into approximate (CA2) coefficient and detailed (CD2) coefficient. A similar process is followed for the third level to divide CA2 into CA3 and CD3 [148, 149]. Three detailed coefficients (CD1, CD2, and CD3) and approximate coefficient CA3 are then used for training of two different neural networks, one of which is trained

with an approximate coefficient and another with detailed coefficients. The methodology for third level decomposition using DWT is shown in figure 5.10. These networks are three-layered feedforward networks with Levenberg-Marquardt training algorithm. The log-sigmoid transfer function and linear transfer function are used in hidden neurons and out-neuron respectively.

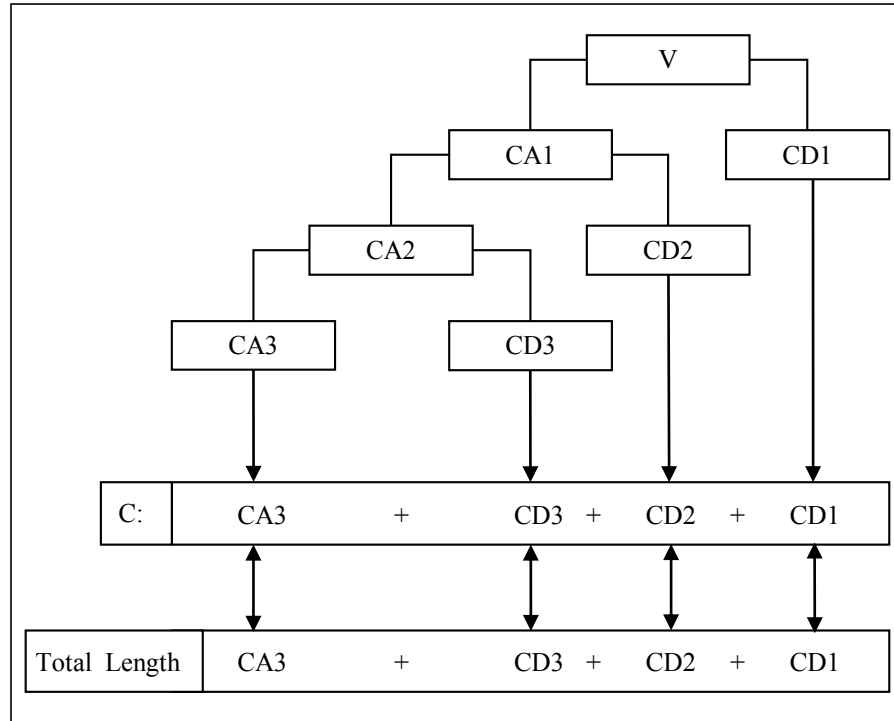


Figure 5.10: Multilevel decomposition by wavelet

5.2.3 Methodology and Results

The extracted shape adaptive discrete wavelet features and DWT features are combined to form a feature vector. The classifier is trained on the combined feature vector and predicts the quality score for test images. Training and testing datasets discussed in chapter 3, section 3.4.6 are used for the experimentation. The schematic representation of the methodology is presented in figure 5.11.

From the above figure 5.7, it is observed that the variations in the magnitude of the SADWT feature is very small and not proportional to the amount of degradation. ANN failed to adapt to such small variations and remained undertrained. Therefore, NWT classifier is used for the further experimentations. NWT, being a neural network, has the same architecture as ANN and differs only in number of input neurons. It has seven neurons in the input layer as there are seven input features.

The architecture of ANN is already discussed in section 5.1.2. The combined feature is also evaluated using NFC.

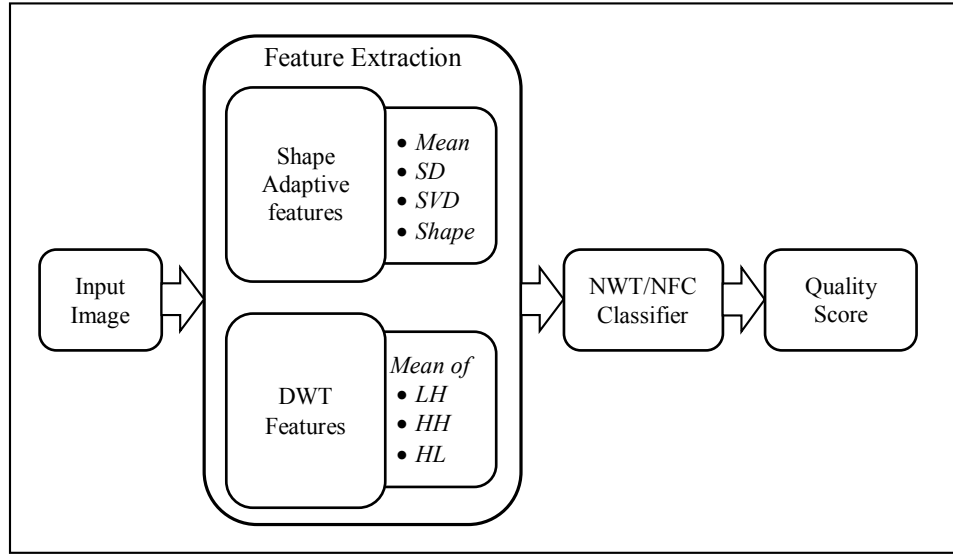


Figure 5.11: Schematic representation of methodology for combination of SADWT and DWT features

The proposed feature metric evaluated with NFC and NWT classifier on LIVE, TID2008, and TID2013 datasets. Performance of the combined metric is measured through LCC and SROCC. The results are presented in tables 5.4(a)-(c). This approach worked well independent of the databases. NWT has performed slightly better than NFC. The regression plots of NWT for LIVE, TID2008, and TID2013 show satisfactory results with $R = 0.91$, $R = 0.95$ and $R = 0.98$ respectively. The plots are presented in figure 5.12.

The experimental results show that the proposed metric performs well for overall quality estimation and also for other distortions considered for the experimentations. Table 5.5 presents the performance comparison of the proposed metric and state-of-the-art metrics which shows that the proposed metric performs reasonably well.

NWT is also tested for block based features which were earlier evaluated with ANN. The results obtained for the block based feature metric presented in the following table 5.6 show that it worked well for JPEG distortion. Regression plots of NWT are depicted in figure 5.13.

Table 5.4(a): Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on LIVE dataset

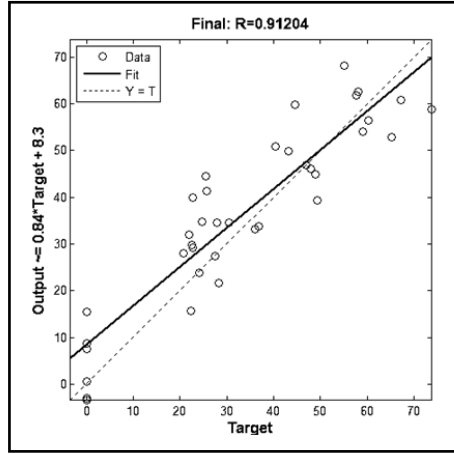
Distortion Type	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.91	0.91	0.97	0.94
JPEG2K	0.93	0.91	0.91	0.86
Gaussian Blur	0.87	0.86	0.90	0.89
White Noise	0.89	0.86	0.97	0.95
Fast Fading	0.87	0.87	0.98	0.96
Overall	0.88	0.87	0.91	0.90

Table 5.4(b): Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on TID2008 dataset

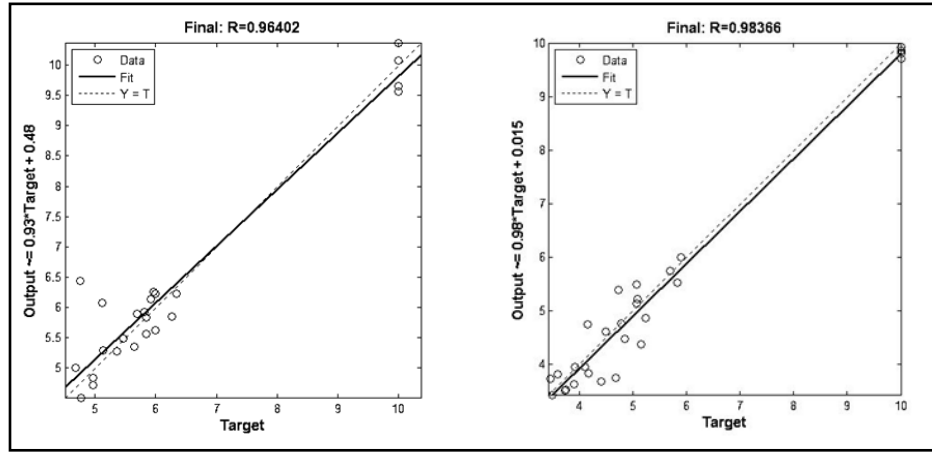
Distortion Type	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.90	0.88	0.97	0.95
JPEG2K	0.91	0.90	0.95	0.92
Gaussian Blur	0.86	0.86	0.97	0.90
White Noise	0.89	0.88	0.95	0.92
Overall	0.87	0.87	0.96	0.93

Table 5.4(c): Performance evaluation of combination of SADWT and DWT feature metrics using NFC and NWT through LCC and SROCC on TID2013 dataset

Distortion Type	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.92	0.92	0.95	0.94
JPEG2K	0.91	0.90	0.96	0.91
Gaussian Blur	0.88	0.87	0.97	0.92
White Noise	0.89	0.88	0.95	0.91
Overall	0.88	0.88	0.98	0.95



(a) LIVE



(b) TID2008

(c) TID2013

Figure 5.12: Regression Plots of NWT classifier using a combination of SADWT and DWT features

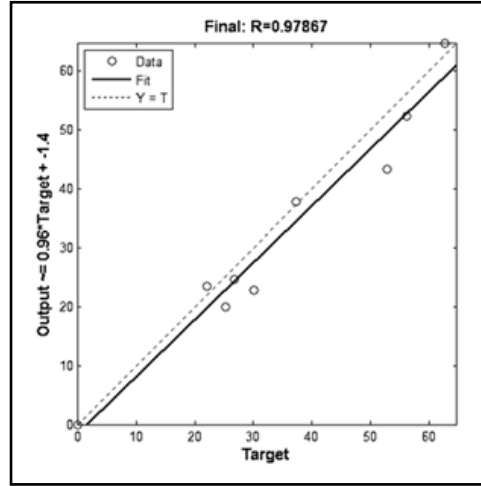
(a) DMOS (target) are plotted against predicted scores (output) for LIVE and (b)-(c) MOS (target) are plotted against predicted scores (output) for TID datasets

Table 5.5: Performance Comparison of the proposed combined feature metric with state-of-the-art metrics on LIVE dataset

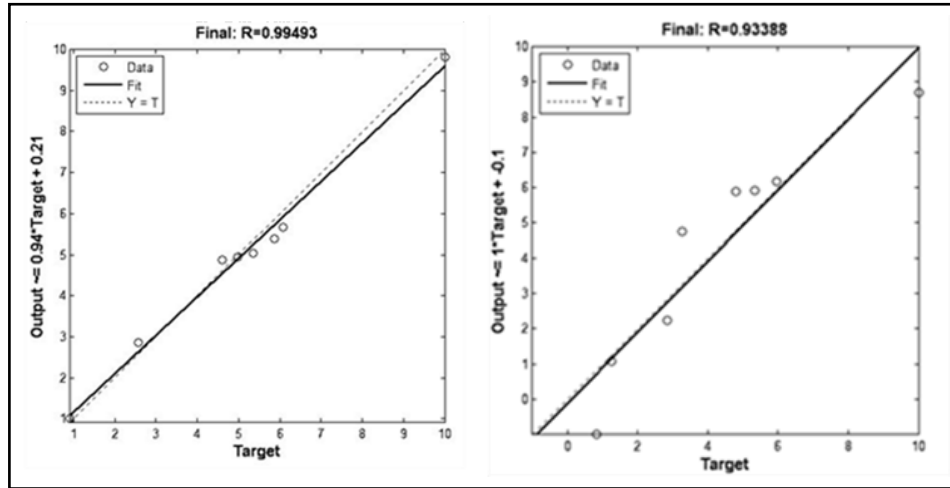
Sr. No.	Metrics	LCC	SROCC
1.	H. R. Sheikh <i>et al.</i> [101]	0.93	-
2.	BIQI [97]	0.82	0.84
3.	L. Liu <i>et al.</i> [107]	0.93	0.93
4.	W. Wu <i>et al.</i> [108]	0.94	0.91
5.	G. Wang <i>et al.</i> [109]	0.92	0.91
6.	L. Li <i>et al.</i> [110]	0.93	0.91
7.	Proposed Metric	0.91	0.90

Table 5.6: Performance evaluation of block based features using NWT through LCC and SROCC for JPEG images from LIVE, TID2008, and TID2013 datasets

Classifier	LIVE		TID2008		TID2013	
	LCC	SROCC	LCC	SROCC	LCC	SROCC
NWT	0.91	0.88	0.90	0.89	0.91	0.92



(a) LIVE



(b) TID2008

(c) TID2013

Figure 5.13: Regression Plots of NWT classifier using block based features for JPEG images

(a) DMOS (target) are plotted against predicted scores (output) for LIVE and (b)-(c) MOS (target) are plotted against predicted scores (output) for TID datasets

5.3 Summary

IG, SFM, and SAM features are implemented to quantify the degradation caused by the blocking artifact. ANN, NFC, and NWT classifiers are used for quality prediction. The obtained results show good performance for LIVE, TID2008, and TID2013 datasets. The results show that the proposed metric has an edge in the performance when compared with the existing algorithms.

Most distortions alter the edge information in an image. The best approach to assess the image quality is to model the change in the edge information. Images may also comprise different objects. Apart from analyzing the image holistically, an image can also be analysed at the object level. Therefore, object based analysis is carried out using SADWT for the present experimentation.

ANN and neurofuzzy classifiers were initially employed for the experimentation. However, as the variations in feature values are small, and feature magnitudes are not proportional to the degradation, hence ANN remain undertrained and failed. Therefore, the neuro-wavelet model is developed for quality prediction. Results are obtained for LIVE, TID2008, and TID2013 datasets using NFC and NWT classifiers. Performance of the proposed metric is comparable to the wavelet-based state-of-the-art metrics.

Chapter 6.

Design of NRIQA Metric using Fusion of Curvelet and SIFT Features

Images are usually classified as natural images and artificial images. Every natural image has strong statistical properties across visual contents such as self-similarity, non-Gaussian, scale invariance, high dimensional singularity, and leading edge. While the leading edge conveys the strong presence of continuous points, non-continuous points convey about the curve through high dimensional singularity [150]. As the intensities of these points are sensitive to the distortions, these points can be extracted to be used as a feature description. The changes in these intensity values can be measured using SIFT and curvelet transforms. As the continuous and non-continuous points have statistical importance for quality estimation of natural images, and edges and curves both are equally important for the quality assessment of natural images, so a fusion of SIFT and curvelet is proposed as a feature for quality prediction.

NFC and NWT are used and successfully tested in the experimentations discussed in the previous chapter. A fusion of the feature vectors is proposed in this chapter to test these classifiers further. A new metric is implemented by the fusion of SIFT key points and the coefficients identified by the curvelet transform. The spatial covariance is also implemented as another feature which is later combined with the fused feature metric. Both the metrics are tested with NFC and NWT separately on LIVE, TID2008, and TID2013 image datasets. Combination of these metrics is also tested.

This chapter is mainly focused on the design of the fusion metric. Its implementation is discussed in section 6.1. Subsection 6.1.2 presents the implementation of spatial covariance metric. Section 6.2 details the methodology and results. Finally, a summary of the chapter is presented in section 6.3.

6.1 Feature Extraction

The hypothesis for the NSS approach is that natural images acquire very small space in the space of all possible images. NSS metric can evaluate the dissimilarity between a subset of pristine natural images and the degraded images [151, 152].

Natural images are high quality images of the visual environment. These images can have a variety of objects such as animals, birds, hills and valleys, rocks and statues, cars, roads, etc. All natural images follow certain statistical characteristics across the visual contents. Leading edge is one such strong characteristic of natural images which gets altered by the distortions induced during the image processing tasks. The key points convey rich information about the changes occurred to the edges due to the distortions. Therefore, SIFT key points are determined to quantify the distorted edges. However, SIFT tends to ignore the points such as those in the low contrast region. Therefore, a fusion of SIFT key points and the points identified by curvelet transform is proposed to generate a fused feature vector.

Spatial covariance computes the relationship between neighbouring pixels. Block based transform and quantization used for the compression, noise, and blur affects this relationship. Thus, spatial covariance is implemented to quantify the degradation to the image caused due to these distortions.

6.1.1 Fusion Feature Vector

SIFT key point detector is robust to noise and invariants to the scale change. The intensities of SIFT key points change as a result of the distortions and hence are useful in quantification of the distortions. SIFT algorithm selects peak in a scale-space by computing maximum and minimum Difference of Gaussian (DoG) using scales. The unstable points in the low contrast region and the points on the edges are ignored by SIFT [153, 154] as shown in figure 6.1(a). The key points ignored by SIFT are identified by curvelet transform as shown in figure 6.1(b). Both these sets of key points provide information about an image and the changes occurred due to the distortions. Therefore, the coefficients identified by curvelet transform are fused with SIFT key points to generate a feature vector.



Figure 6.1: (a) SIFT feature point map and (b) curvelet coefficients map for "Monarch.bmp" from LIVE dataset

6.1.1.1 Scale Invariant Feature Transform (SIFT)

When applied to the image, SIFT generates a large set of local features invariant to scaling, rotation and translation of an image. SIFT algorithm has four step filtering process [150].

Step1: Scale-Space Extrema Detection

The Gaussian scale function is applied to an image to detect the location and scale of a point from different directions. The scale space is computed as convolving input image with variant-scale Gaussian function –

$$S(p, q, \sigma) = g(p, q, \sigma) * i(p, q) \quad (6.1)$$

The even key points in the scale space are identified by calculating the Difference of Gaussian (DoG). The scale space extrema $DG(p, q, \sigma)$ is the difference between two images where the first image is scaled k times than the other. It is mathematically defined as

$$DG(p, q, \sigma) = g(p, q, k\sigma) * g(p, q, \sigma) \quad (6.2)$$

Local maxima and minima of $DG(p, q, \sigma)$ are calculated by comparing each point in the image with its 8 neighbours at the same scale and 9 neighbours at one level up and down. If this value is minimum or maximum of these points, then this point is an extrema.

Step 2: Key points Localization

Key point localization algorithm attempts to refine the location of identified points. Points with low contrast or the points poorly distributed on the edges are removed from the identified key points.

Laplacian values are computed to identify the low contrast points for each key point identified in step 1. The location of extremum l_z , is given as

$$l_z = \frac{\partial D G^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (6.3)$$

If the function value at l_z is less than the contrast threshold value, the point is identified as the local extrema with low contrast, and it is eliminated.

An edge has high principle curvature but low minimal curvature. DoG calculates the difference between the curvatures. If this difference is below the ratio of largest to the smallest eigenvector, from the 2x2 Hessian matrix at the location and scale of the key point, the key point at the edge is eliminated.

Step 3: Orientation Assignment

Based on the properties of an image, orientation is assigned to the key points. The key point descriptors can be represented relative to the orientation.

To find the orientation, the gradient magnitude (M) and orientation (θ) are calculated as,

$$M(p, q) = \sqrt{(S(p+1, q) - S(p-1, q))^2 + (S(p, q+1) - S(p, q-1))^2} \quad (6.4)$$

$$\theta = \tan((S(p, q+1) - S(p, q-1)) / (S(p+1, q) - S(p-1, q))) \quad (6.5)$$

The orientation histogram from the gradient orientation of sample points is plotted. The highest peak in the histogram and other local peaks within 80% of the height of the highest peak are identified to create key points with that orientation.

Step 4: Key point Descriptors

The gradient data is oriented and weighted by Gaussian with variance $1.5 \times$ key point scale. It creates a set of histograms over a window centred on the key point. Key point descriptors are generally defined as the set of 16 histograms aligned in a 4×4 grid, each with 8 bin orientations.

6.1.1.2 Curvelet Transform

Curvelet transform has two alternatives, namely, Unequally Spaced Fast Fourier Transform (USFFT) and wrapping function. In the proposed metric, work wrapping function is used which comprises of several sub-bands at different scales consisting of different orientations and positions in the frequency domain. Curvelet transform, when applied on an image of size $M \times N$, generates curvelet coefficients C , indexed by scale a , orientation b , and spatial location parameters p and q as [155] -

$$C(a, b, p, q) = \sum_{m=0, n=0}^{M, N} f(m, n) \varphi_{a, b, p, q}(m, n) \quad (6.6)$$

6.1.1.3 Implementation of A Fusion Feature Metric

SIFT generates a large number of key points. For efficient implementation of the metric, averaging technique is used for dimensional reduction of the feature. In the averaging technique, the threshold value is computed by averaging the key point values. The values above the threshold contribute as the feature values. Curvelet also generates a large set of coefficients which is reduced to a smaller set of feature values by applying the SVD. SIFT and curvelet coefficients are fused using weighted sum as SIFT identifies maximum key points whereas few points that lie on edges are identified by the curvelet transform. Weight 0.6 is assigned to SIFT key points, and weight 0.2 is assigned to curvelet coefficients. Weighted sum resulted into a feature vector of size 128. The flowchart of the implementation of the fusion vector is shown in figure 6.2.

The change in the obtained features occurred due to the distortions are depicted in figure 6.3 using histograms of original image “bikes.bmp” and its distorted variants from LIVE image dataset. It is observed that the histograms of distorted images are different from that of the original image. Therefore, this feature is used for quantification of distortions.

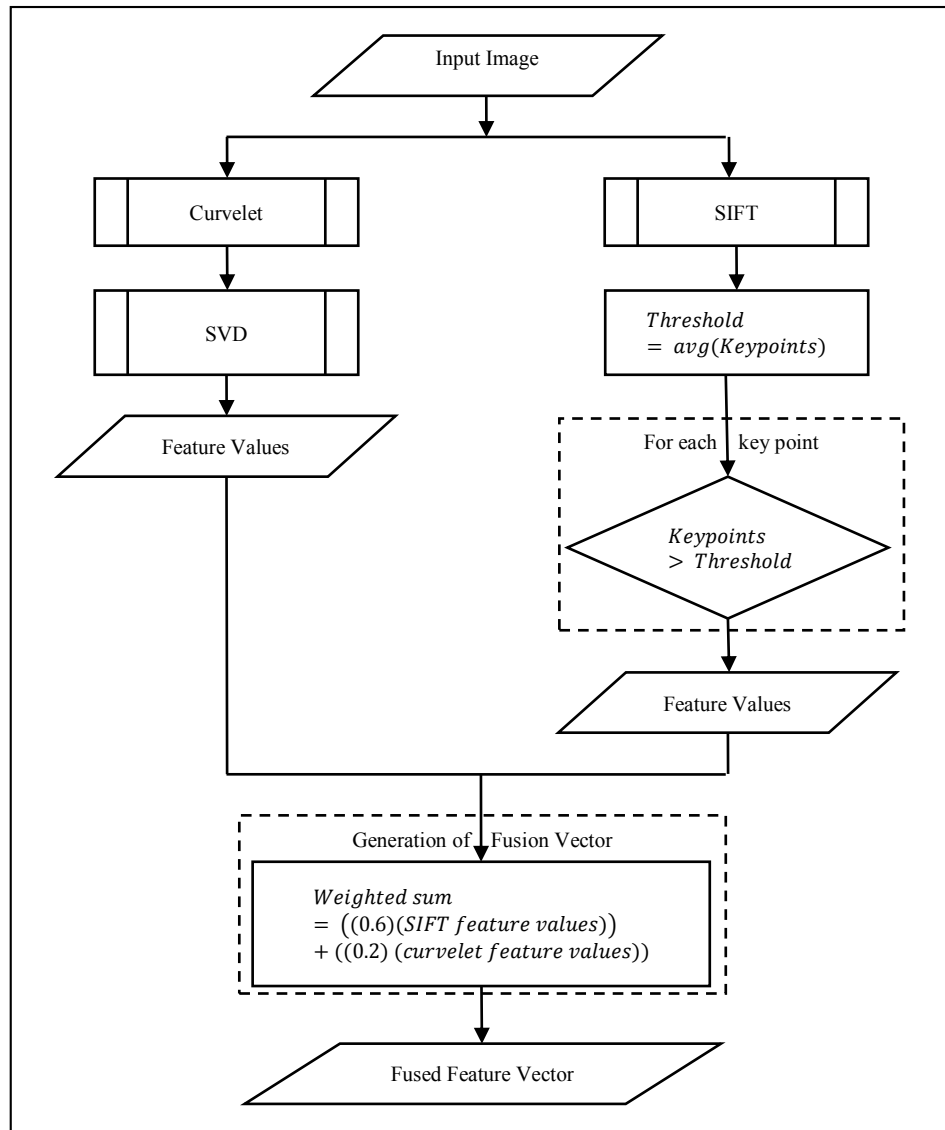
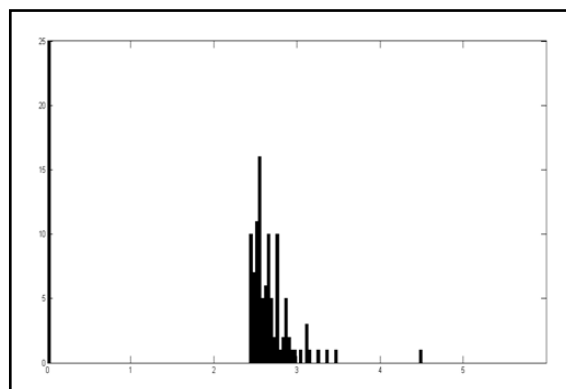
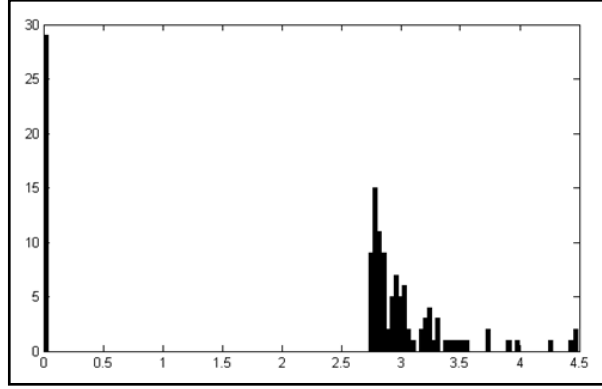


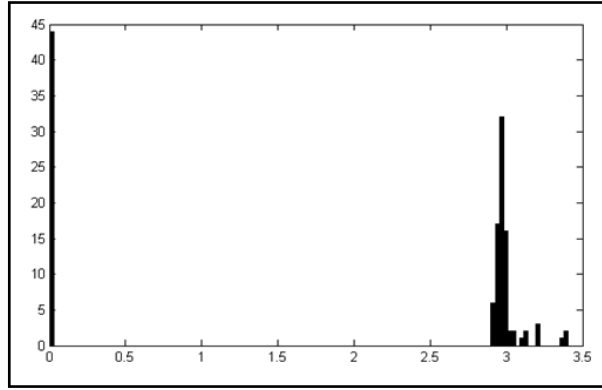
Figure 6.2: Flowchart of the fusion of SIFT key points and Curvelet coefficients



(a)



(b)



(c)

Figure 6.3: Histograms of feature values (a) Original image “bikes.bmp”, (b) Gaussian blur, and (c) JPEG2000 from LIVE dataset

6.1.2 Spatial Covariance

Spatial covariance computes the relationship between neighbouring pixels. Block based transform and quantization used for the compression, noise, and blur affects this relationship. Spatial covariance is used to quantify the degradation in an image.

Statistically, spatial covariance is computed in the following steps. In the first step, the gradient of an image is calculated as

$$\psi g = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial g}{\partial x} \\ \frac{\partial g}{\partial y} \end{bmatrix} \quad (6.7)$$

where $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$ represent gradient in x-direction and y-direction respectively

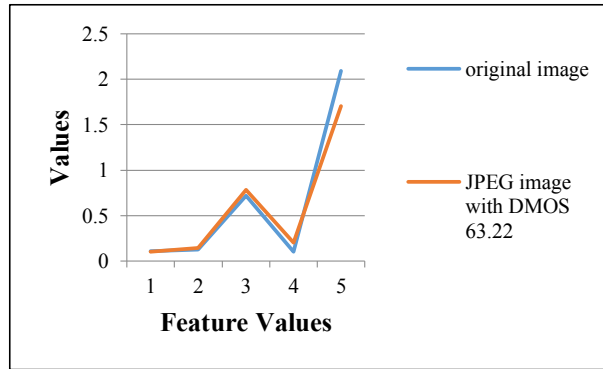
Further, the covariance is computed as

$$Cov(x_1, x_2) = (E(x_1 - \mu_1) \times (x_2 - \mu_2)) \quad (6.8)$$

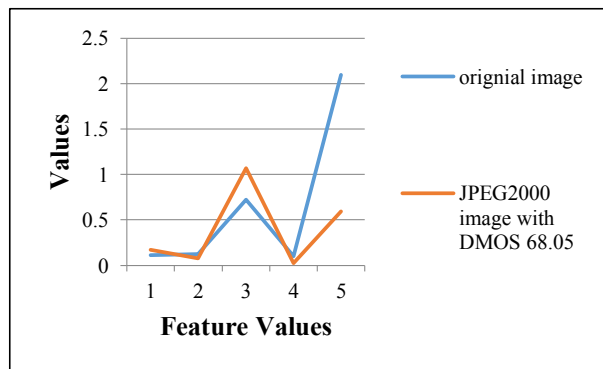
where E is the mathematical expectation and μ is the mean of both the directions.

To calculate the feature value, the image is split into 8×8 blocks, and the covariance between two neighbouring image blocks is calculated. A covariance matrix of size 5×5 is generated by filtering the significant values. The 2D matrix is reduced to 1D vector of size 5×1 by applying the single value decomposition. The resultant 5×1 vector represents the feature vector.

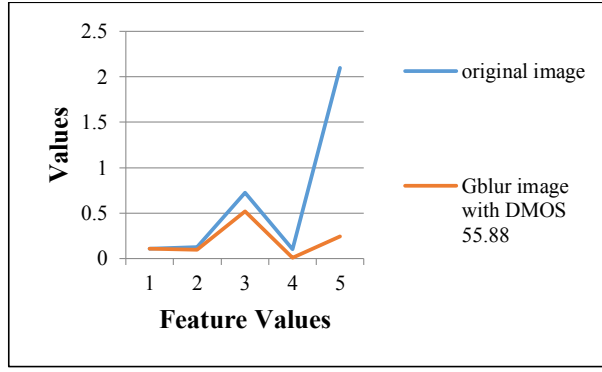
The variations in feature values are depicted in figure 6.4 for “bikes.bmp” and its distortion variants from LIVE dataset. The figure illustrates that this feature shows significant variation for noise, blur and fast fading distortions. In white noise, each pixel value is different from the neighbouring pixel. Therefore, the feature values for the white noise image are larger than the original image. Whereas, the difference between each pixel and its neighbouring pixel is less in case of blur and fast fading distortions resulting in the smaller feature values.



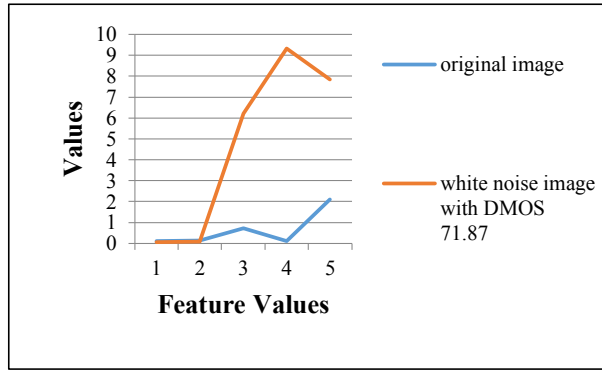
(a) Comparison of the original image and its JPEG variant



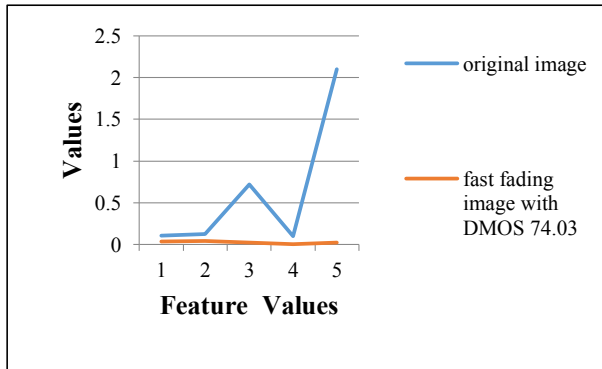
(b) Comparison of the original image and its JPEG2000 variant



(c) Comparison of the original image and its Gblur variant



(d) Comparison of the original image and its white noise variant



(e) Comparison of the original image and its fast-fading variant

Figure 6.4: Comparison of the feature values calculated using spatial covariance for original image “bikes.bmp” and its variants from LIVE dataset

6.2 Methodology and Results

The schematic representation of the methodology is presented in figure 6.5. SIFT key points, and curvelet coefficients are fused to obtain fusion metric. Spatial covariance feature is also extracted and combined with the fusion metric. The classifier trained on the combined feature metric predicts the quality of test images.

Training and testing datasets discussed in chapter 3, section 3.4.6 are used for the experimentation.

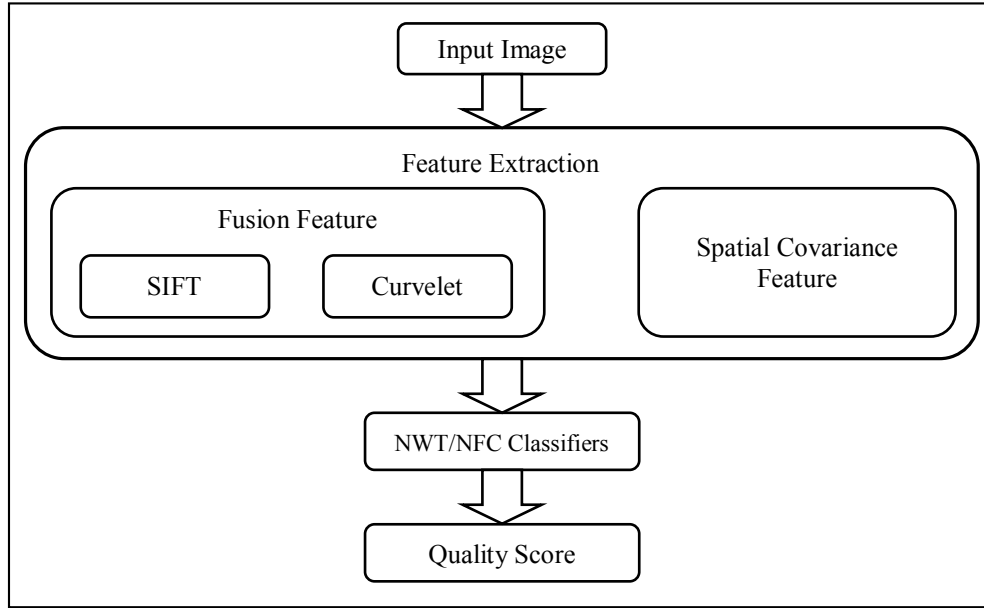


Figure 6.5: Schematic representation of methodology for performance evaluation of a combination of fusion feature and spatial covariance feature

Apart from the results obtained for the combination of the fused feature vector and spatial covariance feature, the model is also tested for fused feature vector and spatial covariance feature separately.

6.2.1 Performance Evaluation of Fusion Metric

The proposed fusion metric is evaluated using NFC and NWT on LIVE, TID2008, and TID2013 datasets. Performance of fusion of SIFT and curvelet feature metric is measured through LCC, SROCC on all the three standard datasets. Results presented in table 6.1(a)-(c) exhibit a good correlation between subjective score and predicted score for LIVE, TID2008, and TID2013 respectively.

It is observed that the fusion metric works well for the different types of distortions independent of the database. NWT has performed slightly better than NFC. Regression plots of NWT for LIVE, TID2008, and TID2013 show satisfactory results with $R = 0.90$, $R = 0.93$ and $R = 0.96$ respectively as shown in figure 6.6. This metric shows good performance for TID2013 dataset than LIVE and TID2008.

Table 6.1(a): Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on LIVE dataset

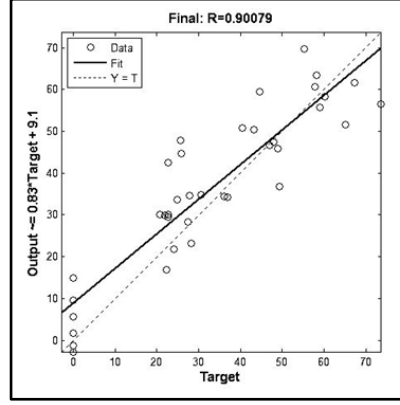
Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.90	0.89	0.97	0.96
JPEG2000	0.87	0.85	0.92	0.90
Gaussian Blur	0.87	0.86	0.98	0.95
White Noise	0.89	0.87	0.96	0.96
Fast Fading	0.87	0.85	0.96	0.95
Overall	0.88	0.87	0.90	0.89

Table 6.1(b): Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on TID2008 dataset

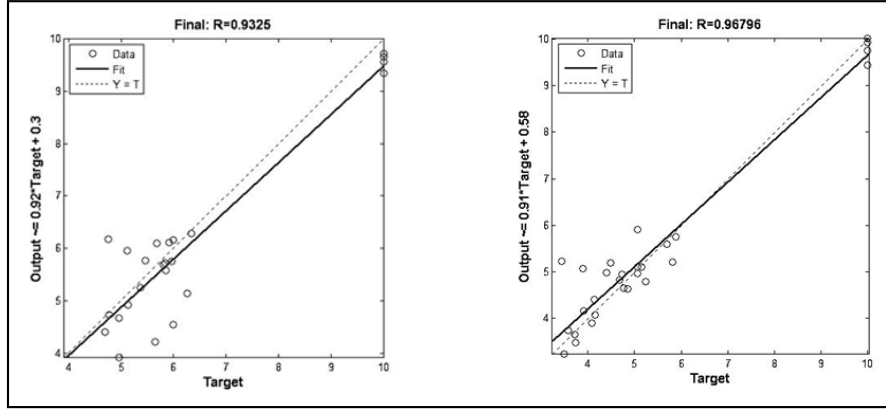
Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.89	0.88	0.95	0.94
JPEG2000	0.86	0.84	0.92	0.90
Gaussian Blur	0.86	0.85	0.96	0.95
White Noise	0.89	0.87	0.94	0.94
Overall	0.87	0.85	0.93	0.91

Table 6.1(c): Performance Evaluation of fusion metric using NFC and NWT through LCC and SROCC on TID2013 dataset

Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.90	0.89	0.95	0.94
JPEG2000	0.88	0.86	0.93	0.92
Gaussian Blur	0.90	0.89	0.93	0.93
White Noise	0.90	0.88	0.95	0.94
Overall	0.92	0.91	0.96	0.94



(a) LIVE



(b) TID2008

(c) TID2013

Figure 6.6: Regression plots of NWT using fusion feature metric
(a) DMOS (target) are plotted against predicted scores (output) for LIVE and
(b)-(c) MOS (target) are plotted against predicted scores (output) for TID
datasets

6.2.2 Performance Evaluation of Spatial Covariance Metric

Similarly, spatial covariance metric is also evaluated for the three dataset and results are presented in table 6.2(a)-(c). The metric works well independent of the dataset. Regression plots of NWT for LIVE, TID2008, and TID2013 shows satisfactory results with $R = 0.94$, $R = 0.94$ and $R = 0.96$ respectively. Plots are shown in figure 6.7.

Table 6.2(a): Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on LIVE dataset

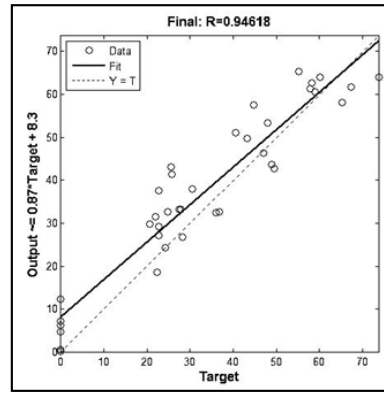
Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.93	0.92	0.96	0.94
JPEG2000	0.90	0.90	0.94	0.93
Gaussian Blur	0.87	0.86	0.89	0.89
White Noise	0.92	0.91	0.96	0.95
Fast Fading	0.92	0.92	0.94	0.94
Overall	0.90	0.89	0.94	0.92

Table 6.2(b): Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2008 dataset

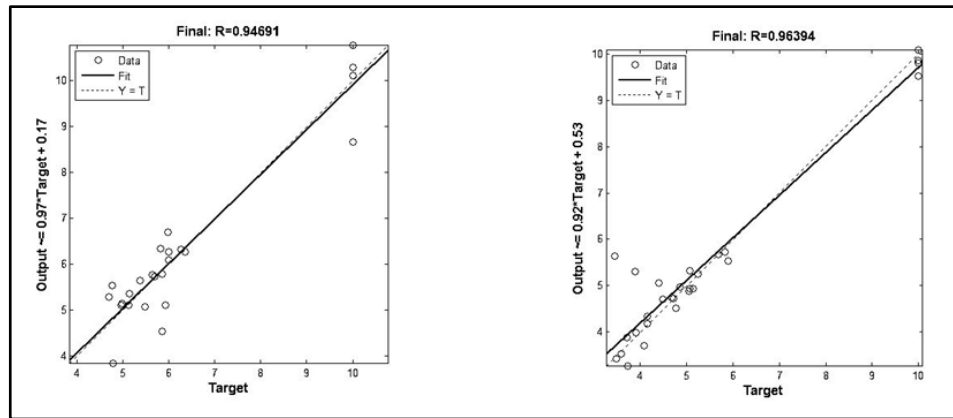
Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.92	0.92	0.98	0.96
JPEG2000	0.91	0.90	0.94	0.93
Gaussian Blur	0.92	0.92	0.94	0.92
White Noise	0.94	0.92	0.97	0.96
Overall	0.91	0.90	0.94	0.93

Table 6.2(c): Performance Evaluation of spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2013 dataset

Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.90	0.89	0.97	0.95
JPEG2000	0.88	0.86	0.94	0.93
Gaussian Blur	0.90	0.89	0.92	0.91
White Noise	0.90	0.88	0.94	0.92
Overall	0.92	0.91	0.96	0.94



(a) *LIVE*



(b) *TID2008*

(c) *TID2013*

Figure 6.7: Regression plots of NWT using spatial covariance metric
(a) DMOS (target) are plotted against predicted scores (output) for LIVE and
(b)-(c) MOS (target) are plotted against predicted score (output) for TID
datasets

6.2.3 Performance Evaluation of Combined Feature Metric

Combination of the fused metric and spatial covariance feature metric is evaluated for test images from LIVE, TID2008, and TID2013. The results are presented in table 6.3(a)-(c). Regression plots of NWT for LIVE, TID2008, and TID2013 shows satisfactory results with $R = 0.90$, $R = 0.91$ and $R = 0.96$ respectively. Plots are shown in figure 6.8.

Table 6.3(a): Performance Evaluation of combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on LIVE dataset

Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.92	0.90	0.94	0.93
JPEG2000	0.89	0.89	0.93	0.91

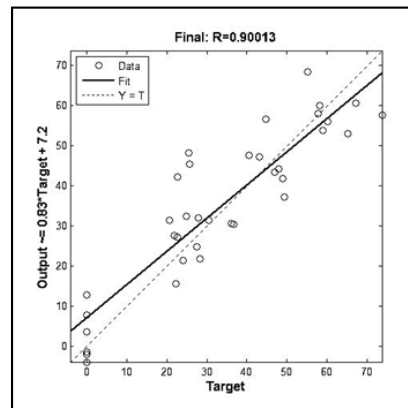
Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
Gaussian Blur	0.92	0.90	0.94	0.93
White Noise	0.91	0.90	0.97	0.95
Fast Fading	0.92	0.91	0.97	0.96
Overall	0.88	0.87	0.90	0.90

Table 6.3(b): Performance Evaluation of the combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2008 dataset

Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.89	0.89	0.90	0.89
JPEG2000	0.90	0.89	0.94	0.92
Gaussian Blur	0.92	0.90	0.97	0.95
White Noise	0.91	0.90	0.97	0.96
Overall	0.88	0.88	0.91	0.90

Table 6.3(c): Performance Evaluation of the combination of fusion metric and spatial covariance feature metric using NFC and NWT through LCC and SROCC on TID2013 dataset

Distortion Types	NFC		NWT	
	LCC	SROCC	LCC	SROCC
JPEG	0.89	0.89	0.90	0.90
JPEG2000	0.90	0.89	0.94	0.92
Gaussian Blur	0.90	0.89	0.93	0.91
White Noise	0.90	0.89	0.93	0.92
Overall	0.90	0.89	0.96	0.95



(a) LIVE

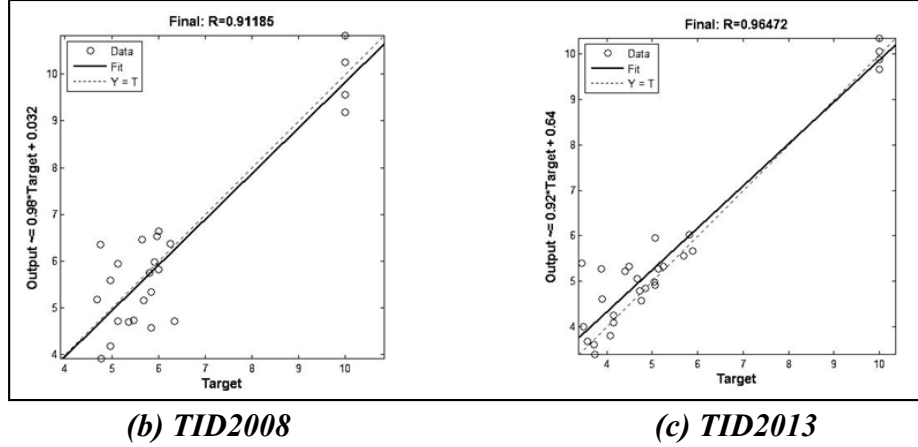


Figure 6.8: Regression plots of NWT using a combination of fused feature metric and spatial covariance metric

(a) DMOS (target) are plotted against predicted scores (output) for LIVE and (b)-(c) MOS (target) are plotted against predicted scores (output) for TID datasets

Performance comparison of fusion metric, spatial covariance metric, and the combined feature metric with the state-of-the-art metrics is presented in table 6.4. The proposed metrics exhibit comparable results.

Table 6.4: Performance comparison of all three proposed metrics with state-of-the-art metrics through LCC and SROCC on LIVE, TID2008, and TID2013 datasets

Metric	LIVE		TID2008		TID2013	
	LCC	SROCC	LCC	SROCC	LCC	SROCC
DIIVINE	0.8659	0.8596	0.9038	0.8930	0.6313	0.5296
C-DIIVINE	0.9000	0.8948	NA	NA	0.6693	0.5563
BLIINDS-II	0.9272	0.9214	0.9219	0.8982	0.6982	0.6072
BRISQUE	0.9311	0.9251	0.9391	0.9357	0.6821	0.5970
GM-LOG	0.9545	0.9497	NA	NA	0.7673	0.7127
NIQE	0.9122	0.9111	0.804	0.795	0.4308	0.3249
IL-NIQE	0.9095	0.9055	NA	NA	0.6494	0.5252
FRIQUEE-II	0.962	0.942	NA	NA	0.704	0.669
CORNIA	0.943	0.942	0.9347	0.8990	0.613	0.549
Proposed (fused metric)	0.900	0.897	0.932	0.919	0.967	0.948
Proposed (Spatial covariance metric)	0.946	0.928	0.946	0.937	0.963	0.929
Proposed (combined)	0.900	0.903	0.911	0.909	0.964	0.951

6.3 Summary

Distortions introduced in an image affect the intensities of the key points which can be identified by SIFT. As SIFT does not identify all the key points along the edges, curvelet transform is employed to obtain curvelet coefficients. The weighted sum of SIFT key points and curvelet coefficients is calculated to generate a fused feature metric. NFC and NWT are used for evaluating the fused feature metric. The results show that the fused metric works well to quantify the different distortions.

Spatial covariance is also implemented as another feature to handle the noise present in an image. The feature metric performs reasonably well for the tested distortions; however, its performance is comparatively lower for blurry images as expected.

The fused feature metric and spatial covariance feature metric are combined and evaluated. Quality is predicted using a machine learning approach. Results are obtained using NFC and NWT. NWT is used to improve the accuracy of the classifier. The models work well for the combined feature vector. The results showed that the correlation between subjective score and the predicted score is comparable to the existing NSS based metrics.

Chapter 7.

Optimum Framework for NRIQA

Right from the acquisition, a digital image undergoes various image processing tasks and applications such as compression, storage, retrieval, enhancement, transmission, reconstruction, and display. Different image processing applications and tasks introduce different types of distortions. These distortions are required to be quantified for image quality assessment. Quality assessment of an image is given paramount consideration in the modern communication systems and multimedia applications for improvement in the user experience. Many such systems, applications, and imaging devices need to estimate the image quality in the absence of the reference image or without prior knowledge of the distortions present in the image. On the fly quality improvement measures are also essential in many applications such as biometric systems based on facial information, remote printing. Objective NRIQA methods play a crucial role in estimating the image quality in such situations. Machine learning based and NSS based are the two approaches usually followed to develop the distortion-generic NRIQA metrics.

NRIQA methods are developed on the fundamental assumption that the quality of an image can be assessed without direct matching with the original image [156]. NRIQA methods are usually classified as distortion specific and distortion-generic. Distortion specific NRIQA methods assume that the distortion present in an image is known. Whereas, distortion-generic or general-purpose IQA methods predict the image quality regardless of the distortions present in the image.

The machine learning based framework for NRIQA metric is proposed in this thesis for estimating the image quality regardless of the distortions. It estimates the ‘overall’ image quality without prior knowledge of the distortion/s present in the image. The framework achieves this by means of collection and combination of the purposefully selected different features relevant to the different distortions.

Different features that have been explored and tested in the previous chapters, which differ in their relevance, extraction methods, and domains, are combined and used for validating the framework. Size of the combined feature vector is reduced by

applying PCA. The reduced feature vector, which remains available in the feature repository, is used to train the quality predictor.

Besides building the features repository, evolved and grown with each extracted, evaluated, and added feature, the machine learning based quality prediction models namely ANN, SVM, NFC, and NWT are also employed and tested. While ANN could not handle the small variations in the feature values in certain situations and hence excluded from further experimentations, NFC and NWT performed satisfactorily and hence proposed as the quality predictors for the framework.

The performance of the combined feature metric is evaluated using NFC and NWT on LIVE, TID2008, and TID2013 datasets. The combined feature metric is optimized by applying PCA and evaluated. The results are compared with the results obtained in the experimentations discussed in the previous chapters 4, 5, and 6. Performance of the combined feature metric and dimensionally reduced feature metric is also compared to confirm the performance gain.

Design of the proposed framework for NRIQA metric is discussed in section 7.1. A short introduction to the motivation behind using a combined feature metric for the proposed framework is given in this section followed by the PCA in section 7.1.1. Section 7.1.2 discusses the methodology. Schematic representation of the framework and its evolution is described in this section through the diagrams. Performance evaluation and results are presented and discussed in section 7.1.3. The last section 7.3 summarizes the chapter.

7.1 Design of The Proposed Framework for NRIQA

A machine learning based NRIQA metric, once trained on the features relevant to the specific distortion, can evaluate an image for that distortion. To make the framework ‘generic’, it has to be trained on the universally applicable features set which are sensitive to the various distortion types. Implementing such universally applicable features set is highly challenging though may not be impossible [114]. The combination of purposefully selected different relevant features, however, works better in most of the scenarios and hence are frequently used in numerous research work carried out and published in the area of IQA and object recognition [157, 158]. Based on these observations, the framework for machine learning

based NRIQA metric is proposed in this chapter for estimating the overall image quality, i.e. regardless of the type of distortions present in the images. It uses a combination of features for the quality estimation of images.

The features extracted and used in the previous experimentations are combined to form a single large feature vector. The design of the proposed framework allows inclusion of new features to the repository to make it more robust to certain distortions or even to new distortions. However, with the inclusion of every feature, grows the size of the combined feature. Such a large set of features may adversely affect the performance of the framework as it requires more processing power to perform computations on large data, it requires more memory and feature space (storage), it slows down the computational efficiency, and most importantly it may cause overfitting of the neural network. In line with the objectives defined for the study for improving the performance of the framework, dimensionality reduction using PCA is thus proposed to tackle this “dimensionality curse” problem. Dimensional reduction of the combined feature metric helps to achieve the optimum performance from the framework. In other words, it helps in faster quality prediction without compromising on the quality prediction accuracy. It also helps in reducing the processing burden and storage requirement.

The feature optimized by PCA is used to train the framework for image quality estimation. The prediction models NFC and NWT are used for evaluating the performance of optimized feature metric.

7.1.1 Principal Component Analysis (PCA)

PCA is applied to reduce the dimensions of a dataset comprising of a large number of correlated variables while preserving the variations present in the dataset as much as possible. It is accomplished by transforming the large set of variables to a new small set of variables which are uncorrelated and ordered so that the first few variables preserve most of the variations present in all original variables [159].

PCA identifies the patterns present in the data and articulates the data in such a way that it emphasizes their similarities and differences. Since patterns in the data can be difficult to find in the data of high dimensions, PCA is used for analysing the data. The algorithm of PCA is given below.

Transform $N \times d$ matrix X into $N \times m$ matrix Y :

- Centralize the data – subtract the mean from each dimension. It results in a dataset with zero mean.
- Calculate the $d \times d$ covariance matrix as

$$C = \frac{1}{N-1} X^T X \quad (7.1)$$

$$ie., C_{i,j} = \frac{1}{N-1} \sum_{q=1}^N X_{q,i} \cdot X_{q,j} \quad (7.2)$$

where $C_{i,i}$ (diagonal) is the variance of variable i , and $C_{i,j}$ (off-diagonal) is the covariance between variables i and j .

- Eigenvectors of the covariance matrix are calculated.

m eigenvectors are selected to form the new basis, which corresponds to the largest m eigenvalues.

7.1.2 Methodology

A primary requirement of the framework is the feature repository built using the combination of purposefully selected and extracted features. The experimentations presented in previous chapters are carried out to estimate the image quality for five common types of distortions namely JPEG, JPEG2K, Gblur, white noise, and fast fading. Different features extracted and implemented in these experimentations are textural, block-based, SADWT and DWT, a fusion of SIFT key-points and curvelet coefficients, spatial covariance. All these features are combined to obtain a single feature vector to test the framework for abovementioned five common types of distortions.

The proposed framework allows inclusion of more features to its repository. Increase in the feature values leads to the high dimension feature dataset which may affect the process intensive supervised algorithms. Although the quality predictors used in the proposed framework can handle feature vector much larger than the one generated during the present study, the size of the feature vector may affect the performance of the framework. Dimensionality reduction is warranted in such a situation. Therefore, Principal Component Analysis (PCA) is employed to reduce

the size of the feature vector. It eliminates the redundant or less significant feature values. The reduced feature vector eases the processing burden and makes the framework lightweight and efficient. NFC and NWT, which are used as the quality predictors for the framework, are trained on the optimized feature vector. The methodology of the framework is presented in figure 7.1.

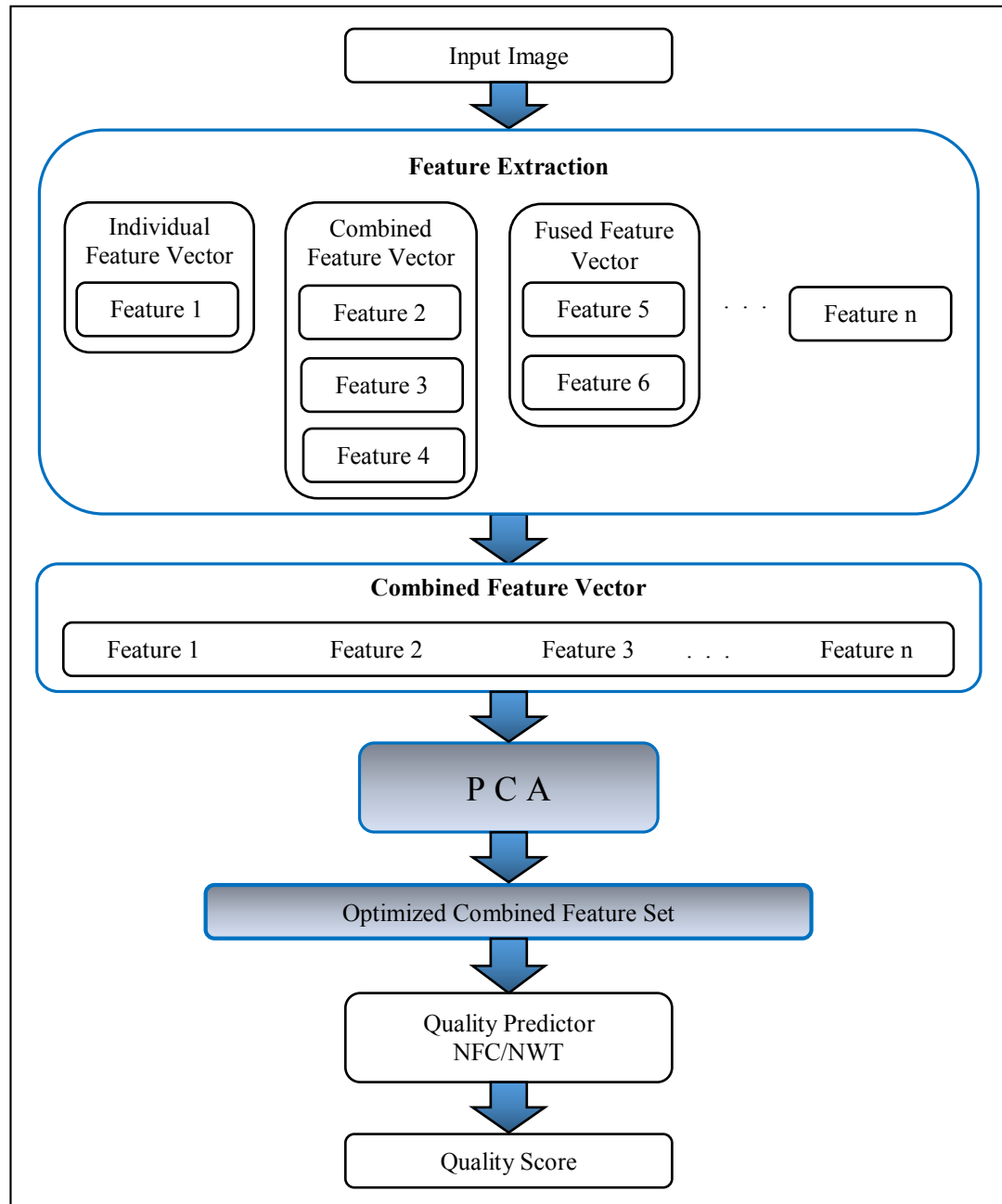


Figure 7.1: Schematic representation of the proposed framework

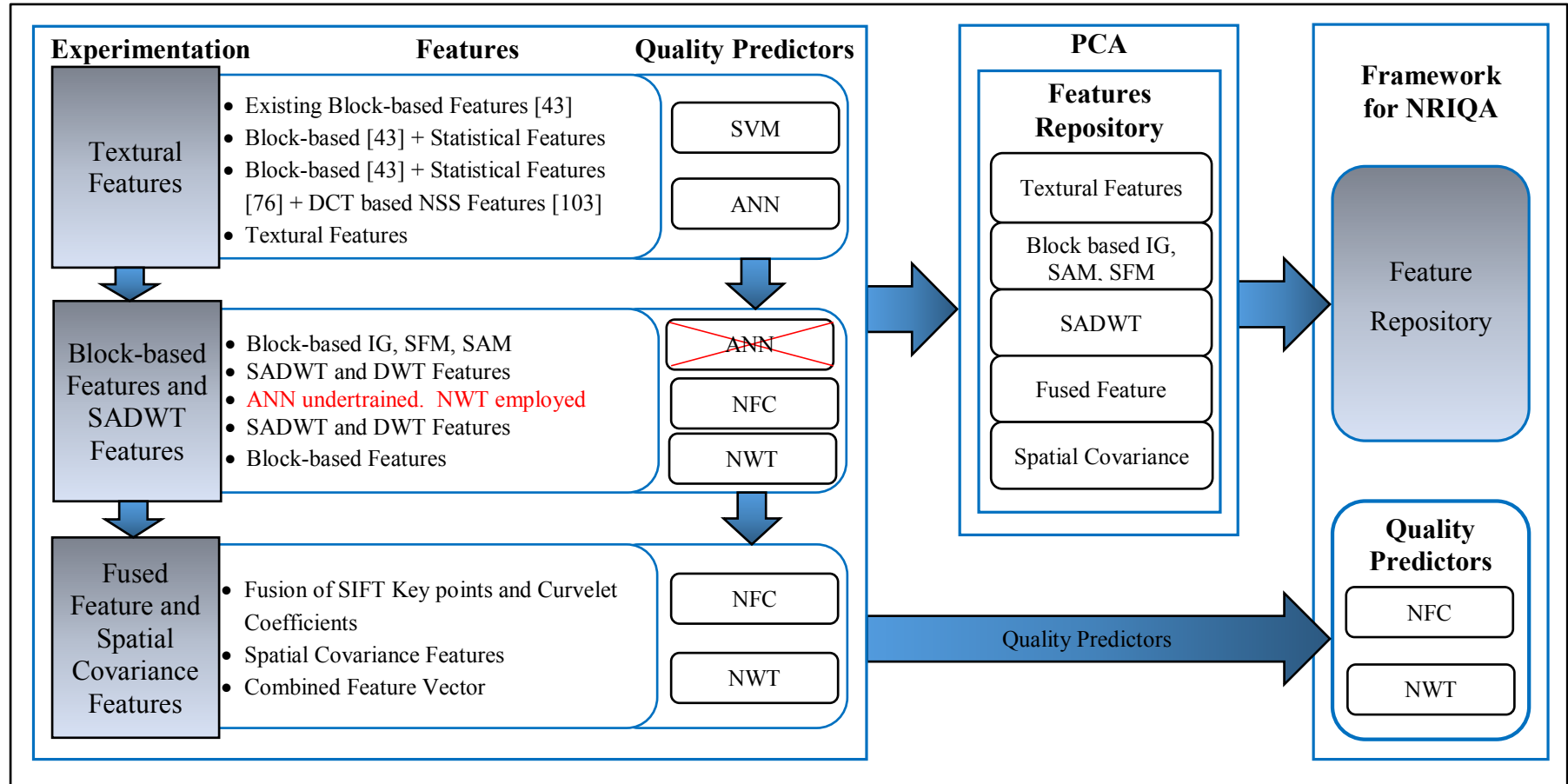


Figure 7.2: Evolution of the proposed framework

The framework comprises two essential ingredients – a feature repository and quality predictor/s. The proposed framework evolved through the experimentations presented in the previous chapters. The features implemented and tested in those experimentations formed the basis for the feature repository. SVM, ANN, NFC, and NWT are tested and validated for different features. While ANN remained undertrained as it failed to handle small variations in feature values of SADWT, NFC and NWT performed well and hence opted as the quality predictors for the framework. Figure 7.2 shows the evolution process of the proposed framework.

7.1.3 Performance Evaluation and Results

To test the framework, different features evaluated in the previous chapters are combined to form a single large feature vector. Performance of the combined feature metric is evaluated using LCC and SROCC. Standard image datasets LIVE, TID2008, and TID2013, each divided into separate sets of training and testing images as described in chapter 3, section 3.4.6 are used for all the experimentations throughout the study. The same training and testing datasets are used in the experimentations presented in this chapter.

Combination of the features produces a feature vector of size 170. PCA is applied to optimize the combined feature vector. The first ten principal components are used for quality prediction. The performance of the combined feature vector before applying PCA is compared with the performance of PCA-optimized combined feature vector. For the performance evaluation through LCC and SROCC of these feature metrics, NFC and NWT are employed. The results are presented in table 7.1(a)-(c) for the distorted images from LIVE, TID2008, and TID2013 datasets respectively. The results exhibit a good correlation between subjective score and predicted score. Regression plots of NWT before and after PCA for LIVE, TID2008, and TID2013 show satisfactory results with $(R=0.938, R=0.945)$, $(R=0.932, R=0.947)$, and $(R=0.934, R=0.987)$ respectively as shown in figure 7.3(a)-(c).

Table 7.1(a): Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on LIVE dataset

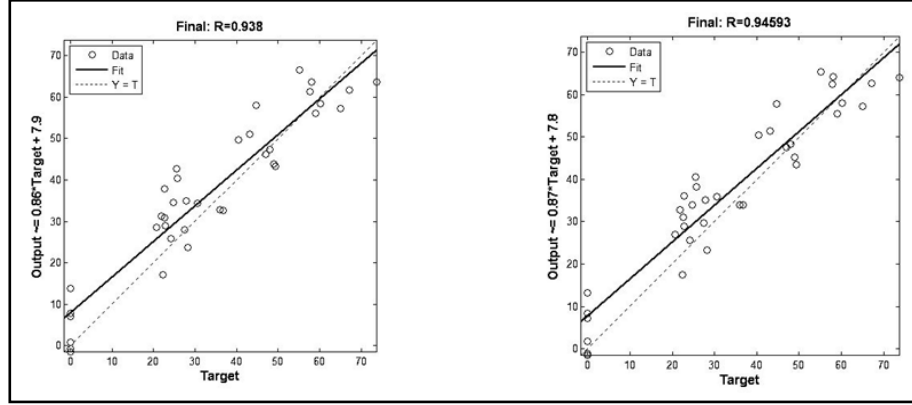
Distortion Type	NWT				NFC			
	LCC		SROCC		LCC		SROCC	
	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA
JPEG	0.981	0.988	0.964	0.973	0.945	0.954	0.934	0.948
JPEG2K	0.921	0.965	0.912	0.959	0.922	0.924	0.918	0.919
GBlur	0.913	0.951	0.900	0.943	0.908	0.925	0.901	0.918
WN	0.993	0.968	0.975	0.957	0.959	0.960	0.947	0.955
FF	0.968	0.969	0.941	0.959	0.948	0.953	0.938	0.948
Overall	0.938	0.945	0.925	0.939	0.964	0.968	0.958	0.960

Table 7.1(b): Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on TID2008 dataset

Distortion Type	NWT				NFC			
	LCC		SROCC		LCC		SROCC	
	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA
JPEG	0.962	0.939	0.956	0.929	0.946	0.955	0.937	0.945
JPEG2K	0.962	0.936	0.947	0.926	0.926	0.935	0.918	0.924
GBlur	0.921	0.932	0.911	0.927	0.9264	0.937	0.916	0.929
WN	0.961	0.968	0.954	0.957	0.944	0.955	0.937	0.947
Overall	0.932	0.947	0.929	0.935	0.965	0.968	0.954	0.959

Table 7.1(c): Performance evaluation of combined feature metric (before and after PCA) using NWT and NFC through LCC and SROCC on TID2013 dataset

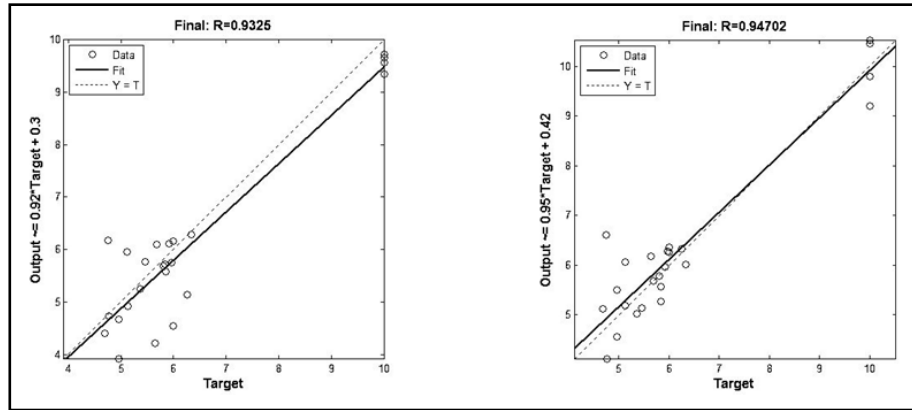
Distortion Type	NWT				NFC			
	LCC		SROCC		LCC		SROCC	
	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA	Before PCA	After PCA
JPEG	0.942	0.969	0.933	0.959	0.949	0.954	0.940	0.948
JPEG2K	0.952	0.946	0.943	0.939	0.914	0.925	0.905	0.918
GBlur	0.943	0.941	0.931	0.938	0.940	0.957	0.937	0.949
WN	0.983	0.963	0.979	0.959	0.944	0.959	0.934	0.950
Overall	0.934	0.984	0.927	0.968	0.966	0.968	0.957	0.959



(a) Before PCA

(b) After PCA

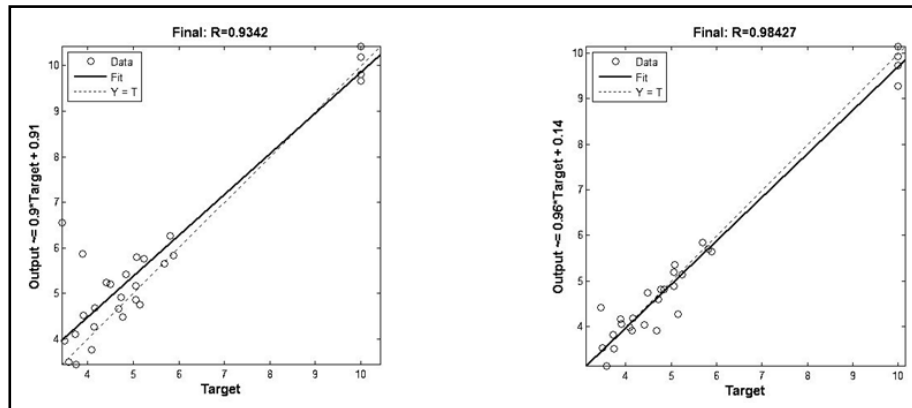
Figure 7.3(a): Regression plots of NWT using combined feature metric (a) before and (b) after dimension reduction by PCA for LIVE dataset. DMOS (target) are plotted against predicted scores (output).



(a) Before PCA

(b) After PCA

Figure 7.3(b): Regression plots of NWT using combined feature metric (a) before and (b) after dimension reduction by PCA for TID2008 dataset. MOS (target) are plotted against predicted scores (output).



(a) Before PCA

(b) After PCA

Figure 7.3(c): Regression plots of NWT using combined feature metric (a) before and (b) after dimension reduction by PCA for TID2013 dataset. MOS (target) are plotted against predicted scores (output).

It is observed from the above results presented in table 7.1(a)-(c) that the optimized feature metric exhibits good performance for overall quality estimation. It also perform reasonably well in quality prediction of images for specific distortions. However, few results obtained for individual distortions show lower performance in case of optimized feature metric. One such example is white noise images from LIVE by NWT. The potential reason for this decline in the performance can be attributed to the computation of principal components by PCA. PCA attempts to preserve the maximum information from the given feature data. It assigns highest variance to the feature values that have highest order of magnitude. In the process it skews the results. Hence, the quality prediction model may predict the quality with lower accuracy than the expected.

Experimental results exhibit the better performance by the proposed optimized metric when evaluated on TID2013 in comparison to TID2008. It should be noted that TID2013 image dataset has one additional level of distortion and 7 additional distortions types. It is important not only for the purpose of sampling the entire space but also for obtaining discriminative rating from the subjects.

The individual features implemented in the previous experimentations performed well for all the three image datasets. The performance of the optimized combined feature metric should necessarily be in the same range or improved to prove its practicability. Therefore, the performance of the dimensionally reduced combined feature metric and the performance of the individual feature metrics implemented in the previous chapters are compared in table 7.2(a)-(c). Comparison table shows that the performance of the reduced combined feature vector is in the comparable range or marginally improved. It proves the robustness of the reduced combined feature vector against the individual distortions also.

Table 7.2(a): Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on LIVE dataset

Distortion Type =>		JPEG	JPEG2k	Gaussian blur	White Noise	Fast Fading	Overall
Textural feature metric							
ANN	LCC	0.88	0.84	0.96	0.97	0.90	0.89
	SROCC	0.87	0.82	0.95	0.98	0.88	0.90
SVM	LCC	0.85	0.85	0.95	0.95	0.80	0.90
SADWT feature metric							
NFC	LCC	0.91	0.93	0.87	0.89	0.87	0.88
	SROCC	0.91	0.91	0.86	0.86	0.87	0.87
NWT	LCC	0.97	0.91	0.9	0.97	0.98	0.91
	SROCC	0.94	0.86	0.89	0.95	0.96	0.90
Fusion of SIFT key points and curvelet coefficients							
NFC	LCC	0.9	0.87	0.87	0.89	0.87	0.88
	SROCC	0.89	0.85	0.86	0.87	0.85	0.87
NWT	LCC	0.97	0.92	0.98	0.96	0.96	0.90
	SROCC	0.96	0.9	0.95	0.96	0.95	0.89
Spatial covariance feature metric							
NFC	LCC	0.93	0.9	0.87	0.92	0.92	0.9
	SROCC	0.92	0.9	0.86	0.91	0.92	0.89
NWT	LCC	0.96	0.94	0.89	0.96	0.94	0.94
	SROCC	0.94	0.93	0.89	0.95	0.94	0.92
Combination of fusion metric and spatial covariance feature metric							
NFC	LCC	0.92	0.89	0.92	0.91	0.92	0.88
	SROCC	0.9	0.89	0.9	0.9	0.91	0.87
NWT	LCC	0.94	0.93	0.94	0.97	0.97	0.90
	SROCC	0.93	0.91	0.93	0.95	0.96	0.90
Optimized combined feature metric							
NFC	LCC	0.954	0.924	0.925	0.960	0.953	0.968
	SROCC	0.948	0.919	0.918	0.955	0.948	0.960
NWT	LCC	0.988	0.965	0.951	0.968	0.969	0.945
	SROCC	0.973	0.959	0.943	0.957	0.959	0.939

Table 7.2(b): Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on TID2008 dataset

Distortion Type =>		JPEG	JPEG2k	Gaussian blur	White Noise	Overall
Textural feature metric						
ANN	LCC	0.87	0.87	0.94	0.92	0.93
	SROCC	0.88	0.87	0.91	0.92	0.93
SVM	LCC	0.85	0.85	0.92	0.91	0.92
SADWT feature metric						
NFC	LCC	0.9	0.91	0.86	0.89	0.87
	SROCC	0.88	0.9	0.86	0.88	0.87
NWT	LCC	0.97	0.95	0.97	0.95	0.96
	SROCC	0.95	0.92	0.9	0.92	0.93
Fusion of SIFT key points and curvelet coefficients						
NFC	LCC	0.92	0.91	0.92	0.94	0.91
	SROCC	0.92	0.9	0.92	0.92	0.9
NWT	LCC	0.98	0.94	0.94	0.97	0.94
	SROCC	0.96	0.93	0.92	0.96	0.93
Spatial covariance feature metric						
NFC	LCC	0.92	0.91	0.92	0.94	0.91
	SROCC	0.92	0.9	0.92	0.92	0.9
NWT	LCC	0.98	0.94	0.94	0.97	0.94
	SROCC	0.96	0.93	0.92	0.96	0.93
Combination of fusion metric and spatial covariance feature metric						
NFC	LCC	0.89	0.9	0.92	0.91	0.88
	SROCC	0.89	0.89	0.9	0.9	0.88
NWT	LCC	0.9	0.94	0.97	0.97	0.91
	SROCC	0.89	0.92	0.95	0.96	0.9
Optimized combined feature metric						
NFC	LCC	0.955	0.935	0.937	0.955	0.968
	SROCC	0.945	0.924	0.929	0.947	0.959
NWT	LCC	0.939	0.936	0.932	0.968	0.947
	SROCC	0.929	0.926	0.927	0.957	0.935

Table 7.2(c): Performance comparison of the optimized combined feature using NFC and NWT with the individual feature metrics implemented in the previous chapters through LCC and SROCC on TID2013 dataset

Distortion Type =>		JPEG	JPEG2k	Gaussian blur	White Noise	Overall
Textural feature metric						
ANN	LCC	0.88	0.85	0.93	0.94	0.91
	SROCC	0.87	0.82	0.94	0.96	0.90
SVM	LCC	0.85	0.85	0.91	0.92	0.90
SADWT feature metric						
NFC	LCC	0.92	0.91	0.88	0.89	0.88
	SROCC	0.92	0.9	0.87	0.88	0.88
NWT	LCC	0.95	0.96	0.97	0.95	0.98
	SROCC	0.94	0.91	0.92	0.91	0.95
Fusion of SIFT key points and curvelet coefficients						
NFC	LCC	0.9	0.88	0.9	0.9	0.92
	SROCC	0.89	0.86	0.89	0.88	0.91
NWT	LCC	0.97	0.94	0.92	0.94	0.96
	SROCC	0.95	0.93	0.91	0.92	0.94
Spatial covariance feature metric						
NFC	LCC	0.9	0.88	0.9	0.9	0.92
	SROCC	0.89	0.86	0.89	0.88	0.91
NWT	LCC	0.97	0.94	0.92	0.94	0.96
	SROCC	0.95	0.93	0.91	0.92	0.94
Combination of fusion metric and spatial covariance feature metric						
NFC	LCC	0.89	0.9	0.9	0.9	0.9
	SROCC	0.89	0.89	0.89	0.89	0.89
NWT	LCC	0.9	0.94	0.93	0.93	0.96
	SROCC	0.9	0.92	0.91	0.92	0.95
Optimized combined feature metric						
NFC	LCC	0.954	0.925	0.957	0.959	0.968
	SROCC	0.948	0.918	0.949	0.950	0.959
NWT	LCC	0.969	0.946	0.941	0.963	0.984
	SROCC	0.959	0.939	0.938	0.959	0.968

Accuracy and efficiency of any machine learning based NRIQA metric depends upon the relevance of the selected features. The features implemented in the previous experimentations are combined and evaluated at every stage. The framework is trained on the optimized combined feature vector and validated using NWT on the test images from all the three image datasets. Results presented in table 7.3 show that the performance is consistent at every stage.

Table 7.3: Performance comparison of the combined feature metrics formed at different stages of evolution of the framework using NWT through LCC on LIVE, TID2008, and TID2013 datasets

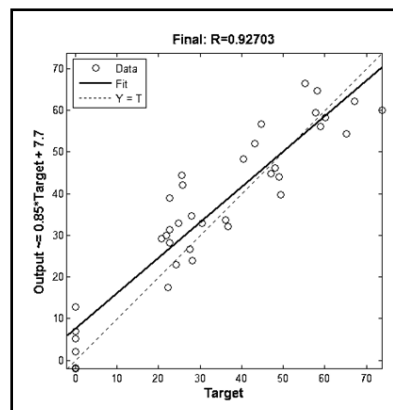
Combined Feature Vector	LIVE	TID2008	TID2013
Textural	0.89	0.93	0.92
Textural + Block based	0.88	0.89	0.91
Textural + Block based + SADWT	0.91	0.94	0.93
Textural + Block based + SADWT + (Fusion of SIFT key points and curvelet coefficients + spatial covariance)	0.94	0.93	0.98

The framework is further validated with the statistical features namely mean, standard deviation, variance, skewness, and kurtosis, implemented in chapter 4. The features are combined with the statistical features defined in the frequency domain. Statistical features in the wavelet domain are also computed and combined with statistical features and DCT based statistical features. Subsequently, GLCM based textural features [133, 134] and edge-based features [74] are combined to form a feature vector. These features are discussed in appendix-I. The obtained feature vector is further combined with existing block based [43], second-order statistical [76], and NSS based [103] features which are also discussed in chapter 4. All the above combinations are evaluated using NWT after applying PCA. Finally, the combined feature vector of size 80 is also reduced by applying PCA and tested with NWT. The results are presented in table 7.4 for LIVE, TID2008, and TID2013 datasets respectively. It can be evident from the results that the performance of the framework is consistent for the combinations of features distinguished by different approaches such as distortion specific and NSS-based, and the domains such as frequency and spatial. All these features are statistical features. Regression plots of NWT for the

combined feature vector is shown in figure 7.4 which exhibit a good correlation between subjective and objective score with $R=0.927$, $R=0.942$, and $R=0.985$ for LIVE, TID2008, and TID2013 respectively.

Table 7.4: Performance comparison of the existing combined statistical feature metrics (Appendix-I) using NWT through LCC on LIVE, TID2008, and TID2013 datasets

Features	LIVE	TID2008	TID2013
Statistical	0.91	0.94	0.95
Statistical + DCT based statistical	0.90	0.89	0.94
Statistical + DCT based statistical + Wavelet statistical	0.91	0.92	0.95
Statistical + DCT based statistical + Wavelet statistical + Textural	0.90	0.92	0.95
Statistical + DCT based statistical + Wavelet statistical + Textural + Edge based	0.91	0.91	0.94
Statistical + DCT based statistical + Wavelet statistical + Textural + Edge based + Block based	0.91	0.90	0.96
Statistical + DCT based statistical + Wavelet statistical + Textural + Edge based + Block based + Second order statistical	0.90	0.91	0.96
Statistical + DCT based statistical + Wavelet statistical + Textural + Edge based + Block based + Second order statistical + NSS based	0.92	0.94	0.98



(a)

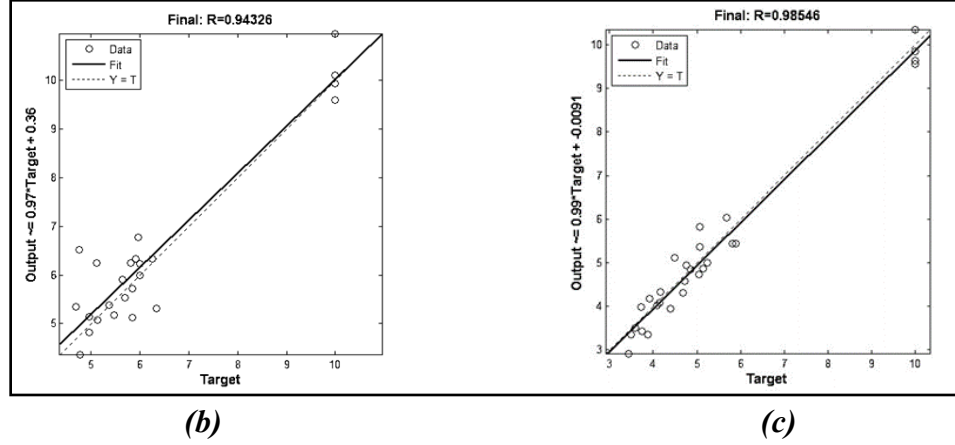


Figure 7.4: Regression plots of NWT using existing combined statistical features (given in Appendix-I)

- (a) DMOS (target) are plotted against predicted score (output) for LIVE and**
(b) MOS (target) are plotted against predicted score (output) for TID2008
(c) MOS (target) are plotted against predicted score (output) for TID2013

The proposed framework is designed to work with grayscale images. Because of the unavailability of the frameworks that match with the characteristics of the proposed framework, its performance is benchmarked against the state-of-the-art NSS based distortion generic frameworks developed for the quality estimation of colour images.

Most of the distortion-generic IQA methods developed by the researchers are NSS based which estimate the image quality by measuring the deviations from NSS models that capture the statistical “naturalness” of non-distorted images. Some such state-of-the-art frameworks are BIQI, BRISQUE, BLIINDS-I and BLIINDS-II. These frameworks usually work in two stages. In the first stage, it tries to identify each distortion present in the image and then predict the quality in the second stage. The machine learning based NRIQA framework proposed through this study is also distortion generic as it estimates the overall image quality without prior knowledge of distortions present in an image. The framework achieves this by means of collection and combination of the selected different features relevant to the different distortions. The performance comparison of the proposed framework with state-of-the-art frameworks is presented in the following table 7.5. It is observed from the table that the performance of the proposed framework is comparable with the existing state-of-the-art frameworks. The observations are encouraging and support the hypothesis of the framework.

Table 7.5: Performance comparison of the proposed framework using NFC and NWT with state-of-the-art frameworks through LCC and SROCC on LIVE, TID2008, and TID2013 datasets

Metric	LIVE		TID2008		TID2013	
	LCC	SROCC	LCC	SROCC	LCC	SROCC
BIQI	0.825	0.819	0.870	0.843	NA	NA
DIVIINE	0.865	0.859	0.903	0.893	0.631	0.529
C-DIVIINE	0.900	0.894	NA	NA	0.669	0.556
BLIINDS-II	0.927	0.921	0.921	0.898	0.698	0.607
BRISQUE	0.931	0.925	0.939	0.935	0.682	0.597
I. Nizami <i>et al.</i> [113]	0.962	0.961	0.951	0.942	NA	NA
Proposed (NFC)	0.968	0.960	0.968	0.959	0.968	0.959
Proposed (NWT)	0.945	0.939	0.947	0.935	0.984	0.963

7.2 Summary

The framework is deployed with machine learning based approach to estimate overall quality of an image regardless of the distortion type. The framework achieves distortion generic characteristic by combining the features to address different distortions. The combination of feature values raises the dimensionality curse problem. The combined feature vector is thus optimised by applying PCA. The performance of the optimized combined feature metric is evaluated using NWT and NFC. Results show that the quality score predicted by the proposed framework is consistent with the subjective score. The results of the combined feature metric and dimensionally reduced feature metric are compared. Combined feature metric shows performance gain after the optimization.

Performance of the optimised combined feature metric is compared with the individual feature metrics implemented in the previous chapters. It is observed that the results are consistent which proves the robustness of the optimised feature metric for the individual distortions as well.

Chapter 8.

Conclusion and Future Scope

NRIQA has been an area of interest for the researchers. Since last couple of decades, researchers have carried out intensive research in this field and proposed a number of different approaches and algorithms for quality assessment of the digital images. However, the growth of generalized NRIQA methods has accelerated recently. The generalized or generic NRIQA methods attempt to estimate image quality irrespective of the types of distortions present in the image. Most of the generalized NRIQA methods, which follow NSS approach, heavily use the transforms and hence time consuming, computationally complex, and resource hungry. On the other hand, machine learning based IQA methods are lightweight, less complex, and less resource demanding. The machine learning based image quality predictors are required to be trained on the features extracted from the training images prior to the quality prediction.

8.1 Summary

The framework proposed in this thesis follows the machine learning based approach for the quality prediction of digital images. NWT and NFC are employed as the quality predictors for the framework. As no single feature is applicable to the different distortions, so the combination of the different features is proposed to train the quality predictors. Therefore, different features have been extracted, and examined individually and in combinations. The results show that the combinations of the features work well for individual distortions and prediction of the overall quality as well. It proves the usability of this approach for preparing the features repository for the framework. The observations of such experimentations are summarized below.

Study of an existing NRIQA metric for blockiness shows that quality prediction using mathematical model sometimes lead to the incorrect prediction. The model parameters used by this mathematical model are same for all the test images. These model parameters are estimated using subjective scores as a training dataset. Thus mathematical models sometimes do not establish a correlation between subjective score, extracted features, and objective score. This relationship is nonlinear and

composite for the computation. The machine learning method has the potential to address this problem. ANN is used to predict the quality of an image with 96% accuracy.

Brightness, contrast, local amplitude, and texture of an image can be affected by the compression. For the assessment of these distortions, first order and second order features such as mean, standard deviation, variance, skewness, kurtosis, and entropy are calculated in the frequency domain along with block based features. The results show 95% accuracy. These experiments are carried out only for specific distortions.

Combination of different features for the quality assessment of different distortions is implemented to support the hypothesis for this framework. Implementation of this metric using a combination of block based, second-order histogram features obtained from co-occurrence and frequency based features derived from the energy of frequency content, and NSS based features shows promising results. The linear correlation between subjective and objective scores is 0.92 for LIVE dataset. From the results, it is observed that the performance of this metric is slightly lower for fast fading type of distortion.

GLCM based textural features are unable to handle isotropic images due to their heavy dependency on the directional information. Therefore, textural features are implemented using energy maps. Noise present in an image also affects the texture. GLCM heavily depends on the magnitude of gray level differences but fall short in handling noise and change in luminance. Therefore, pattern strength is used to quantify noise in the image. Statistical features are computed for energy maps and pattern strength for the quality evaluation. This metric shows good correlation between subjective score and objective score. The linear correlation coefficient for the images from LIVE, TID2008, and TID2013 datasets are 0.89, 0.93, and 0.91 respectively. This metric shows better accuracy when compared with existing NR metrics; however, it shows slightly lower performance for JPEG and fast fading.

Hence, to quantify the blockiness in JPEG images, block based features, namely image gradient, spatial frequency measure, and spectral activity measure are implemented. The accuracy of quality prediction using ANN classifier is 94% for JPEG images from LIVE dataset.

Understanding the functionality of ANN is difficult as it hides the relation between inputs and outputs in the weights of the neurons behind its hidden layers. Determination of number of layers and number of neurons of an ANN is also difficult. Therefore, the neurofuzzy classifier is implemented which applies the fuzzy system to symbolize knowledge in an understandable way and learning capability of ANN to optimize its parameters. This classifier is tested with block based features, and it shows good accuracy.

An image may consist of different objects which are also subjected to the distortions. SADWT is applied to analyse the shape of objects and compute the wavelet coefficients. Standard deviation, mean, SVD, and shape features are extracted. From the extracted feature values, it has been observed that all feature output magnitudes are not in the specified range and the variation with respect to different distortion types do not follow a linear relationship. Fitting the features and the subjective score becomes complex as artificial neural networks do not handle these variations. Therefore, despite sufficient training data, ANN remains undertrained. To overcome this drawback, NWT model is used for the quality prediction. Results obtained using NWT classifier show better accuracy when compared with existing wavelet and curvelet based metrics.

The fusion feature vector is developed by combining the key points identified through SIFT and curvelet coefficients. Spatial covariance is also extracted as another feature. Spatial covariance calculates the relationship between neighbouring pixels. This feature is found useful in estimating the quality of compressed, noisy, and blurred images. The fusion and spatial covariance feature metric are combined and evaluated using NWT and NFC classifiers. The results show better performance over existing NSS based metrics.

Finally, the features implemented and evaluated during the study are combined to serve as a feature metric for quality estimation of the images. This combined feature metric is optimized by dimensionality reduction using PCA. The optimized feature metric is evaluated using NWT and NFC on LIVE, TID2008, and TID2013 datasets. Performance of the optimized combined feature metric is measured through LCC and SROCC. The optimized feature metric exhibits a good correlation between subjective and objective scores. The results show that it predicts the overall quality with quite a good accuracy. It also works well for the different individual distortions. Its

performance is found consistent when compared with the performance of other metrics that are evaluated individually. It indicates the robustness of the optimized feature vector to different distortions.

8.2 Conclusion

The proposed framework uses a combination of different features to form a feature vector which is used for quality prediction. Combination of features increases the size of the feature vector which, at a certain point, would raise the dimensionality curse problem and may degrade the computational efficiency of the framework. Therefore, to better the computational efficiency of the framework without compromising on the quality prediction accuracy, the combined feature metric is optimized using PCA. The optimized feature vector is then stored in the features repository which is used by the framework for quality prediction.

The framework facilitates the use of “ready to use” feature metric. If the images are known to suffer from certain possible distortions, an optimized feature metric created by combining different relevant features can be placed into the repository. This feature metric is then used for quality prediction of those set of images. This adaptability of the framework opens up the possibility to employ the framework for quality estimation in such devices where the images are to be assessed for certain possible distortions. An optimized and small set of feature values make “Over-The-Air” (OTA) update easy and quick to enhance the performance of the imaging devices.

Being lightweight and less resource demanding, the framework can be ported to the imaging devices or the devices that have imaging functionality. Also, it just accepts any optimized feature thrown to it, gets trained, and start predicting the image quality. The framework acquires Plug-n-Play functionality through these characteristics.

8.3 Future Scope

The study presented in this thesis has focused on proposing a framework for NRIQA. The study is limited to grayscale images and five common types of distortions namely JPEG, JPEG2000, Gaussian blur, white noise, and fast fading. Below are the few suggestions for improving upon the work presented in this thesis.

- The feature-rich repository is the primary requirement of the framework for quality evaluation with good accuracy. Collection of appropriate features relevant to different distortions is necessary to make the framework capable of handling different distortions. Hence, it is suggested to build the useful repository by analysing and extracting the appropriate features, and examining their viability for predicting the image quality.
- Colour information is crucial for quality estimation of colour images. The distortion to the colour image affects the chromatic information which is essential for the accurate quality prediction of a colour image. Grayscale based IQA metrics ignore this chromatic information and underestimate the degradation. Therefore, it is essential to update the feature-repository with appropriate colour features so that the framework can accurately estimate the quality of colour images.
- The framework proposed in the thesis for estimating the image quality is designed based on the assumption that the image is affected by only one type of distortion. Hence, it is validated on the public image datasets, namely LIVE, TID2008, and TID2013, which comprise such distorted images. However, a digital image may acquire multiple distortions during various stages it undergoes such as acquisition, compression, transmission, and display. Though, the proposed framework estimates the image quality irrespective of the distortions present in the image, its applicability for quality estimation of the multi-distorted images has not been examined. Therefore, it is suggested to test the framework on image datasets such as “*LIVE Multiply Distorted Image Quality Database (2012)*”, “*Multiply Distorted Image Database (MDID) (2016)*”. The work may be further extended by implementing the appropriate features for estimating the quality of multi-distorted images.
- The proposed framework does not identify the distortions present in the image. However, the framework can be further improved so that it can identify and report the types and degrees of the distortions present in the image. It will help in the automatic selection of appropriate features or feature sets and thereby improve the efficiency and accuracy of the framework.

References

1. M. Joshi, M. Raval, Y. Dandawate, K. Joshi, S. Metkar, "Image and Video Compression: Fundamentals, Techniques, and Applications", *CRC Press, first edition*, 2014.
2. Z. Wang, A.C. Bovik, "Why is Image Quality Assessment so Difficult", *IEEE Signal Processing Letters*, vol. 4, pp.3313-3316, September 2002.
3. J. Ruikar, A. Sinha, S. Chaudhary, "Image quality assessment algorithms: study and performance comparison", *International Conference on Electronic and Communication Systems*, Coimbatore, India, 13-14 February 2014.
4. Z. Wang, A. C. Bovik, "A lecture book on Modern Image Quality Assessment", *Morgan and Claypool edition publisher*, 2006.
5. B. W. Keelan, "Handbook of image quality, characterization and prediction", *Marcel Dekker Inc.*, ISBN 0-8247-0770-2.
6. M. Oszust, "Full reference image quality assessment with linear combination genetically selected quality measures", *PLoS ONE*, vol. 11, June 2016, DOI:10.1371/journal.pone.0158333.
7. I. Avcibas, B. Sankur, K. Sayood, "Static evaluation of image quality measures", *Journal of Electronic Imaging*, vol. 11, pp 206-223, April 2010.
8. A. M. Eskicioglu, P. S. Fisher, "Image quality measures and their performance", *IEEE Transactions on Communications*, vol.43, no. 12, pp. 2959-2965, December 1995.
9. A. Srivastava, A. B. Lee, E. P. Simoncelli, S. C. Zha, "On advances in the statistical modelling of natural images", *Journal of Mathematical Imaging and Vision*, vol. 18, no. 1, pp. 17-33, January 2003.
10. N. Kendrick, T. Kuy, R. J. Naselaris, J. Prenger, L. Gallant, "Identifying natural images from human brain activity", *Nature Letter*, vol. 452, pp. 352-356, March 2008.
11. E. P. Simoncelli, B. A. Olshausen, "Natural image statistics & neural representation", *Annual Review of Neuroscience*, vol. 24, May 2001.
12. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.
13. Z. Wang, A. C. Bovik, "A universal image quality index", *IEEE Signal Processing Letters*, vol. 9, no.3, pp. 81-84, March 2002.

14. J. Ruikar, A. Sinha, S. Chaudhary, "Image quality assessment using edge correlation", *International Journal of Electronics & Telecommunications*, vol. 63, pp. 99-107, January 2017.
15. R. Soundararajan, A. C. Bovik, "RRED Indices: Reduced reference entropic differencing for image quality assessment," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 517-526, February 2012.
16. Q. Li, Z. Wang, "Reduced-reference image quality assessment using divisive normalization based image representation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 202-211, April 2009.
17. A. Rehman, Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation", *IEEE Transactions on Image Processing*, vol. 21, no.8, pp. 3378-3389, May 2012.
18. R. Herzog, M. Cadik, T. O. Aydin, K. I. Kim, K. Myszkowski, H. Seidel, "NoRM: No-reference image quality metric for realistic image synthesis", *Computer Graphics Forum*, vol. 31, pp. 545-554, May 2012.
19. M. A. Saad, A. C Bovik, C. Charier, "A DCT statistics-based blind image quality index", *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 583-586, June 2010.
20. X. Zhang, X. Feng, W. Wang, W. Xue, "Edge strength similarity for image quality assessment", *IEEE Signal Processing Letters*, vol. 20, pp. 319-322, April 2013.
21. B. Liao, Y. Chen, "An image quality assessment algorithm based on dual-scale edge structure similarity", *Second International Conference on Innovative Computing, Information and Control*, pp. 56-59, Kumamoto, Japan, 5-7 September 2007.
22. B. Wang, Z. Wang, Y. Liao, X. Lin "HVS based structural similarity for image quality assessment", *Proceedings of IEEE International Conference on Signal Processing*, pp. 1194-1197, Beijing, China 26-29 October 2008.
23. C. Li, A. C. Bovik, "Content-partitioned structural similarity index for image quality assessment", *Signal Processing: Image Communication Journal*, vol. 25, pp. 517-526, August 2010.
24. C. L. Yang, W. R. Gao, L. M. Po, "Discrete wavelet transform-based structural similarity for image quality assessment", *IEEE International Conference on Image Processing*, pp. 377-380, San Diego, CA, USA, 12-15 October 2008.
25. W. Xue, L. Zhang, X. Mou, A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index", *IEEE Transactions on Image Processing*, vol. 23 no. 2, pp. 684-695, February 2014.

26. J. Zhao, H. Feng, Z. Xu, Q. Li, "An image quality assessment based on regional weight", *International Journal for Light and Electron Optics*, vol. 123, no. 6, Pages 494–497, March 2012.
27. L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment", *Transactions on Image Processing*, vol. 20, pp. 2378-2386, January 2011.
28. A. Shnayderman, A. Gusev, A. M. Eskicioglu, "An SVD based grayscale image quality measure for local global assessment", *IEEE Transactions on Image Processing*, vol. 15, pp. 422–429, January 2006.
29. A. Shnayderman, A. Gusev, A. M. Eskicioglu, "An SVD based grayscale image quality measure for local and global assessment", *Proceedings of SPIE Image Quality and System Performance*, vol. 5492, pp. 82-92, February 2004.
30. S. Wang, C. Deng, W. Lin, B. Zhao, J. Chen, "A novel SVD-based image quality assessment metric", *Proceedings of 20th IEEE International Conference on Image Processing*, pp. 423-426, Melbourne, VIC, Australia, 15-18 September 2013.
31. M. Pedersen, J. Hardeberg, "Full-reference image quality metric: classification and evaluation", *Foundations and Trends in Computer Graphics and Vision*, vol. 7, pp. 1-80, January 2012, DOI: 10.1561/06000000037.
32. A. Chetouani, A. Beghdadi, M. Deriche, "Statistical modelling of image degradation based on quality metrics", *20th International Conference on Pattern Recognition*, pp. 714-717, Istanbul, Turkey, 23-26 August 2010.
33. A. Chetouani, A. Beghdadi, M. Deriche, "A hybrid system for distortion classification and image quality evaluation", *Signal Processing: Image Communication*, vol. 27, pp. 948-960, October 2012.
34. Z. Wang, A. C. Bovik, "Reduced and no reference image quality assessment", *IEEE Signal Processing Magazine*, pp. 29-40, November 2011.
35. L. Ma, S. Lee, K. N. Ngan, "Reduced-reference image quality assessment in recognized DCT domain", *Signal Processing: Image Communication*, vol. 28, no. 8, pp. 884-902, September 2013.
36. X. Li, D. Tao, X. Gao, W. Lu, "A Natural image quality evaluation metric", *Journal of Signal Processing*, vol. 89, pp. 548-555, April 2009.
37. G. Cheng, L. Cheng, "Reduced reference image quality assessment based on dual derivative priors", *Electronic Letters*, vol. 45 no 18, pp. 937-939, September 2009.
38. G. Cheng, J. Huang, Z. Liua, C. Lizhi, "Image quality assessment using natural image statistics in gradient domain", *International Journal of*

Electronics and Communications (AEÜ), vol. 65, no 5, pp. 392–397, May 2011.

39. Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. H. Yang, A. C. Bovik, “Quality aware images”, *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1680-1689, May 2006.
40. A. K. Moorthy, A. C. Bovik, “Visual quality assessment algorithms: What does the future hold?” *International Journal of Multimedia Tools and Applications*, vol. 51, no. 2, pp. 675-696, February 2011.
41. N. H. B. Ismail, Soong Der Chen, L. S. Ng, A. R. Ramli, “Analysis of image quality assessment algorithm to detect the presence of unnatural contrast enhancement”, *Journal of Theoretical and Applied Information Technology*, vol. 83, pp. 415-422, January 2016.
42. V. Kamble, K. M. Bhurchandi, “No reference image quality assessment algorithm: A survey”, *Optik International Journal for Light & Electron Optics*, vol. 126, pp. 1090-1097, June 2015.
43. Z. Wang, H. R. Sheikh, A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images", *Proceedings of IEEE International Conference on Image Processing*, vol. 1, pp. I-477-I-480, Rochester, NY, USA, 22-25 September 2002.
44. S. Suthaharan, “No-reference visually significant blocking artifact metric for natural scene images”, *Signal Processing*, vol. 89, pp. 1647-1652, August 2009.
45. S. Lee, S. J. Park, “A new image quality assessment method to detect and measure strength of blocking artifacts”, *Journal on Signal Processing: Image Communication*, vol. 27, pp. 31-38, January 2012 .
46. C. Chen, J. Bloom, “A blind-reference free blockiness measures”, *Pacific-Rim conference on Advances in multimedia information processing –PCM 2010*, lecture notes in computer science, Springer, Berlin, Heidelberg, vol. 6297, pp. 112-123, 2010.
47. S. Amiri, H. Hassanpor, O. Marouzi, “No-reference image quality assessment based on localised discrete cosine transform for JPEG compressed images”, *Multimedia Tools and Applications*, vol. 77, pp. 787-803, January 2017. <https://doi.org/10.1007/s11042-016-4246-9>.
48. Y. Zhan, R. Zhang, “No reference JPEG quality assessment based on blockiness and luminance change”, *IEEE Signal Processing Letters*, vol. 24, pp. 760-764, March 2017.
49. Z. M. Parvez Sazzad, Y. Kawayoke, Y. Horita, “No reference image quality assessment for JPEG 2000 based on spatial features”, *Journal of Signal Processing: Image Communication*, vol. 23, no. 4, pp. 257-268, April 2008.

50. L. Liang, S. Wang, J. Chen, S. Ma, , D. Zhao, W. Gao, "No reference perceptual image quality metric using gradient profile for JPEG 2000", *Signal Processing: Image Communication*, vol. 25, no.7, pp. 502-516, August 2010.
51. J. Zhang, M. L. Thinh, S. H. Ong, T. Q. Nguyen, "No-reference image quality assessment using structural activity", *Signal Processing*, vol. 91, no. 11, pp. 2575-2588, November 2011.
52. J. Zhang, S. H. Ong, M. Le Thinh, "Kurtosis based no reference quality assessment of JPEG2000 images", *Signal Processing: Image Communication Journal*, vol. 26, no. 1, pp. 13-23, January 2011.
53. H. Liu, N. Klomp, I. Henyderickx, "A No-reference metric for perceived ringing artifacts in images", *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 20, pp. 529-539, April 2010.
54. Fatin E. M. Al-Obaidi, "Image quality assessment for defocused blur images", *American Journal of Signal Processing*, vol. 5 No. 3, pp. 51-55, May 2015, doi: 10.5923/j.ajsp.20150503.01.
55. P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "A no reference perceptual blur metric", *Proceedings of IEEE International Conference on Image Processing*, vol. 3, pp. III-57-III-60, Rochester, NY, USA, 22-25 September 2002.
56. P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi "Perceptual blur and ringing metrics: application to JPEG 2000", *Signal Processing: Image Communication Journal*, vol. 19, pp. 163-172, February 2004.
57. A. Chetouani, A. Beghdadi, M. Deriche, "A new reference-free image quality index for blur estimation in the frequency domain", *IEEE International Symposium on Signal Processing and Information Technology*, pp. 155-159, December 2009.
58. R. Ferzli, L. Karam, "A No-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)", *IEEE Transactions on Image Processing*, vol. 18, no. 4, pp. 717-728, April 2009.
59. N. Narvekar, L. J. Karam, "A No-Reference image blur metric based on the Cumulative Probability of Blur Detection (CPBD)", *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2678-2683, September 2011.
60. F. Kerouh, A. Serir, "A no reference blur image quality measure based on wavelet transform", *Proceedings of Second International Conference on Digital Information Processing and Communications (ICDIPC)*, pp. 127-132, Klaipeda City, Lithuania, 10-12 July 2012.
61. F. Kerouh, A. Serir, "A no reference quality metric for measuring image blur in wavelet domain", *International Journal of Digital Information and Wireless Communications*, vol. 1, pp. 767-776, January 2012.

62. L. Li, W. Lin, X. Wang, G. Yang, K. Bharami, A. Kot, "No reference image blur assessment based on discrete orthogonal moments", *IEEE Transactions on Cybernetics*, vol. 46, no. 4, pp. 39-49, January 2016.
63. C. Lim Paramesran, W. Jassim, Y. Yu, K. Ngan, "Blind image quality assessment for Gaussian blur images using exact Zernike moments and gradient magnitude", *Journal of The Franklin Institute*, vol. 353, pp. 4715-4733, September 2016.
64. C. Vu, T. Phan, D. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images", *IEEE Transactions on Image Processing*, vol. 21, pp. 934-945, September 2011.
65. P. Vu, D. Chandler, "A fast wavelet-based algorithm for global and local image sharpness estimation", *IEEE Signal Processing Letters*, vol. 9, pp. 423-426, May 2012.
66. R. Hassen, Z. Wang, M. Salama, "No reference image sharpness assessment based on local phase coherence measurement", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 10)*, Dallas, TX, USA, 14-19 March 2010.
67. E. Marvidaki, V. Mezaris, "No reference blur assessment in natural images using Fourier transform and spatial pyramids", *IEEE International Conference on Image Processing*, pp. 566-570, Paris, France, 27-30 October 2014.
68. M. A. Qureshi, M. Deriche, A. Beghdadi, "Quantifying blur in color images using higher order singular values", *IET Electronics Letter*, August 2016, DOI: 10.1049/el.2016.1792.
69. G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, "No-reference noticeable blockiness estimation in images", *Signal Processing*, vol. 23, pp. 417-432, July 2008.
70. T. Brandao, M. P. Queluz, "No-reference image quality assessment based on DCT domain statistics", *Signal Processing*, vol. 88, no. 4, pp. 822-833, April 2008.
71. J. Shen, Q. Li, G. Erlebacher, "Hybrid no-reference natural image quality assessment of noisy, blurry, JPEG2000, and JPEG images", *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2089-2098, August 2011.
72. J. Bagade, Y. Dandawate, Kulbir Singh, "No reference image quality assessment using block based features and artificial neural network", *Global Trends in Information Systems and Software Applications*, Springer Berlin Heidelberg Publisher, pp. 128-138, January 2012.
73. J. Bagade, Kulbir Singh, Y. Dandawate, "No reference image quality assessment using block-based and frequency domain statistical features:

A machine learning approach", *International Journal of Communication Networks and Distributed Systems*, vol. 12, no. 1, pp. 95-112, March 2014.

74. S. Suresh, R. Venkatesh Babu, H. J. Kim, "No reference image quality assessment using modified extreme learning machine classifier", *Journal on Applied Soft Computing*, vol. 9, pp. 541-552, March 2009.
75. R. V. Babu, S. Suresh, A. Perkis, "No-reference JPEG-image quality assessment using GAP-RBF", *Signal Processing*, vol. 87, pp. 1493-1503, June 2007.
76. P. Gastaldo, R. Zunino, I. Heynderickx, E. Vicario, "Objective quality assessment of displayed images by using neural networks", *Signal Processing: Image Communication*, vol. 20, no.7, pp. 643-661, August 2005.
77. A. Chetouani, A. Beghdadi, "A no-reference image quality metric for blur and ringing distortions based on a neural weighting process", *Image Quality and System Performance IX*, vol. 8293, pp. 82930D, January 12.
78. F. Shao, S. Gu, G. Jiang, M. Yu, "A novel no reference stereoscopic image quality assessment method", *Symposium on photonics and optoelectronics*, pp.1-4, May 2012.
79. C. Li, A.C. Bovik, X. Wu, "Blind image quality assessment using a General Regression Neural Network", *IEEE Transactions on Neural Networks*, vol. 22, no. 5, pp. 793-799, May 2011.
80. L. Liu, B. Liu, H. Huang, A. Bovik, "No-reference image quality assessment based on spatial and spectral entropies", *Signal Processing: Image Communication*, vol. 29, pp. 856-863, January 2014.
81. S. Wang, C. Den, B. Zhao, G. Huang, B. Wang, "Gradient-based no-reference image blur assessment using extreme learning machine", *Neurocomputing*, vol. 174, pp. 310-321, January 2016.
82. M. Zhang, C. Muramatsu, X. Zhou, T. Hara, H. Fujita, "Blind image quality assessment using the joint statistics of generalised local binary pattern", *IEEE Signal Processing Letters*, vol. 22, pp. 207-210, May 2015.
83. F. Rezaie, M. Helfroush, H. Danyali, "No-reference image quality assessment using local binary pattern in the wavelet domain", *Multimedia Tools and Applications*, vol. 77, no. 2, pp. 2529-2541, January 2018, <https://doi.org/10.1007/s11042-017-4432-4>.
84. P. Feritas, W. Akamine, M. Farias, "No-reference image quality assessment using texture information bank", *Brazilian Conference on Intelligence Systems*, Recife, Brazil, 9-12 October 2016.
85. P. Feritas, W. Akamine, M. Farias, "Blind image quality assessment using multiscale local binary patterns", *Journal of Imaging Science and Technology*, vol. 60, pp. 60405-1-60405-8, December 2016.

86. P. Feritas, W. Akamine, M. Farias, "No-reference image quality assessment based on statistics of Local Ternary Pattern", *2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal, 6-8 June 2016.
87. P. Freitas, W. Akamine, M. Farias, "No reference image quality assessment using salient local binary pattern", *IS&T International Symposium on Electronic Imaging, Image quality and System Performance*, pp. 367-1-367-6, January 2018.
88. S. Corchs, F. Gasparini, R. Schettini, "No reference image quality classification for JPEG-distorted images", *Digital Signal Processing*, vol. 30, pp. 86-100, 2014.
89. X. Shang, X. Zhao, Y. Ding, "Image quality assessment based on joint quality-aware representation construction in multiple domains", *Handawi Journal of Engineering*, vol. 2018, Article ID 1214697, February 2018.
90. L. Kang, P. Ye, Y. Li, D. Doermann, "Convolution neural networks for no-reference image quality assessment", *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 23-28 June 2014.
91. S. Bianco, L. Celona, P. Napoletano, R. Schettini, "On the use of deep learning for blind image quality assessment", *Signal Image and Video Processing*, vol. 12, no. 2, pp. 355-362, August 2017, DOI 10.1007/s11760-017-1166-8.
92. X. Lv, M. Qin, X. Chen, G. Wei, "No reference image quality assessment based on statistics of convolution of feature maps", *AIP Conference Proceeding 1995*, pp. 040034-1-040034-5, April 2018, <https://doi.org/10.1063/1.5033698>.
93. H. Zeng, L. Zhang, A. Bovik, "A probabilistic quality representation approach to deep blind image quality prediction", [online] <https://arxiv.org/abs/1708.08190>, August 2017.
94. A. Mittal, A. Moorthy, A. C. Bovik, "Blind/referenceless image spatial quality evaluator", *IEEE Conference on Signals, System and, Computers*, pp. 723-727, Pacific Grove, CA, USA, 6-9 November 2011.
95. A. Mittal, A. K. Moorthy, A. C. Bovik, "No reference image quality assessment in the spatial domain", *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4707 December 2012.
96. A. Mittal, R. Soundararajan, A. Bovik, "Making a completely blind image quality analyzer", *IEEE Signal processing letters*, vol. 20, no. 3, pp. 209-213, March 2013.
97. A. K. Moorthy, A. C. Bovik, "A two step framework for constructing blind image quality indices", *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513-516, May 2010.

98. Y. Zhang, A. Moorthy, D. Chandler, A. Bovik, "C-DIIVINE: No reference image quality assessment based on local magnitude and phase statistics of natural scenes", *Signal Processing: Image Communication*, vol. 29, no. 10, pp. 725-747, May 2014.
99. Y. Jhang, M. Damon, Chandler, "An algorithm for no-reference image quality assessment based on log-derivative statistics of natural scenes", *SPIE Proceedings: Image Quality and System Performance*, vol. 8653, pp. 86530J-86530J-10, February 2013.
100. D. Zhang, Y. Ding , N. Zheng, "Nature scene statistics approach based on ICA for no reference image quality assessment", *Proceedings of International Workshop on Information and Electronics Engineering (IWIEE)*, vol. 29, pp. 3589- 3593, 2012.
101. H. R. Sheikh, A. C. Bovik, L. K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1918–1927, November 2005.
102. M. J. Chen, A. C. Bovik, "No. reference image blur assessment using multiscale gradient", *Proceeding of IEEE Quality of Multimedia Experience*, pp. 70-74, San Diego, CA, USA , July 2009.
103. M. A. Saad, A. C. Bovik, C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain", *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339-3351, August 2012.
104. Y. Fang, K. Ma, Z. Wang, W. Lin, Z. Fang, G. Zhai, " No reference quality assessment of contrast distorted images based on natural scene statistics", *IEEE Signal Processing Letters*, vol. 22, no. 7, pp. 838-842, July 2015.
105. K. Gu, G. Zhai, X. Yang, W. Zhang, " Using free energy principle for blind image quality assessment", *IEEE Transactions on Multimedia*, vol. 17, pp. 50-63, January 2015.
106. L. Zhang, L. Zhang, A. Bovik, "A feature-enrich completely blind image quality evaluator", *IEEE Transactions on image Processing*, vol. 24, pp. 2579-2591, August 2015.
107. L. Liu, H. Dong, H. Huang, A. Bovik, "No reference image quality assessment in curvelet domain", *Signal Processing: Image Communication*, vol. 29, no. 4, pp. 494-505, March 2014.
108. W. Lu, K. Zeng, D. Tao, Y. Yuan, X. Gao, "No-reference image quality assessment in contourlet", *Neurocomputing*, vol. 73, no. 4-6, pp. 784-794, January 2010.
109. G. Wang, Z. Wu, H. Yan, M. Cui, " No reference image quality assessment based on non-subsample shearlet transform and natural scene statistics", *Optoelectronics Letters*, vol. 12, no. 2, pp. , March 2016.

110. L. Li, D. Wu, J. Wu, J. Qian, B. Chen, “ No reference image quality assessment with a gradient-induced dictionary”, *KSII Transactions on Internet and Information*, vol. 10, no. 1, pp. 288-306, January 2016.
111. T. Feng, D. Deng, J. Yan, W. Zhang, W. Shi, I. Zou, “ Sparse representation of salient regions for no reference image quality assessment”, *International Journal of Advance Robotics Systems*, October 2016, DOI: 10.1177/1729881416669486.
112. M. Qin, X. Lv, X. Chen W. Wang, “Hybrid NSS features for no-reference image quality assessment”, *IET Image Processing*, vol. 11, pp. 443-449, June 2017.
113. I. Nizami, M. Masid, K. Khursid, “Feature selection for no-reference image quality assessment using natural scene statistics”, *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 26, pp. 2163-2177, September 2018.
114. C. Fan, Y. Zhang, L. Feng, A. Jiang, “ No reference image quality assessment based on multi-expert convolution neural network”, *IEEE Access*, vol. 6, pp. 8934 – 8943, February 2018.
115. M. Shahid, A. Rossholm, B. Lovstrom, H. Zepernick, “No reference image and video quality assessment: a classification and review of recent approaches”, *EURASIP Journal on Image and Video Processing*, August 2014.
116. D. M. Chandler, “Seven challenges in image quality assessment: Past, present, and future research”, *ISRN Signal Processing*, vol. 2013, Article ID 905685, 53 pages, November 2012, doi:10.1155/2013/905685.
117. D. M. Chandler, M. Damon, Md. M. Alam, T. D. Phan, "Seven challenges for image quality research", *IS&T/SPIE Electronic Imaging*, vol. 9014, pp. 901402-901402-14, 2014.
118. International Telecommunication Union. Bt-500-11: Methodology for the Subjective Assessment of the Television Pictures, International Telecommunication Union Recommendation.
119. H. R. Sheikh, M. F. Sabir, A. C. Bovik, “A statistical evaluation of recent full reference image quality assessment algorithms”, *IEEE Transactions on Image Processing*, vol. 15, pp. 3440-3445, May 2006.
120. H. R. Sheikh, Z. Wang, L. Cormack , A.C. Bovik, "LIVE Image Quality Assessment Database Release 2", <http://live.ece.utexas.edu/research/quality>.
121. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, “TID2008- A database for evaluation of full-reference visual quality assessment metrics”, *Advances of Modern Radio electronics*, vol. 10, pp. 30-45, 2009

122. N. Ponomarenko, F. Battisti, K. Egiazarian, J. Astola, V. Lukin, "Metrics performance comparison for colour image database", *Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, vol. 27, pp. 1-6, 2009.
123. N. Ponomarenko, L. Jin, O. Ieremeiev, V. Lukin, K. Egiazarian, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, J. Kuo, "Image database TID2013: Peculiarities, results and perspectives", *Signal Processing: Image Communication*, vol. 30, pp. 57-77, January 2015.
124. N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, L. Jin, J. Astola, B. Vozel, K. Chehdi, M. Carli, F. Battisti, J. Kuo, "Color image database TID2013: Peculiarities and preliminary results", *4th European Workshop on Visual Information Processing*, EUVIP, pp. 106-111, June 2013.
125. S. Winkler, "Analysis of public image and video databases for quality assessment", *IEEE Journal on Selected Topics in Signal Processing*, vol. 6, October 2012.
126. VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment", March 2000, <http://www.vqge.org/>.
127. B. Yegnanarayana, "Artificial Neural Network", *PHI Learning Pvt. Ltd*, New Delhi, INDIA, 2009.
128. M. Hagan, H. Demuth, M. H. Beale, 'Neural Network Design', 4th Indian reprint 2011, *Cengage Learning India Pvt. Ltd.*, New Delhi, India, 1996.
129. R. C. Gonzalez, R. E. Woods, "Digital Image Processing", *Pearson education*, second edition.
130. J. Bagade, K. Singh, Y. Dandawate, "Generalized reference-less image quality assessment framework using texture energy measures and pattern strength features", *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 25, pp. 3947-3959, March 2017.
131. A. Mir, M. Hanmandlu, S. Tandon, "Texture analysis of CT image", *IEEE Engineering in Medicine & Biology Magazine*, vol. 14, pp. 781-786, November 1995.
132. T. Pradeep Kumar, M. Veena Prasad Reddy, P. K. Bora, "Leaf identification using shape and texture features", *Advances in Intelligent Systems & Computing*, vol. 460 AISC, pp. 531-547, December 2017.
133. R. M. Haralick, "Statistical and structural approaches to texture", *Proceedings of the IEEE*, vol. 67, no. 5, pp.786-804, May 1979.
134. R. M. Haralick, K. Shanmugam & I. Dinstein, "Textural features for image classification", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 611-621, November 1973.

135. K. Laws “Textured image segmentation”, *Ph.D. Dissertation, University of Southern California*, 1980.
136. T. Ojala, M. Pietikäinen, “Unsupervised texture segmentation using feature distributions”, *Pattern Recognition*, vol. 32 pp. 477-486, March 1999.
137. A. K. Jain, “Fundamentals of Digital Image Processing”, *Prentice Hall Publishers*, 1989.
138. JPEG official site, Available at <http://www.jpeg.org/>.
139. S. Li, W. Li, “ Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding”, *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 5, pp. 725- 743, August 2000.
140. M. Cagnazzo, G. Poggi, L. Verdoliva, and A. Zinicola, “Region-oriented compression of multispectral images by shape-adaptive wavelet transform and SPIHT,” *Proceedings of IEEE International Conference on Image Processing (ICIP '04)*, vol. 4, pp. 2459–2462, Singapore, 24-27 October 2004.
141. L. Liu, Y. Hua, Q. Zhao, H. Huang, A. Bovik, “Blind image quality assessment by relative gradient statistics and adaboosting neural network”, *Signal Processing Image Communication*, vol. 40, pp. 1-15, January 2016.
142. W. Xue, X. Mou, L. Zhang, A.C. Bovik, X. Feng, “Blind image quality assessment using joint statistics of gradient magnitude and laplacian features”, *IEEE Transaction on Image Processing*, vol. 23, no. 11, pp. 4850–4862, September 2014.
143. S. Grgic, M. Mrak, M. Grgic, “Comparison of JPEG image coder”, *Proceeding Of the 3rd International Symposium on Video Processing & Multimedia Communication, VIPromCom-2001*, pp. 79-85, 2001.
144. M. Black, G. Sapiro, D. Marimont, D. Heeger, “Robust Anisotropic Diffusion”, *IEEE Transactions on Image Processing*, vol. 7, no. 3, March 1998.
145. Q. Hung Do, J. Chen, “A neuro-fuzzy approach in the classification of students’ academic performance”, *Computational Intelligence and Neuroscience*, vol. 2013, Article ID 179097, September 2013.
146. F. Qiu, “Neuro-fuzzy based analysis of hyperspectral imagery”, *Photogrammetric Engineering & Remote Sensing*, vol. 74, pp. 1235-1247, October 2008.
147. J. Benitz, J. Castro, I. Requena, “ Are artificial neural networks black boxes?”, *IEEE Transactions on Neural Networks*, vol. 8, no. 5, pp. 1156-1164, 1997.
148. S. Mallat, “A Wavelet Tour of Signal Processing: The sparse way”, *Academic Press, Elsevier, Burlington*, 2008.

149. V. Gadre, A. Abhyankar “Multiresolution & multirate signal processing: Introduction, principles & applications”, *McGraw Hills Education*, July 2017.
150. D. Wei, Y. Li, “No-Reference Image Quality Assessment Based on SIFT Feature Points”, *International Journal of Simulation - Systems, Science & Technology*, vol. 17, no. 17, pp. 7.1-7.5, January 2016.
151. A. K. Moorthy, A. C. Bovik, “Blind image quality assessment: From natural scene statistics to perceptual quality”, *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3350-3364, December 2011.
152. A. K. Moorthy, A. C. Bovik, “Statistics of natural image distortions”, *Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing*, pp. 962-965, Dallas, TX, USA , 14-19 March 2010.
153. J. Liu, X. Yu, “Research on SAR image matching technology based on SIFT”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII. Part B1. Beijing, 2008.
154. W. Liu, C. Li, Y. Chi, X. Sun, “Image quality assessment based on SIFT and SSIM”, *Advances in Image and Graphics Technologies, IGTA 2014, Communications in Computer and Information Science*, vol. 437, Springer, Berlin, Heidelberg, 2014.
155. A. Djimeli, D. Tchiotsop, R. Tichinda, “Analysis of interest points of curvelet coefficients Contributions of microscopic images and improvement of edges”, *Signal and Image Processing: An International Journal (SIPIJ)*, vol.4, no. 2, May 2013.
156. G. Ciocca, S. Corchs, F. Gasparini, R. Schettini, “How to assess image quality within a workflow chain: an overview”, *International Journal on Digital Libraries*, vol. 15, no. 1, pp. 1-25. 25p, November 2014.
157. D. Lisin, M. Mattar, M. Blaschko, M. Benfield, E. Learned-Miller, “Combining local and global image features for object class recognition”, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 21-23 September 2005.
158. B. Wu, R. Nevatia, “Optimizing discrimination-efficiency tradeoff in integrating heterogeneous local features for object detection”, *IEEE Conference on Computer Vision and Pattern Recognition*, 23-28 June 2008.
159. M. Richardson, “Principal Component Analysis”, <http://people.maths.ox.ac.uk/richardsonm/SignalProcPCA.pdf>, 2009.

Appendix I

List of Features used in Experimentation

The features extracted for this study are statistically represented as follows.

1. Statistical Features

These features are implemented statistically. Also these features are computed in frequency domain.

1. Mean

$$m = \sum_{i=0}^{L-1} r_i p(r_i) \quad (1)$$

2. Standard Deviation

$$\sigma = \sqrt{\sum_{i=0}^{L-1} (r_i - m)^2 p(r_i)} \quad (2)$$

3. Variance

$$\sigma^2(r) = \sum_{i=0}^{L-1} (r_i - m)^2 p(r_i) \quad (3)$$

4. Skewness

$$s = \frac{1}{\sigma^3(r)} \sum_{i=0}^{L-1} (r_i - m)^3 p(r_i) \quad (4)$$

5. Kurtosis

$$k = \frac{1}{\sigma^4(r)} \sum_{i=0}^{L-1} (r_i - m)^4 p(r_i) \quad (5)$$

6. Entropy

$$E = -\sum_{i=0}^{L-1} p(r_i) \log_2 p(r_i) \quad (6)$$

where, r denotes image grey levels and $p(r_i)$, $i=0,1,2,\dots,L-1$, be corresponding histogram, where L is the number of distinct grey levels.

2. Textural Features

These features are derived from Grey Level Co-occurrence Matrices (GLCM)

1. Contrast

$$C = \sum_{i,j} |i - j|^2 p(i, j) \quad (7)$$

2. Correlation

$$Corr = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) p(i, j)}{\sigma_i \sigma_j} \quad (8)$$

3. Homogeneity

$$H = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (9)$$

4. Energy

$$E = \sum_{i,j} p(i, j)^2 \quad (10)$$

5. Inverse Difference Moment

$$IDM = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i - j)^2} p(i, j) \quad (11)$$

6. Smoothness

$$R = 1 - \frac{1}{1 + \sigma^2} \quad (12)$$

3. Edge based feature

This feature is proposed by S. Suresh *et al.* [68].

Horizontal/Vertical Map

$$\hat{E}_h = I * P_h \quad (13)$$

$$P_h = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \quad (14)$$

$$E_h(i, j) = \begin{cases} \hat{E}_h(i, j) & \text{if } \hat{E}_h(i, j) < \tau_e \\ 0 & \text{Otherwise} \end{cases} \quad (15)$$

where $\tau_e=35$.

Appendix II

Image Parameters for Images from LIVE dataset.

LIVE is a standard image dataset prepared and made publically available for the research in image quality assessment. A “readme.txt” file provided with the dataset shares the different parameters for different distortions. Following table is prepared from the same information.

Image	JPEG (Bit rate)	JPEG2000 (Bit rate)	Gblur (Sigma)	WN (Sigma)	FF (Strength)
bikes.bmp	0.30737 (img29.bmp)	1.6219 (img15.bmp)	2.624972 (img40.bmp)	1.996094 (img87.bmp)	16.5 (img31.bmp)
	0.56447 (img81.bmp)	0.6957 (img78.bmp)	1.307261 img26.bmp	0.320313 img29.bmp	18.9 img32.bmp
	0.22872 (img188.bmp)	0.13174 (img115.bmp)	0.96351 img129.bmp	0.257813 img97.bmp	20.3 img33.bmp
	1.0323 (img203.bmp)	1.1466 (img129.bmp)	0.906218 img112.bmp	0.089844 img68.bmp	22.7 img34.bmp
	1.4875 (img226.bmp)	0.33249 (img133.bmp)	0.619759 img60.bmp	0.035156 img95.bmp	25.1 img35.bmp
building2.bmp	0.39385 (img36.bmp)	0.12981 (img30.bmp)	2.052054 (img25.bmp)	0.035156 (img118.bmp)	16.5 (img51.bmp)
	0.64297 (img60.bmp)	0.72182 (img41.bmp)	1.164031 (img80.bmp)	1.5 (img136.bmp)	18.9 (img52.bmp)
	2.0938 (img171.bmp)	1.0458 (img132.bmp)	0.934864 (img33.bmp)	0.421875 (img121.bmp)	20.3 (img53.bmp)
	0.28835 (img205.bmp)	0.32443 (img163.bmp)	0.72002 (img8.bmp)	0.3125 (img36.bmp)	22.7 (img54.bmp)
	1.1092 (img225.bmp)	1.7229 (img216.bmp)	0.533821 (img137.bmp)	0.117188 (img102.bmp)	26.1 (img55.bmp)
buildings.bmp	0.58384 (img43.bmp)	0.40141 (img13.bmp)	2.624972 (img73.bmp)	1.996094 (img4.bmp)	16.5 (img66.bmp)
	1.7772 (img157.bmp)	0.20136 (img29.bmp)	1.536428 (img134.bmp)	0.40625 (img61.bmp)	17.9 (img67.bmp)
	0.26731 (img162.bmp)	0.37358 (img33.bmp)	0.934864 (img62.bmp)	0.289063 (img130.bmp)	20.3 (img68.bmp)
	1.0394 (img163.bmp)	1.6665 (img199.bmp)	0.848926 (img7.bmp)	0.058594 (img46.bmp)	22.7 (img69.bmp)
	0.24723 (img204.bmp)	0.84836 (img222.bmp)	0.562467 (img45.bmp)	0.03125 (img103.bmp)	26.1 (img70.bmp)
caps.bmp	0.85118 (img24.bmp)	0.60354 (img49.bmp)	5.833312 (img70.bmp)	1 (img80.bmp)	15.5 (img86.bmp)
	0.40535 (img63.bmp)	0.40137 (img68.bmp)	3.083306 (img88.bmp)	0.125 (img144.bmp)	18.9 (img87.bmp)
	0.42483 (img105.bmp)	0.19944 (img83.bmp)	1.708303 (img115.bmp)	0.0625 (img110.bmp)	20.3 (img88.bmp)
	0.15312 (img159.bmp)	0.098741 (img122.bmp)	1.164031 (img1.bmp)	0.03125 (img132.bmp)	23.7 (img89.bmp)
	0.1993 (img207.bmp)	0.050378 (img193.bmp)	0.677051 (img28.bmp)	0.015625 (img64.bmp)	25.1 (img90.bmp)
carnivaldolls.bmp	2.5748 (img14.bmp)	0.7153 (img18.bmp)	2.854139 (img52.bmp)	1.996094 (img47.bmp)	16.5 (img101.bmp)
	0.3273 (img80.bmp)	0.12349 (img79.bmp)	1.765595 (img6.bmp)	0.207031 (img135.bmp)	18.9 (img102.bmp)
	0.3273 (img104.bmp)	0.43582 (img108.bmp)	1.278615 (img59.bmp)	0.15625 (img71.bmp)	20.3 (img103.bmp)
	0.20726 (img176.bmp)	0.25592 (img124.bmp)	0.877572 (img27.bmp)	0.046875 (img90.bmp)	22.7 (img104.bmp)
	1.0085 (img216.bmp)	1.5053 (img148.bmp)	0.505175 (img95.bmp)	0.019531 (img14.bmp)	26.1 (img105.bmp)

Image	JPEG (Bit rate)	JPEG2000 (Bit rate)	Gblur (Sigma)	WN (Sigma)	FF (Strength)
cemetery.bmp	0.85407 (img55.bmp)	0.12538 (img46.bmp)	2.166638 (img83.bmp)	1.996094 (img23.bmp)	15.5 (img116.bmp)
	0.30013 (img78.bmp)	0.12274 (img48.bmp)	1.249969 (img135.bmp)	0.25 (img65.bmp)	17.9 (img117.bmp)
	1.5697 (img117.bmp)	0.66567 (img98.bmp)	0.96351 (img109.bmp)	0.179688 (img9.bmp)	20.3 (img118.bmp)
	0.53292 (img128.bmp)	0.98942 (img125.bmp)	0.762988 (img143.bmp)	0.0625 (img63.bmp)	22.7 (img119.bmp)
	0.21773 (img213.bmp)	0.32443 (img178.bmp)	0.619759 (img13.bmp)	0.039063 (img45.bmp)	25.1 (img120.bmp)
churchandcapitol.bmp	0.83865 (img2.bmp)	0.12615 (img5.bmp)	7.66665 (img104.bmp)	1.996094 (img20.bmp)	15.5 (img131.bmp)
	0.29125 (img4.bmp)	0.38308 (img7.bmp)	1.565074 (img66.bmp)	0.234375 (img93.bmp)	18.9 (img132.bmp)
	0.57502 (img21.bmp)	0.104 (img45.bmp)	1.249969 (img91.bmp)	0.140625 (img62.bmp)	21.3 (img133.bmp)
	1.6916 (img132.bmp)	0.69118 (img176.bmp)	0.906218 (img2.bmp)	0.035156 (img18.bmp)	22.7 (img134.bmp)
	0.21903 (img195.bmp)	1.9567 (img198.bmp)	0.734343 (img142.bmp)	0.019531 (img13.bmp)	26.1 (img135.bmp)
coinsinfountain.bmp	0.58801 (img49.bmp)	0.70372 (img21.bmp)	2.510388 (img123.bmp)	1 (img24.bmp)	16.5 (img1.bmp)
	0.46101 (img73.bmp)	0.10943 (img28.bmp)	1.708303 (img19.bmp)	0.25 (img92.bmp)	18.9 (img2.bmp)
	1.8247 (img101.bmp)	3.1539 (img77.bmp)	1.45049 (img22.bmp)	0.171875 (img124.bmp)	21.3 (img3.bmp)
	0.25896 (img129.bmp)	0.18741 (img168.bmp)	0.848926 (img101.bmp)	0.050781 (img31.bmp)	23.7 (img4.bmp)
	0.18796 (img184.bmp)	0.32846 (img182.bmp)	0.562467 (img32.bmp)	0.03125 (img8.bmp)	25.1 (img5.bmp)
dancers.bmp	0.60116 (img39.bmp)	0.25079 (img4.bmp)	2.166638 (img41.bmp)	0.625 (img15.bmp)	16.5 (img16.bmp)
	0.43061 (img41.bmp)	0.11189 (img65.bmp)	1.479136 (img128.bmp)	0.390625 (img50.bmp)	17.9 (img17.bmp)
	2.6988 (img61.bmp)	0.87657 (img165.bmp)	0.96351 (img93.bmp)	0.25 (img51.bmp)	21.3 (img18.bmp)
	1.2678 (img119.bmp)	0.534 (img170.bmp)	0.906218 (img21.bmp)	0.085938 (img27.bmp)	23.7 (img19.bmp)
	0.24753 (img186.bmp)	1.6262 (img205.bmp)	0.562467 (img51.bmp)	0.042969 (img30.bmp)	26.1 (img20.bmp)
flowersonih35.bmp	0.92722 (img27.bmp)	0.20353 (img44.bmp)	2.624972 (img107.bmp)	0.75 (img53.bmp)	16.5 (img36.bmp)
	2.9792 (img74.bmp)	0.12805 (img86.bmp)	1.135385 (img85.bmp)	0.296875 (img49.bmp)	18.9 (img37.bmp)
	0.27056 (img79.bmp)	1.8386 (img111.bmp)	0.891895 (img65.bmp)	0.25 (img122.bmp)	20.3 (img38.bmp)
	1.769 (img198.bmp)	0.48161 (img167.bmp)	0.762988 (img87.bmp)	0.070313 (img72.bmp)	23.7 (img39.bmp)
	0.50574 (img229.bmp)	0.94912 (img192.bmp)	0.447884 (img90.bmp)	0.019531 (img133.bmp)	25.1 (img40.bmp)
house.bmp	0.56641 (img51.bmp)	0.21506 (img47.bmp)	3.999975 (img138.bmp)	0.25 (img83.bmp)	16.5 (img71.bmp)
	0.32699 (img75.bmp)	0.68106 (img80.bmp)	1.393198 (img113.bmp)	0.875 (img38.bmp)	17.9 (img72.bmp)
	1.4029 (img120.bmp)	0.08262 (img144.bmp)	1.049447 (img122.bmp)	0.125 (img109.bmp)	20.3 (img73.bmp)
	0.15798 (img147.bmp)	0.40101 (img187.bmp)	0.992156 (img49.bmp)	0.046875 (img70.bmp)	22.7 (img74.bmp)
	1.8851 (img151.bmp)	1.8156 (img212.bmp)	0.533821 (img94.bmp)	0.019531 (img54.bmp)	26.1 (img75.bmp)

Image	JPEG (Bit rate)	JPEG2000 (Bit rate)	Gblur (Sigma)	WN (Sigma)	FF (Strength)
lighthouse.bmp	0.42931 (img44.bmp)	0.65051 (img51.bmp)	14.999972 (img24.bmp)	1 (img66.bmp)	16.5 (img91.bmp)
	2.6187 (img57.bmp)	0.24826 (img91.bmp)	1.479136 (img97.bmp)	0.171875 (img96.bmp)	17.9 (img92.bmp)
	0.3919 (img86.bmp)	0.24231 (img106.bmp)	1.135385 (img15.bmp)	0.0625 (img40.bmp)	20.3 (img93.bmp)
	0.19213 (img161.bmp)	0.36474 (img149.bmp)	0.82028 (img102.bmp)	0.039063 (img114.bmp)	23.7 (img94.bmp)
	1.2943 (img231.bmp)	1.5456 (img174.bmp)	0.447884 (img4.bmp)	0.019531 (img43.bmp)	26.1 (img95.bmp)
lighthouse2.bmp	2.2014 (img64.bmp)	0.23974 (img74.bmp)	2.739555 (img118.bmp)	1.996094 (img137.bmp)	16.5 (img106.bmp)
	0.40802 (img69.bmp)	2.9056 (img103.bmp)	1.59372 (img53.bmp)	0.164063 (img32.bmp)	18.9 (img107.bmp)
	2.607 (img97.bmp)	2.3683 (img114.bmp)	1.278615 (img55.bmp)	0.121094 (img108.bmp)	21.3 (img108.bmp)
	1.2186 (img141.bmp)	0.52595 (img164.bmp)	1.078093 (img75.bmp)	0.035156 (img6.bmp)	23.7 (img109.bmp)
	0.22353 (img181.bmp)	1.4569 (img211.bmp)	0.591113 (img133.bmp)	0.019531 (img125.bmp)	26.1 (img110.bmp)
manfishing.bmp	0.56739 (img33.bmp)	2.2558 (img26.bmp)	2.166638 (img78.bmp)	1.996094 (img41.bmp)	16.5 (img121.bmp)
	0.33917 (img93.bmp)	0.60047 (img63.bmp)	1.822887 (img16.bmp)	0.242188 (img73.bmp)	18.9 (img122.bmp)
	1.6232 (img135.bmp)	0.58961 (img94.bmp)	1.106739 (img47.bmp)	0.171875 (img143.bmp)	20.3 (img123.bmp)
	0.21134 (img148.bmp)	0.36474 (img143.bmp)	0.791634 (img14.bmp)	0.042969 (img37.bmp)	23.7 (img124.bmp)
	0.95764 (img174.bmp)	0.09471 (img200.bmp)	0.505175 (img67.bmp)	0.015625 (img59.bmp)	26.1 (img125.bmp)
monarch.bmp	0.32107 (img18.bmp)	0.20004 (img56.bmp)	11.333325 (img11.bmp)	1 (img34.bmp)	15.5 (img136.bmp)
	0.77547 (img94.bmp)	0.4277 (img61.bmp)	2.854139 (img37.bmp)	0.1875 (img140.bmp)	17.9 (img137.bmp)
	2.6482 (img96.bmp)	2.7899 (img105.bmp)	1.851533 (img3.bmp)	0.117188 (img17.bmp)	21.3 (img138.bmp)
	0.60179 (img178.bmp)	0.042317 (img160.bmp)	1.708303 (img103.bmp)	0.035156 (img1.bmp)	23.7 (img139.bmp)
	0.1814 (img209.bmp)	0.10277 (img208.bmp)	0.906218 (img96.bmp)	0.027344 (img82.bmp)	26.1 (img140.bmp)
ocean.bmp	0.881 (img8.bmp)	0.58689 (img31.bmp)	5.833312 (img117.bmp)	1.996094 (img42.bmp)	16.5 (img6.bmp)
	3.3364 (img30.bmp)	0.40169 (img37.bmp)	1.479136 (img57.bmp)	0.179688 (img120.bmp)	18.9 (img7.bmp)
	1.4233 (img143.bmp)	2.5311 (img69.bmp)	1.078093 (img119.bmp)	0.109375 (img21.bmp)	20.3 (img8.bmp)
	0.15037 (img164.bmp)	0.19144 (img120.bmp)	0.791634 (img106.bmp)	0.011719 (img99.bmp)	23.7 (img9.bmp)
	0.48142 (img194.bmp)	0.062469 (img219.bmp)	0.562467 (img92.bmp)	0 (img147.bmp)	25.1 (img10.bmp)
paintedhouse.bmp	0.29938 (img23.bmp)	0.41508 (img14.bmp)	7.66665 (img116.bmp)	1.996094 (img94.bmp)	16.5 (img21.bmp)
	0.62819 (img56.bmp)	0.1909 (img40.bmp)	1.421844 (img81.bmp)	0.25 (img39.bmp)	18.9 (img22.bmp)
	0.42863 (img152.bmp)	2.437 (img82.bmp)	1.249969 (img110.bmp)	0.125 (img26.bmp)	21.3 (img23.bmp)
	0.18284 (img215.bmp)	0.24383 (img196.bmp)	0.949187 (img39.bmp)	0.058594 (img111.bmp)	23.7 (img24.bmp)
	1.4535 (img228.bmp)	0.78388 (img210.bmp)	0.562467 (img23.bmp)	0.03125 (img142.bmp)	26.1 (img25.bmp)

Image	JPEG (Bit rate)	JPEG2000 (Bit rate)	Gblur (Sigma)	WN (Sigma)	FF (Strength)
parrots.bmp	0.33221 (img32.bmp)	0.38193 (img85.bmp)	7.66665 (img69.bmp)	1 (img104.bmp)	16.5 (img41.bmp)
	0.92118 (img72.bmp)	0.22432 (img96.bmp)	3.999975 (img105.bmp)	0.128906 (img69.bmp)	17.9 (img42.bmp)
	0.31453 (img103.bmp)	0.68851 (img101.bmp)	2.166638 (img31.bmp)	0.0625 (img67.bmp)	21.3 (img43.bmp)
	0.21 (img149.bmp)	0.054408 (img158.bmp)	1.249969 (img56.bmp)	0.03125 (img22.bmp)	23.7 (img44.bmp)
	0.15717 (img196.bmp)	0.13904 (img197.bmp)	0.791634 (img12.bmp)	0.011719 (img75.bmp)	25.1 (img45.bmp)
plane.bmp	2.1241 (img17.bmp)	0.11053 (img53.bmp)	4.916644 (img5.bmp)	1 (img105.bmp)	15.5 (img56.bmp)
	0.42568 (img25.bmp)	2.3766 (img58.bmp)	3.312473 (img139.bmp)	0.1875 (img129.bmp)	18.9 (img57.bmp)
	0.9083 (img40.bmp)	0.60958 (img113.bmp)	1.708303 (img30.bmp)	0.109375 (img139.bmp)	20.3 (img58.bmp)
	0.27649 (img165.bmp)	0.40504 (img173.bmp)	1.078093 (img86.bmp)	0.042969 (img10.bmp)	23.7 (img59.bmp)
	0.1628 (img201.bmp)	0.20756 (img203.bmp)	0.562467 (img63.bmp)	0.019531 (img78.bmp)	26.1 (img60.bmp)
rapids.bmp	0.43058 (img9.bmp)	0.72312 (img3.bmp)	7.66665 (img17.bmp)	1 (img77.bmp)	15.5 (img76.bmp)
	0.29138 (img35.bmp)	0.1227 (img32.bmp)	1.421844 (img141.bmp)	0.25 (img116.bmp)	17.9 (img77.bmp)
	0.85997 (img110.bmp)	0.41906 (img81.bmp)	1.020802 (img34.bmp)	0.125 (img74.bmp)	20.3 (img78.bmp)
	1.4648 (img146.bmp)	1.5456 (img126.bmp)	0.791634 (img29.bmp)	0.046875 (img107.bmp)	23.7 (img79.bmp)
	0.17405 (img199.bmp)	0.21562 (img146.bmp)	0.562467 (img9.bmp)	0.019531 (img101.bmp)	25.1 (img80.bmp)
sailing1.bmp	0.9069 (img102.bmp)	2.5288 (img19.bmp)	2.166638 (img125.bmp)	1.996094 (img52.bmp)	15.5 (img96.bmp)
	3.2723 (img106.bmp)	0.11565 (img43.bmp)	1.249969 (img130.bmp)	0.1875 (img141.bmp)	17.9 (img97.bmp)
	0.54289 (img121.bmp)	0.12526 (img88.bmp)	1.078093 (img74.bmp)	0.128906 (img79.bmp)	21.3 (img98.bmp)
	0.2724 (img175.bmp)	0.48564 (img136.bmp)	0.734343 (img35.bmp)	0.046875 (img56.bmp)	22.7 (img99.bmp)
	0.17035 (img210.bmp)	1.0781 (img214.bmp)	0.591113 (img46.bmp)	0.023438 (img76.bmp)	25.1 (img100.bmp)
sailing2.bmp	0.55609 (img67.bmp)	0.20057 (img55.bmp)	11.333325 (img10.bmp)	1 (img113.bmp)	16.5 (img111.bmp)
	0.39285 (img111.bmp)	2.511 (img59.bmp)	2.166638 (img48.bmp)	0.132813 (img128.bmp)	17.9 (img112.bmp)
	0.16377 (img124.bmp)	1.849 (img62.bmp)	1.421844 (img72.bmp)	0.0625 (img12.bmp)	20.3 (img113.bmp)
	1.3781 (img144.bmp)	1.1063 (img141.bmp)	1.078093 (img43.bmp)	0.027344 (img117.bmp)	23.7 (img114.bmp)
	0.17097 (img177.bmp)	0.52997 (img185.bmp)	0.734343 (img84.bmp)	0.015625 (img81.bmp)	26.1 (img115.bmp)
sailing3.bmp	0.91991 (img38.bmp)	0.38246 (img12.bmp)	3.541641 (img144.bmp)	1.996094 (img58.bmp)	15.5 (img126.bmp)
	0.59313 (img83.bmp)	0.10603 (img38.bmp)	2.510388 (img124.bmp)	0.164063 (img123.bmp)	18.9 (img127.bmp)
	0.33148 (img108.bmp)	0.37665 (img109.bmp)	1.45049 (img114.bmp)	0.109375 (img89.bmp)	21.3 (img128.bmp)
	1.8247 (img206.bmp)	0.22771 (img130.bmp)	0.96351 (img79.bmp)	0.03125 (img84.bmp)	23.7 (img129.bmp)
	0.16271 (img223.bmp)	1.3804 (img183.bmp)	0.562467 (img111.bmp)	0.019531 (img5.bmp)	25.1 (img130.bmp)

Image	JPEG (Bit rate)	JPEG2000 (Bit rate)	Gblur (Sigma)	WN (Sigma)	FF (Strength)
sailing4.bmp	0.55225 (img15.bmp)	0.59673 (img39.bmp)	3.999975 (img121.bmp)	0.28125 (img134.bmp)	16.5 (img141.bmp)
	0.30951 (img37.bmp)	0.26196 (img52.bmp)	1.507782 (img127.bmp)	0.15625 (img126.bmp)	17.9 (img142.bmp)
	0.42826 (img47.bmp)	1.8547 (img90.bmp)	1.049447 (img76.bmp)	0.125 (img86.bmp)	21.3 (img143.bmp)
	0.16614 (img197.bmp)	0.030227 (img172.bmp)	0.791634 (img89.bmp)	0.046875 (img100.bmp)	22.7 (img144.bmp)
	1.2926 (img214.bmp)	0.11486 (img221.bmp)	0.677051 (img136.bmp)	0.015625 (img127.bmp)	26.1 (img145.bmp)
statue.bmp	0.29625 (img11.bmp)	0.3777 (img92.bmp)	3.999975 (img120.bmp)	1 (img145.bmp)	16.5 (img11.bmp)
	2.7772 (img13.bmp)	2.4112 (img102.bmp)	2.166638 (img54.bmp)	0.203125 (img91.bmp)	17.9 (img12.bmp)
	2.1953 (img70.bmp)	0.22253 (img116.bmp)	1.822887 (img131.bmp)	0.109375 (img25.bmp)	20.3 (img13.bmp)
	1.103 (img130.bmp)	0.74761 (img162.bmp)	1.307261 (img77.bmp)	0.046875 (img55.bmp)	23.7 (img14.bmp)
	0.16597 (img221.bmp)	0.070529 (img169.bmp)	0.848926 (img98.bmp)	0.015625 (img85.bmp)	25.1 (img15.bmp)
stream.bmp	0.57467 (img16.bmp)	0.71533 (img8.bmp)	3.083306 (img58.bmp)	0.5 (img131.bmp)	16.5 (img26.bmp)
	0.41781 (img85.bmp)	0.40832 (img17.bmp)	1.020802 (img50.bmp)	0.3125 (img106.bmp)	18.9 (img27.bmp)
	0.29203 (img100.bmp)	0.19204 (img71.bmp)	0.834603 (img126.bmp)	0.1875 (img2.bmp)	20.3 (img28.bmp)
	1.6833 (img137.bmp)	0.050378 (img175.bmp)	0.762988 (img18.bmp)	0.0625 (img88.bmp)	23.7 (img29.bmp)
	1.0008 (img185.bmp)	1.4851 (img217.bmp)	0.419238 (img71.bmp)	0.03125 (img138.bmp)	26.1 (img30.bmp)
studentsculpture.bmp	0.82276 (img45.bmp)	1.6466 (img2.bmp)	1.651011 (img100.bmp)	1.996094 (img57.bmp)	16.5 (img46.bmp)
	2.338 (img53.bmp)	2.7031 (img34.bmp)	1.364552 (img38.bmp)	0.4375 (img44.bmp)	17.9 (img47.bmp)
	0.88289 (img59.bmp)	0.20383 (img60.bmp)	1.020802 (img20.bmp)	0.234375 (img112.bmp)	20.3 (img48.bmp)
	0.22008 (img127.bmp)	0.46549 (img142.bmp)	0.677051 (img44.bmp)	0.078125 (img16.bmp)	23.7 (img49.bmp)
	0.44665 (img200.bmp)	1.0136 (img171.bmp)	0.533821 (img108.bmp)	0.023438 (img19.bmp)	26.1 (img50.bmp)
woman.bmp	2.7803 (img58.bmp)	0.40027 (img10.bmp)	2.052054 (img68.bmp)	1.996094 (img33.bmp)	15.5 (img61.bmp)
	0.38194 (img65.bmp)	0.59812 (img23.bmp)	1.651011 (img99.bmp)	0.21875 (img7.bmp)	17.9 (img62.bmp)
	3.266 (img116.bmp)	0.3754 (img104.bmp)	1.020802 (img140.bmp)	0.15625 (img115.bmp)	21.3 (img63.bmp)
	1.1368 (img150.bmp)	1.5375 (img150.bmp)	0.934864 (img64.bmp)	0.046875 (img35.bmp)	22.7 (img64.bmp)
	0.23546 (img167.bmp)	0.13904 (img152.bmp)	0.648405 (img145.bmp)	0.015625 (img98.bmp)	26.1 (img65.bmp)
womanhat.bmp	0.3263 (img1.bmp)	0.21173 (img42.bmp)	3.541641 (img132.bmp)	1 (img60.bmp)	16.5 (img81.bmp)
	0.45313 (img89.bmp)	1.708 (img75.bmp)	1.93747 (img36.bmp)	0.136719 (img119.bmp)	18.9 (img82.bmp)
	0.60102 (img107.bmp)	0.2356 (img99.bmp)	1.479136 (img42.bmp)	0.09375 (img28.bmp)	21.3 (img83.bmp)
	0.15771 (img138.bmp)	0.96524 (img195.bmp)	1.020802 (img61.bmp)	0.039063 (img3.bmp)	23.7 (img84.bmp)
	1.4703 (img168.bmp)	0.1068 (img226.bmp)	0.677051 (img82.bmp)	0.019531 (img11.bmp)	25.1 (img85.bmp)

Source: Laboratory of Image and Video Engineering (LIVE) Image Dataset