IMAGE QUALITY MEASUREMENT THROUGH STRUCTURAL SIMILARITY BASED ON HIGH ORDER MOMENTS

A DISSERTATION

Submitted in partial fulfillment of the requirements for the award of degree of

MASTER OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

By

SWARNJEET KOUR

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY

ROORKEE – 247 667 (INDIA)

MAY – 2016

DECLARATION

I declare that the work presented in this dissertation with title, "**Image quality measurement through structural similarity based on high order moments**", towards the fulfillment of the requirements for award of the degree of **Master of Technology** in **Computer Science & Engineering**, submitted to the **Department of Computer Science and Engineering**, **Indian Institute of Technology-Roorkee**, India, is an authentic record of my own work carried out during the period from **June 2015 to May 2016** under the guidance of **Dr. Debashis Sen,** Assistant Professor, Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology-Kharagpur and **Dr. Vaskar Raychoudhury**, Assistant Professor, Department of Computer Science and Engineering, Indian Institute of Technology, Roorkee.

The matter presented in this dissertation has not been submitted by me for the award of any other degree of this or any other institute.

Date:

Place: Roorkee **(Swarnjeet Kour)**

CERTIFICATE

This is to certify that the statement made by the candidate in the declaration is correct to the best of my knowledge and belief.

Date:

Place: Roorkee

Dr. Vaskar Raychoudhury

Assistant Professor

Department of Computer Science and Engineering

Indian Institute of Technology, Roorkee

ABSTRACT

Image quality assessment means estimating the quality of an image. Image quality is a characteristic of an image that measures the perceived image degradation. The quality of image gets affected due to the noise or distortion occurred during the acquisition, transmission, storage and compression. Broadly, quality can be measured in two ways - subjective and objective. In subjective methods humans are asked to rate the video on different scales according to the perceived quality. Objective methods eliminate human involvement by determining the quality of an input image automatically using some algorithm or mathematical model. With the advancement of digital technology, assessing quality automatically becomes more important.

We propose a simple yet efficient objective quality assessment method, structural similarity based on high order moments. SSIM is based on the assumption that human eye is capable of extracting structural information by viewing the image and this structural information is a good measure of quality. Loss of structural information is considered as loss of quality. We attempt to extend the SSIM by incorporating shape parameters of distributions. Quality of image is loss if shape of objects is not preserved. High order central and joint moments are used as shape descriptors in our new approach. We show that a high order moment adds useful extra information to SSIM, which is relevant in quantification of local structures. We also show that this additional information improves the correspondence of SSIM with human perception. Results are taken on various types of distorted images of a standard dataset and new SSIM is validated against SSIM index, mean square error and subjective ratings.

ACKNOWLEDGEMENTS

I would never have been able to complete my dissertation without the guidance of my supervisor, help from friends, and support from my family and loved ones.

I would like to express my deepest gratitude to my supervisor, **Dr. Vaskar Raychoudhury**, for his excellent guidance, meaningful insights and moral support.

I am grateful to the **Dr. Debashis Sen** for his guidance and moral support. He has been supportive since the day I began working on this dissertation and gave me the freedom I needed to explore this area of research on my own, while pointing me in the right direction in the times of need. His comprehensive knowledge in the area of image processing and hardworking nature has been a constant source of inspiration.

I am also grateful to the **Dept. of Computer Science, IIT-Roorkee** for providing valuable resources to aid my research.

I would like to thank **Sartaj Kanwar, Nimita Mangal**, **Harpreet Kaur, Sweta Arya** and other friends who supported me, were always willing to help and give me their best suggestions.

Finally, hearty thanks to **my parents and siblings**, who encouraged me in good times, and motivated me in the bad times, without which this dissertation would not have been possible.

Dedication

To my parents, for giving me the best education they could

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Demand of image-based applications is increasing day–by-day, thus the importance of simple and reliable assessment of image quality is increased. These images pass through various pre-processing stages before they reach human observers. Different type of distortion occurs during different pre-processing stages, e.g., capturing, compressing, and transmitting images from one device to another. For example, during image compression blurring effects arises which leads to degradation in quality. Moreover, in the transmission stage, some data is loss due to limited bandwidth which also results in quality loss of the received image. In order to enhance and maintain the quality of images, it is essential to estimate the quality at each stage.

The task of image quality assessment (IQA) methods is to predict the quality of images and videos in agreement with human perception. A lot of work has been done on IQA over the past years. To estimate quality, we have to understand how human being perceives images or videos. Quality also varies according to the interest, experience and expectation of the viewer. Broadly quality assessment methods are divided in two categories, subjective and objective assessment methods.

1.1. Subjective quality assessment methods

In this approach, quality assessment is performed by human beings. Video clips are shown to evaluators and they are asked to rate the video on different scales according to perceived quality. The advantage of this method is that, it is reliable but it is expensive, consumes lot of time and

not applicable to real time applications. The different Subjective quality assessment methods [1] used are:

1.1.1 Double Stimulus Continuous Quality Scale

The distorted and original video sequences are shown to the evaluators in random order and order is not known to the evaluators. Rating of video quality is done on scale of 0–100 (with 100 denoting excellent rating and 0 as bad rating).

1.1.2 Double Stimulus Impairment Scale

In this procedure, the evaluators are aware of the reference and processed video/image sequence, and rate the sequences from very annoying to good.

1.1.3 Single Stimulus Continuous Quality Evaluation

As its name suggests, only processed video/image sequence is shown to the evaluator. Evaluator rates each video sequence according to the perceived quality, from bad to excellent. Mean is calculated over all sequences to rate the overall quality of video.

1.1.4 Absolute Category Rating

Only processed video is shown to evaluator for rating. One rating is provided by evaluator for the overall quality of video using scale ranging from zero to five.

1.1.5 Difference mean opinion score

DMOS give the difference between original and distorted sequence's mean opinion score (MOS). Less the value of DMOS score better will be quality of distorted image.

Advantage

- Subjective methods are most accurate and reliable quality assessment measures.
- They serve as a benchmark for objective quality assessment methods.

Disadvantage

- They are expensive and time consuming because they need a large number of observers.
- Real time quality assessment is not possible.

- Results of subjective rating depend upon the emotional state of the observer and physical conditions such as viewing distance, lighting condition, etc [17].

1.2 Objective quality assessment methods

In Objective methods, quantitative measures and mathematical models are developed that can automatically measure the image quality according to human perception. Objective methods are classified into the three categories, depending upon availability of a reference image. The first classification is full-reference, second is no-reference and third is reduced reference. In full reference approach, original video is available for comparing with distorted video. No-reference quality assessment method is used when reference image is not available. In reduced reference method, some attributes are extracted from original image and are used for comparison with distorted image [2].

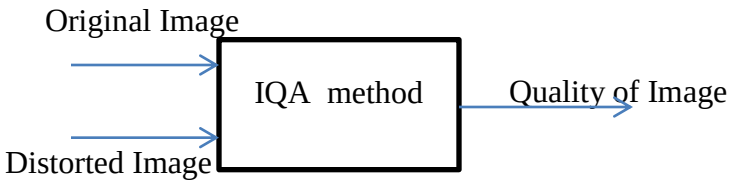


Fig.1.1. Flow diagram of Full Referenced IQA.

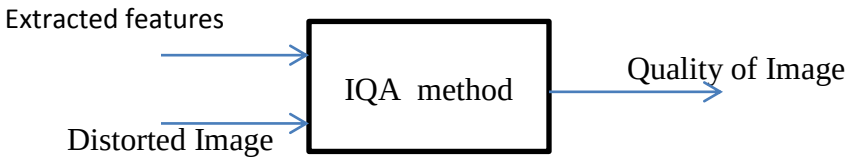


Fig.1.2. Flow diagram of Reduced Referenced IQA.

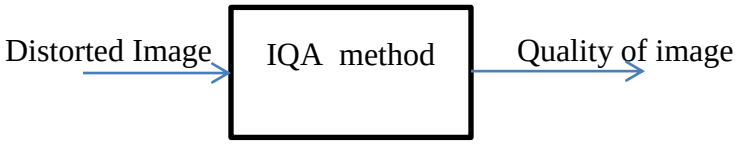


Fig.1.3. Flow diagram of No Referenced IQA.

1.2.1 Type of objective assessment metrics

Various types of metrics are formulated for quality determination. Some of them are described below:

1.2.1.1 Data metrics:

Data metrics are the metrics which do not consider the content of the signal during quality determination. Computation is done by comparing original and reference image, pixel-by-pixel. Mean square error (MSE) and Peak signal to noise ratio (PSNR) are two commonly used data metrics.

1.2.1.2 Picture metrics

These metric are designed using human visual system and try to include features like color perception, contrast sensitivity. Picture metrics specifically report the impacts of distortion on perceived quality. SSIM (Structural Similarity Index Measure) is one of the popular picture metric used for quality assessment.

1.2.1.3 Packet and bit stream- based metrics

It measures the effect of network losses on the image quality. Parameters from the transport and bit stream are analyzed to gain information like packet loss probability, delay, jitter etc. This information is used while modeling metrics that predicts the quality of image/video. V-Factor metric belong to this class.

1.3 Applications of quality assessment

- One of the direct applications of this work, image quality assessment (IQA) is to use it as a standard measure for many of the image processing algorithms. As there are many algorithms related to denoising and restoration is available that removes noise and blurriness from the image and gives a clear image to us. IQA is used to choose one of the best perceptual quality images from the restored images.
- Quality assessment is a technique used for analyzing the trial design and used to prevent this design from errors and biases. Quality assessment is required for the performance measurement of capturing devices (camera), processing or computing algorithms, and various compression algorithms and also can be used to measure the quality of service of networks. These tools find

to be very useful if it embedded with recommender systems and search engines and in result these systems will provide better quality images to the users.

- The objective of image quality assessment is to build a measurable metric which is having an intelligence of computing the image quality properly. Network visual communication is another interesting application where IQA measures are used for monitoring the quality-of-service (QoS) in network. If this better IQA measure is available then one can use this technique not only for evaluating the performance of systems and algorithms, but also used to optimize these systems and algorithms to produce more promising results to users.
- Many of the medical technologies are now using image as a data and now it becomes a challenging task in hospital to how to efficiently store these images in database, and how to transmit and retrieve this medical image information from the large database. Lossy image compression is a technique used for storing the image as this technique reduces the data rate. The risk with this technique is losing some continued diagnostic information that is present in medical images. So, some specific IQA technique is needed that maximizes the level of compression but not affect any diagnostic value of image.
- Images of fingerprints, handwritings, faces, palm prints, and hand shapes are used in many of the biometric methods. These biometric systems have to work correctly under the circumstances where images may contain some noise or distortion. It is necessary to know how much level of quality degradation in image is needed under certain application scenarios. IQA measure is a technique different from the traditional computing techniques used till now, and provides better result in predicting that how much image quality is degrade. Once it is found that how much quality is degraded than some preprocessing algorithm is applied on the image which enhances the quality of image and then some pattern recognition algorithm is applied.

1.4 Performance Metrics

Predicated quality score (X) is compared with subjective score (Y) by using following correlation metrics [11].

1.4.1 Pearson Linear correlation coefficient (PLCC or CC)

It is a measure of how much two distributions (predicated quality and subjective score) are varying linearly. Range of its value is from -1 to +1 inclusive, where one is for positive

correlated, zero is for no correlation, and minus one is for negative correlated. The CC criteria help to define for prediction accuracy.

For the n^{th} image in database of size N, CC is given as:

$$CC = \frac{\sum_{n=1}^{N}(p_{n-}\bar{p})*(q_{n-}\bar{q})}{\sqrt{\sum_{n=1}^{N}(p_{n-}\bar{p})^2 * \sum_{n=1}^{N}(q_{n-}\bar{q})^2}} \quad (1)$$

Where $p_n$is subjective score and q_n is objective score of n^{th} image in database.

1.4.2 Spearman's rank correlation coefficient (SROCC)

It measures strength of association among two distributions X and Y. Direction of association of variables X and Y is indicated by the sign of Spearman's correlation coefficient. If Y tends to increase with increase in X, sign is positive and if Y tends to decrease with increase in X, sign is negative. Zero value of SROCC indicates both variables are not correlated to each other.

SROCC is defined as:

$$SROCC = 1 - \frac{6\sum_{n=1}^{N} {d_n}^2}{N(N-1)^2} \quad (2)$$

Where d_n is difference of objective and subjective score of the nth image in the database and N is total number of images.

1.4.3 Kendall rank correlation coefficient (KROCC)

It is a measure of dependence of two variables on each other. KROCC provides prediction monotonicity [12]. It is given by following mathematical equation.

$$KROCC = \frac{\text{no. of concordant pair} - \text{no.of discordant pair}}{{}^1/_2\text{N(N}-1)} \quad (3)$$

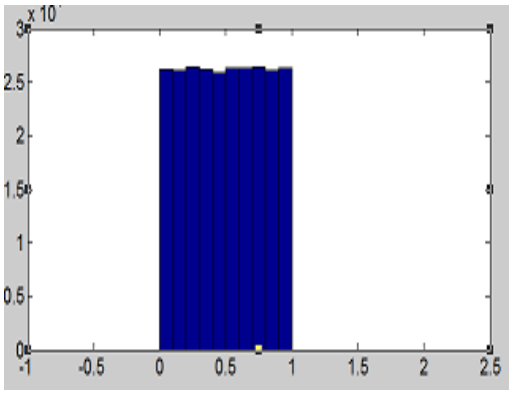
where N is total number of images in dataset. For any pair (x_k, y_k) and (x_l, y_l) under consideration, is *concordant* pair if the both $x_k > x_k$ and $y_l > y_l$ or both $x_k < x_l$ and $y_k < y_l$. They are said to be *discordant*, if $x_k < x_l$ and $y_k > y_l$. or if $x_k > x_l$ and $y_k < y_l$.

1.5 Motivation

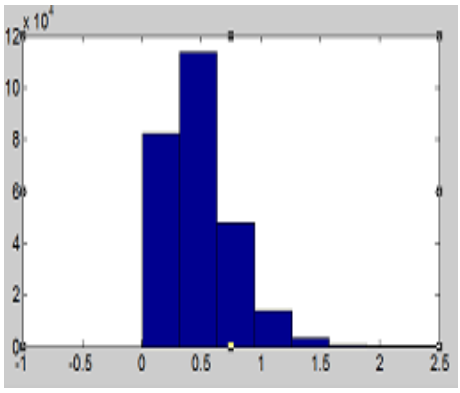
The aim is to find an easy and more direct way to compare the structures of the original and the distorted images. Digital image is nothing but it is representation of how the values of pixels are distributed in 2-D or 3-D space. Local and global distribution of pixel values is one way of representing local and global image structure, respectively. This representation is particularly useful as attributes of distributions like mean, variance and correlation represents what human visual system perceives quite well. However, the moments and their orders which will most effectively relate to human perception are unknown.

Zhou and Wang [5], proposed Structural Similarity model (SSIM) which considers the loss of structural information as degradation in quality. The approach divides structural similarity into three parts, namely, similarity in luminance, contrast and structural similarity. In SSIM metric, only mean, standard deviation and correlation is used to represent image structure. Distributions, local or global, of two images may have same mean, variance and correlation even if their shapes are different.

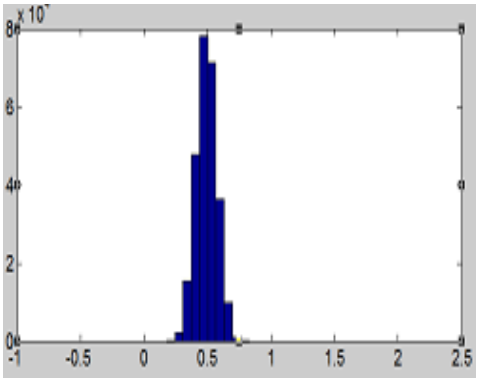
For elaboration, consider the distributions of different shapes given in Fig.1.4. The distributions denoted as A, B, C and D are generated by using different random variable generator functions with mean as 0.5 and variance as 0.07**.** A is a truncated uniform random distribution, B is a gamma distribution, C is a beta distribution and D is a Gaussian distribution.



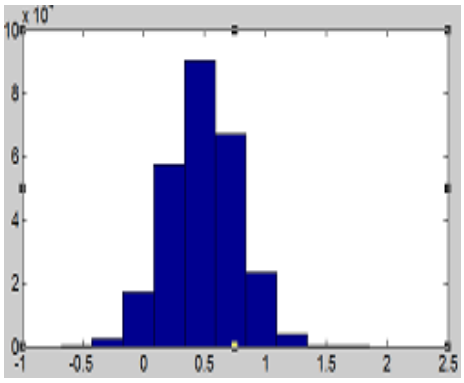
A



B



C



D

Fig.1.4. Shows histograms of A, B, C, D distributions.

TABLE I. COMPARISON OF SSIM VALUE OF DIFFERENT DISTRIBUTIONS

Comparison between Two distributions	**SSIM Value**
SSIM (A,B)	0.9965
SSIM (A,C)	0.9981
SSIM(B,C)	0.9965
SSIM(A,D)	0.9981
SSIM(B,D)	0.9965
SSIM(C,D)	0.9987

As seen from Table I, although the distributions are of fairly different shapes, SSIM takes very high similarity values. It serve as a motivation for our new approach that there is need to consider shape parameter along with the mean, variance and covariance. Shape parameters are introduced by computing and comparing high order moments of both signals.

1.6 Problem Statement

Given: original image and a processed image (distorted image)

Constraints:

- Time bounded - quality to be evaluated in real time.
- High correlation with human perception or judgments.

Objective:

Evaluate the quality or determine the level of distortion present in distorted image by comparing it with the available original image**.**

1.7 Contributions

- We have proposed quality assessment method that can be embedded with search engines and image recommender systems to recommend high quality images to the users from a database of large number of images.
- Proposed measure goes with human perception better than existing quality assessment measures.
- We have compared proposed method with other existing quality assessment methods and results show that our method outperforms them in terms of consistency with subjective ratings.

1.8 Dissertation Overview

The dissertation comprises of five chapters:

- ☐ Chapter 1 introduces image quality assessment, its applications and the quality metrics we have used for our performance analysis. We briefly formulate the problem and summarize the contributions made in this dissertation.
- ☐ Chapter 2 reviews the work related to objective image quality assessment methods and their limitations.
- ☐ Chapter 3 introduces proposed method, structural similarity index based on high order moments.

- Chapter 4 compares the proposed method with other existing methods. Results are taken on various types of distorted images of standard IQA databases and it is shown that the proposed method outperforms all of them. Performance of proposed method is evaluated by using quality metrics
- Chapter 5 summarizes the dissertation and provides some open problems in this area for future extensions of this work.

CHAPTER 2

RELATED WORK

Demand of image based applications is increasing day by day, so there is great need of quality prediction approach that can estimate the quality of image or video automatically in real time. Various quality assessment metrics have been proposed with different accuracy and complexity. The reliable assessment plays a major role in meeting the QoS standards and improving the end user quality. Thus, in the recent years, a lot of work is done on the development of quality assessment models that can closely match the human perception.

2.1 Mean Square Error and Peak Signal-to-Noise Ratio

MSE is a most widely used full reference quality metric, calculated by taking average of the squared intensity differences of processed and original image pixels.

$$MSE = \frac{1}{N}\sum_{i=0}^{N}(y_i - x_i)^2 \quad (4)$$

Where x_i is i^{th} pixel of orginal image, y_i is i^{th} pixel of distorted signal and N is number of pixels.

2.1.1 Advantage

- Easy to compute and understand.
- Have clear physical significance.

2.1.2 Limitations

MSE has many drawbacks such as:

1. Distortion-agnostic

MSE and PSNR both are distortion-agnostic [4]. They do not consider the type of distortion present in image. Some distortions are more or less appealing than other depending upon the type of distortion. Both images shown in Fig.2.1, have same PSNR value but the quality of both the images are different because right image (b) contains high frequency noise and human eye is not capable to notice the high-frequency noise whereas left image (c) has low frequency noise which is localized and visible to human eye.

2. Content-agnostic

Human perception also depends upon the location of distortion. Left image contained noise in bottom region where we have lot of content so distortion cannot stand out where as in image (c) distortion is present in a region having less content activity which makes it clearly visible.

(a) (b) (c)

Fig.2.1. Image (a) is reference image and (b) and (c) has different distortion but has same PSNR [4].

PSNR value for both images (b) and (c) is same but they have different quality. So PSNR and MSE are not consistent with human eye perception [3].

2.2 IQA model based on error sensitivity

It is a Full–reference model, signal whose quality has to be determined is thought as a sum of error signals and undistorted original signals. It is assumed that the quality loss of an image is directly related to the perceptibility of error signals. MSE and PSNR is a simplest implementation of this. Two different quality images may have same MSE. Some errors are more visible than other, so there is need to weight different type of errors according to their visibility [16]. Mannos and Sakrison [5] proposed IQA model using human visual system to weight different errors according to the visibility.

2.2.1 Framework

Generic IQA or VQA model using HVS framework is shown below:

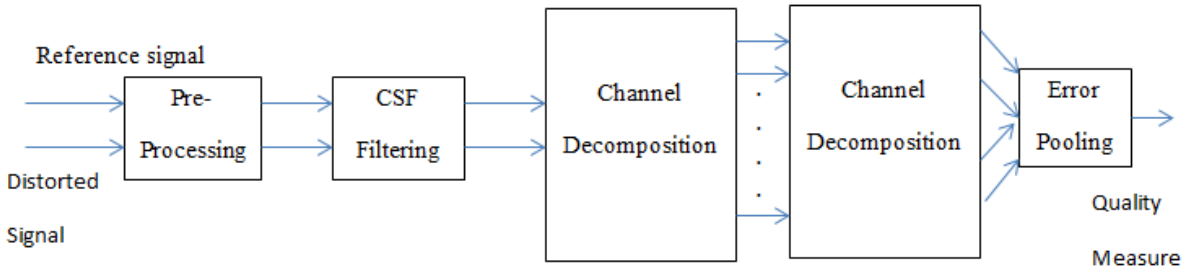


Fig.2.2. System diagram of error sensitivity based IQA system [5]

2.2.2 Stages of error sensitivity VQA model:

Stages of VQA model is described below:

2.2.2.1 Pre-processing:

In this stage, various operations are performed to eliminate distortion:

1. The processed and original signals are aligned properly.

2. The input signals are transformed into a color space visible to human eye [13].

3. Digital pixel value is converted into luminance values of pixels.

2.2.2.2 CSF Filtering:

The contrast sensitivity function (CSF) describes the sensitivity of the HVS to different spatial and temporal frequencies that are present in the visual stimulus [14]. Input signals are weighted according to this function.

2.2.2.3 Channel Decomposition:

The input signals are divided into sub-bands according to the spatial, temporal frequency and orientation. Some methods also implement channel decompositions based upon the neural responses. Discrete cosine or wavelet transforms are used for channel decomposition.

2.2.2.4 Error Normalization

The error in each sub-band is computed and normalized according to its visibility.

2.2.2.5 Error Pooling

In this stage, error signals from different channels are combined to give single value. Usually minkowski norm is used for pooling.

$$E(e_{i,j}) = (\sum_i \sum_j |e_{i,j}|^{\alpha})^{1/\alpha} \quad (5)$$

Where $e_{i,j}$ is the error of the i[th] component in the j[th] channel, and α is a exponent which can take value from 1 to 4.

2.2.3 Research gaps

This error-sensitivity approach depends upon a lot of assumptions. Perceptional quality is best estimated by weighting different noise signals according to their visibility but still some of limitation exists in this VQA model.

2.2.3.1 Quality definition problem

There is no standard definition of image quality. It is not stated anywhere that degradation of quality should be directly related to the visibility of errors. Some errors are clearly visible but not objectionable.

2.2.3.2 Real world images Complexity

Mostly psychophysical experiments are performed by injecting simple figures, such as square waves, bars, or sine waves as an input but images in real world are more complex. So more work has to be done by considering the complexities of real images.

2.2.3.3 Decorrelation Problem

It is assumed that all the dependencies in the input signal are removed during preprocessing but it is not true all the time. When linear channel decomposition like wavelet transform is used strong dependencies still exist.

2.2.3.4 Cognitive Interaction Problem

Quality of image is also influenced by viewing distance, content of the image, interest of the viewer etc. These factors are not considered during assessment of the quality of images.

2.3 IQA model based on structural similarity

Natural image signals are highly structured; their pixels exhibit strong dependencies and these dependencies carry information about the structure of the objects [5]. Wang et al. proposed the model [14] which uses the structural information to determine quality of images. It is assumed that human visual system views and processes an image by extracting its structural information. Therefore, change in the structural information is considered as loss of quality. The approach divides structural similarity into three parts, namely, similarity in luminance, contrast and variation as shown in Fig.2.3. A main difference of the structural similarity index from the previous approach is that quality loss is considered as change in structural information instead of visibility of errors

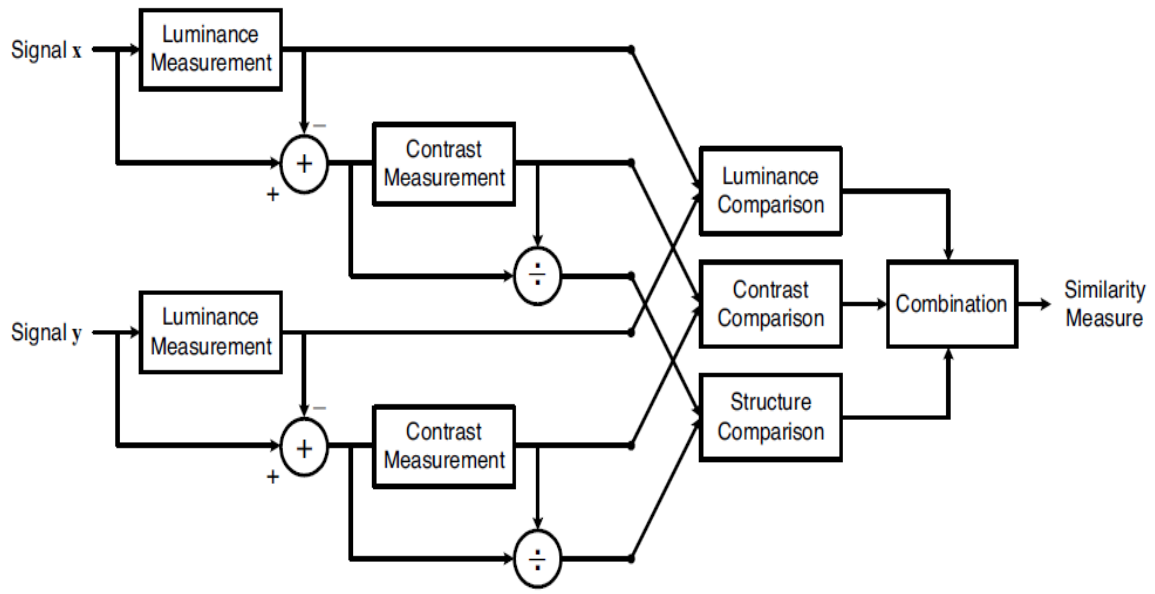


Fig.2.3. System diagram of SSIM [5]

Luminance of signal is calculated by finding the average of pixels intensity.

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{6}$$

Luminance comparison function compares mean of original and distorted signals.

$$l(x,y) = \frac{2\,\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{7}$$

Standard deviation is used to determine contrast of the signal and given by:

$$\sigma_x = \left(\frac{\sum_{i=1}^{N}(x_i - \mu_x)^2}{N-1}\right)^{\frac{1}{2}} \tag{8}$$

Similarly contrast comparison function is given by

$$c(x,y) = \frac{2\,\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{9}$$

Structural comparison is calculated by determining correlation coefficient between original and distorted signal. The correlation coefficient is a quantitative measure of how much one signal

varies in comparison to the other signal. One signal will equal to other when both have the same structure

$$s(x,y) = \frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3} \quad (10)$$

Where

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y) \quad (11)$$

The resulting value of luminance, contrast, and structure comparisons are combined to give SSIM index value.

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\Upsilon} \quad (12)$$

Putting α=1, β= 1 and ϒ =1 we get following equation.

$$SSIM(x,y) = \frac{(2\,\mu_x\mu_y+C_1)(2\,\sigma_{xy}+C_2)}{(\,\mu_x^2+\,\mu_y^2+C_1)(\,\sigma_x^2+\,\sigma_y^2+C_2)} \quad (13)$$

Where C_3, C_2 and C_1 are constants equal to K_1L, K_2L and K_3L respectively. L is a range of pixel intensities. K_1, K_2 and K_3 are constants <<1. SSIM is calculated locally by taking window size of 8 * 8 and window is moved pixel by pixel. The overall quality is given by taking average of all SSIM measures computed locally.

SSIM satisfies the following conditions:

1. **Symmetry**: SSIM give same resultant value when A is compared to B or B is compared to A.

   i.e SSIM(B, A) = SSIM(A, B);

2. **Bounded**: Resultant value of the SSIM measure lies between 0 and 1. SSIM(A, B) ≤ 1.
3. **Unique maximum**: The value of SSIM is one (maximum value) if and only if A is equal to B.

   SSIM(A, B) = 1 if and only if A = B.

2.3.1 Advantage

- The problem of suprathreshold is avoided as it does not depend upon psychophysics to measure quality.

- The problems of decorrelation and real world image complexity are also avoided because it does not predict the quality of image by accumulating errors related with psychophysically understood simple patterns. Instead, quality is determined by comparing structures of both original and distorted images.

  In Fig.2.4, conventional methods fail to explain why the perceived quality of these images is different but it can be easily explain by SSIM. Quality of images keeps on decreasing, as more structural information losses.

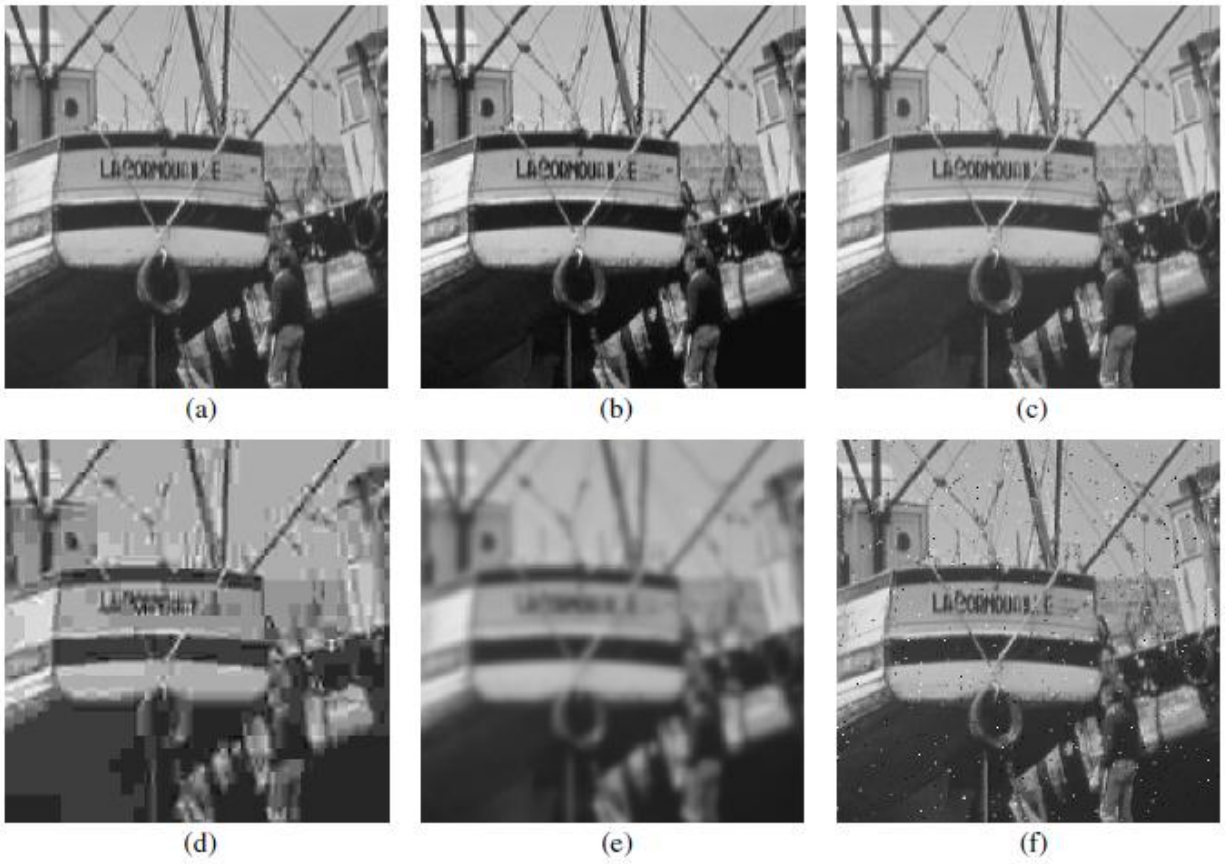


Fig.2.4. Shows images of Boat having same mean square error MSE=210[5].

In Fig.2.4 Image (a) is original image of boat, (b) is contrast stretch image having similarity index equal to 0.9168, (c) is image with shifted mean having SSIM=0.9900, (d) is distorted image of (a) compressed by JPEG compression, having SSIM =0.6949, (e) is blur boat image having SSIM=0.7052 and (f) is boat image containing salt- pepper noise, SSIM=0.7754.

All the images have same mean square error i.e 210. Despite this, perceptual quality of all images is different. It is difficult to explain with error sensitivity method why the images have different quality but it is easily explained with the structural similarity approach as quality of images keeps on decreasing, as more structural information losses.

2.3.2 Limitations

As in SSIM metric, author is using only mean, standard deviation and co-relation to compare the structure of images but two images may have different shape even if they have same mean, variance and covariance. Let us compare proposed SSIM values of pairs of original and processed real world images shown in Fig. 2.5 to DMOS.

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Fig.2.5. Assessment of 6 images from CSIQ [10] database

(a) is the original butter_ flower image and (b) is a blurred image. (c) is an original snow leaves image and (d) is JPEG 2000 compressed image of (c). Similarly, (e) is the original turtle image and (f) contains adding Gaussian white noise. (g) is original bridge images and (h) is a contrast stretch image of (g).

TABLE II. SSIM AND DMOS VALUES OF IMAGES SHOWN IN FIG. 2.5

Comparison between two images	**SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.5131	0.953
SSIM((c),(d))	0.5152	0.757
SSIM((e),(f))	0.5105	0.329
SSIM((g),(h))	0.5123	0.777

As can be seen from the Table II, the SSIM values between all the original and distorted images are nearly the same, whereas the corresponding DMOS values are quite different. Values show that, SSIM is not consistent with subjective ratings.

2.4 Multi-Scale Structural Similarity Index (MSSIM) [7]

Human perceivability of image depends upon the distance of the observer from an image, sampling density and the perceptual capability of user's visual system. Depending upon these factors, subjective rating also varies. A single scale SSIM may appropriate for specific settings only. To incorporate more details of image, Multi Scale SSIM [7] is a convenient way. Original and distorted signals are given as an input to the system. In this low pass filter is applied and then image is down sample by a factor of two, this process is repeated iteratively. The highest scale image is indexed as Scale M, and the original image as Scale 1. At the n-th scale, the structure and contrast comparison are computed and denoted as $c_n(x, y)$ and $s_n(x, y)$, respectively. The luminance, $l_m(x, y)$ is compared only at highest scale M.

All three measures, luminance, structure and contrast comparison, at different scales are combined together to provide overall SSIM evaluation.

$$SSIM(x,y) = [l_m(x,y)]^{\gamma} * \prod_j^M [c_j(x,y)]^{\pi_j} * [s_j(x,y)]^{\beta_j} \quad (14)$$

Importance of components are adjust by setting value of π_j, β_j and Υ. To simplify, all parameters are set to same value $\pi_j = \beta_j = \Upsilon$. Single –Scale SSIM is considered as special case by setting π_j and β_j equal to zero.

2.5 Three-Component Weighted SSIM (3-SSIM) [8]

The performance of SSIM is less effective on the noisy and blurred images. Three weighted component is proposed to address this problem. Image is decomposed into 3 regions smooth, edges and textures region. These three regions have different importance for human perception, thus different weights are given to the SSIM according to the region in which it is calculated. SSIM is calculated in following steps.

1. Compute the SSIM map

2. Decompose the original image into edge, texture and smooth region as shown in Fig.2.6. Region where gradient magnitude is large is marked as edge region, region with low gradient magnitude is marked as smooth region. Textured regions are regions having gradient value in between these two thresholds.

3. Compute SSIM over these 3 regions. The weight for edge regions was taken as 0.5, for textured and smooth regions it was taken as 0.25.

4. Combine the weighted SSIM values, by taking their average to get single value of quality measure.

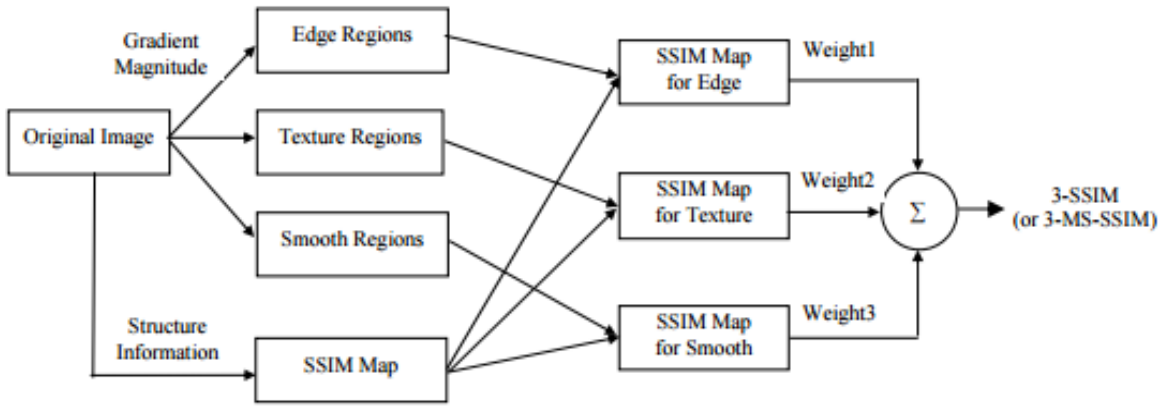


Fig.2.6. System diagram of 3-SSIM [8]

CHAPTER 3

HIGH ORDER MOMENT BASED SSIM

Digital image is nothing but it is representation of how the values of pixels are distributed in 2-D or 3-D space. Two images are structurally similar, if the distribution of intensities or pixel values are same along all directions. Two distributions are exactly same if all moments up to nth order are same. In addition to measurement with the lower order moments like mean, variance and correlation coefficient, measurements based on high order moments are required to represent information about the shape of the distribution. Shape of a distribution might be an important measure for perceptual quality assessment. Obviously, quality of image is degraded when shape of image distribution is not preserved.

High order moments of distribution are used in our algorithm to compute the similarity between the original and processed signals. These moments are descriptors of shape of a distribution.

Third order moment is normalized with respect to standard deviation, known as skewness. It measures the symmetry of the distribution about its mean. It is given by following equation:

$$sk_x = \frac{E[(x-\mu_x)^3]}{{\sigma_x}^3} \tag{15}$$

Two signals X and Y are compared by Skewness measure as:

$$\text{Skewness}(x,y) = \frac{2\,sk_x sk_y + C}{sk_x^2 + sk_y^2 + C} \tag{16}$$

Flatness or peakness of the distribution compared to the normal distribution with the same variance is measured by fourth order moment. It is normalized by dividing it by standard deviation, also known as kurtosis.

$$k_x = \frac{E[(x-\mu_x)^4]}{{\sigma_x}^4} \tag{17}$$

Where k_x is a Kurtosis of distribution X.

Similarly, Kurtosis of two distribution X and Y, i.e distorted and referenced, is compared by :

$$\text{Kurtosis}(x,y) = \frac{2\,k_x k_y + C}{k_x^2 + k_y^2 + C} \tag{18}$$

Fifth and sixth moments can be interpreted in terms of low order moments. Fifth moment measures the importance of tail with respect to mode in causing skew. Fifth moment normalized with respect to standard deviation, is known as hyperskewness, can be given as:

$$hsk_x = \frac{E[(x-\mu_x)^5]}{{\sigma_x}^5} \tag{19}$$

Hyperskewness of original signal X and distorted signal Y can be compare as:

$$\text{hyperskewness}(x,y) = \frac{2\,hsk_x hsk_y + C}{hsk_x^2 + hsk_y^2 + C} \tag{20}$$

Sixth moment can be used to describe further shape parameters.

In the same way we calculate and compare sixth standardized moment (hyperflattness).

$$hk_x = \frac{E[(x-\mu_x)^6]}{{\sigma_x}^6} \tag{21}$$

$$\text{hyperflattness}(x,y) = \frac{2\,hk_x hk_y + C}{hk_x^2 + hk_y^2 + C} \tag{22}$$

Kurtosis, Skewness, hyperskewness and hyperflattness value may go to infinity; to deal with these pathological cases, (16), (18), (20) and (22) would satisfy following conditions.

General representation of (16), (18), (20) and (22) can be given as

$$\frac{2\,AB + Const}{A^2 + B^2 + Const} \tag{23}$$

TABLE III. SHOWS THE PATHOLOGICAL CASES FOR (23)

Value of A Considered	**Value of B considered**	**Value of expression ( )**
∞	Real value	0
Real Value	∞	0
∞	∞	1
- ∞	-∞	1
∞	-∞	-1
- ∞	∞	-1

Along with the high order standardized moment, central moments, coskewness and cokurtosis, are also used in our approach to measure how the both original and distorted signals change together.

Coskewness is a measure of how the shape of two distributions is changing together. It measures the symmetry of probability distribution of one signal with respect to the probability distribution of another signal.

Coskewness of two variables X and Y are defined by two equations:

$$\frac{E[(x-\mu_x)^2(y-\mu_y)]}{\sigma_x{}^2\sigma_y} \tag{24}$$

$$\frac{E[(x-\mu_x)(y-\mu_y)^2]}{\sigma_x\sigma_y{}^2} \tag{25}$$

These measures are normalized with respect to skewness. All values are shifted to positive after subtracting mean from each pixel values to obtain the value of coskewness between 0 and 1.

Let $x-\mu_x$ is denoted by x_m, $y-\mu_y$ by y_m and min_m is the minimum value of x_m and y_m values. After applying all the normalization equation can be given as

$$\text{Coskewness1(X, Y)} = \frac{E[(x_m-min_m)^2(y_m-min_m)]+C}{\sigma_x{}^2\sigma_y(\sqrt[3]{sk_x})^2\sqrt[3]{sk_y}+C} \tag{26}$$

$$\text{Coskewness2}(X,Y) = \frac{E[(x_m - min_m)(y_m - min_m)^2] + C}{\sigma_x \sigma_y{}^2 \sqrt[3]{sk_x}(\sqrt[3]{sk_y})^2 + C} \quad (27)$$

Cokurtosis measures the degree of flatness or peakness of one distribution with respect to other distribution. Higher cokurtosis value means both distributions tend to go their extreme values at same time. Cokurtosis values are normalized with respect to kurtosis.

Normalized Cokurtosis of two distributions 'X' and 'Y' can be given by three equations.

$$\text{Cokurtosis1}(X,Y) = \frac{E[(x-\mu_x)^3(y-\mu_y)] + C}{\sigma_x{}^3 \ \sigma_y \ (\sqrt[4]{k_x})^3 \sqrt[4]{k_y} + C} \quad (28)$$

$$\text{Cokurtosis2}(X,Y) = \frac{E[(x-\mu_x)(y-\mu_y)^3] + C}{\sigma_x \ \sigma_y{}^3 \ \sqrt[4]{k_x} \ (\sqrt[4]{k_y})^3 + C} \quad (29)$$

$$\text{Cokurtosis3}(X,Y) = \frac{E[(x-\mu_x)^2(y-\mu_y)^2] + C}{\sigma_x{}^2 \ \sigma_y{}^2 \ (\sqrt[4]{k_x})^2 \ (\sqrt[4]{k_y})^2 + C} \quad (30)$$

Where E[X] is an expectation value of X.

Along with mean, variance and covariance in (7), (9) and (10) are combined with (16), (18), (20), (22), (26), (27), (28), (29) and (30) to get a new SSIM index as shown in Fig 3.1, which is given as

$$SSIM(x,y) = [l(x,y)]^{\alpha}.[c(x,y)]^{\beta}.[s(x,y)]^{\Upsilon}.[S(x,y)]^{\rho}.[K(x,y)]^{\rho}.$$
$$[HS(x,y)]^{\rho}.[HK(x,y)]^{\rho}.[CS_1(x,y)]^{\rho}.[CS_2(x,y)]^{\rho}.$$
$$[CK_1(x,y)]^{\rho}.[CK_2(x,y)]^{\rho}.[CK_3(x,y)]^{\rho} \quad (31)$$

Where α, β, ϒ and ρ are constants used to set comparable importance of these components. C is constant added to avoid a situation when numerator or denominator converges to zero. C is equal to $(kL)^2$, range of pixel values is denoted by L and k << 1. For grey scale images pixel ranges varies from 0 to 255, so L is 255.

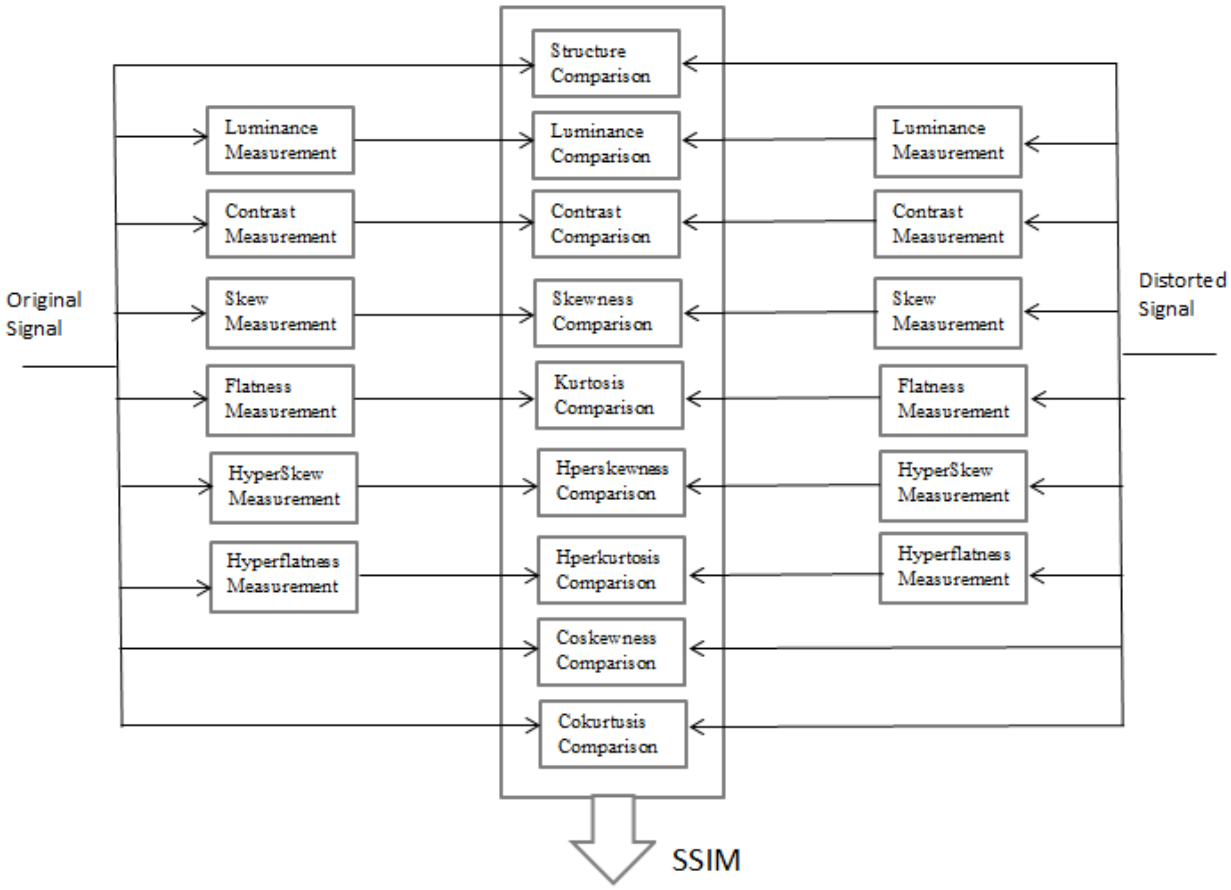


Fig.3.1. System diagram of propose method

3.1 Limitations overcome

The proposed SSIM is computed locally, taking window size of 8*8, for the images shown in Fig. 1.4 and the various SSIM values obtained are given in Table IV.

TABLE IV. PROPOSED SSIM VALUES FOR DIFFERENT PAIRS OF DISTRIBUTIONS SHOWN IN FIG. 1.4

Comparison between Two distributions	**SSIM Value (proposed method)**
SSIM(A,B)	0.002
SSIM(A,C)	*0.011*
SSIM(B,C)	*0.004*
SSIM(A,D)	*0.021*
SSIM(B,D)	*0.013*
SSIM(C,D)	*0.008*

As seen in the histograms, shown in Fig.1.4, all the distributions are of different shapes. New proposed SSIM does better than the existing SSIM [5] in distinguishing the distributions. Results are also taken on the real world images shown in Fig.2.5.

TABLE V. PROPOSED SSIM VALUES AGAINST DMOS VALUES OF IMAGES SHOWN IN FIG.2.5

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.3816	0.953
SSIM((c),(d))	0.4661	0.757
SSIM((e),(f))	0.5063	0.329
SSIM((g),(h))	0.3697	0.777

According to DMOS values, image 2(b) has more distortion compared to other images and 2(f) has least distortion present. Same is indicated with proposed SSIM values unlike the existing SSIM [5] seen from Table I.

CHAPTER 4

RESULTS AND DISCUSSION

4.1. Experiment1:

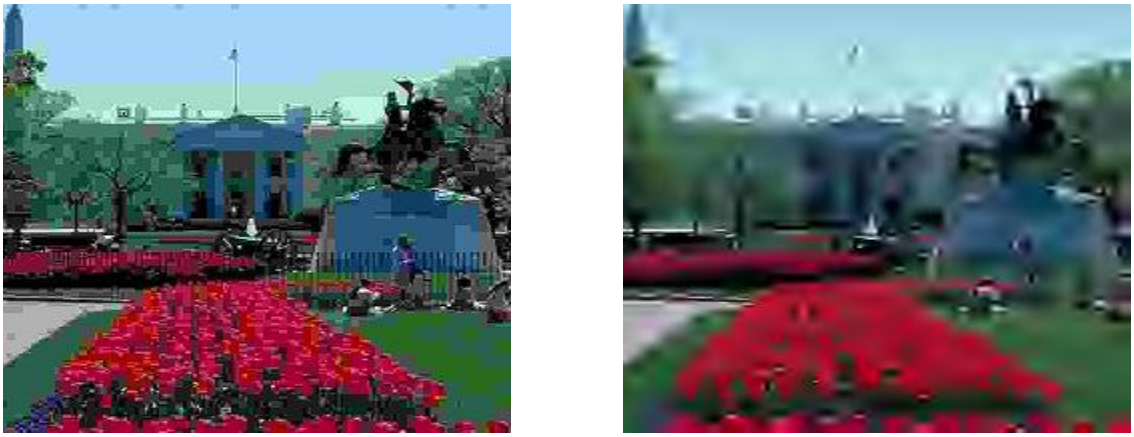
Proposed SSIM is validated against existing methods and subjective ratings by taking results on various kinds of distortion. The proposed SSIM is calculated only on luminance component of the shown images in Fig.4.1. The values of proposed SSIM are given in Table VI.

(a)

(b)

(c)

(d)



(e) (f)

Fig.4.1. Visual assessment of 6 set of images from CSIQ [10] database.

Image (a) is original image and (b) is a blur image, (c) is the image with different contrast, (d) contains additive pink noise, (e) is compressed image compress by JPEG and (f) is further compressed image by JPEG2000.

The proposed SSIM is calculated on images shown in Fig.4.1. Results are shown in Table VI.

TABLE VI. PROPOSED SSIM VALUES AGAINST DMOS VALUES OF IMAGES SHOWN IN FIG.4.1

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.3411	0.750
SSIM((a),(c))	0.7293	0.371
SSIM((a),(d))	0.6549	0.565
SSIM((a),(e))	0.5293	0.657
SSIM((a),(f))	0.355	0.827

It is clearly seen from the results, the proposed SSIM measure is consistent with subjective ratings, shown in Table VI.

Scatter plots of the subjective ratings (DMOS) versus (a) Mean Square Error (MSE) (b) SSIM applied locally on the window size 8*8 (C) proposed SSIM applied locally on the window size

8*8 over 900 images of CSIQ database are shown respectively in Fig. 4.2, 4.3 and 4.4 to compare the performance of proposed SSIM with MSE and existing SSIM.

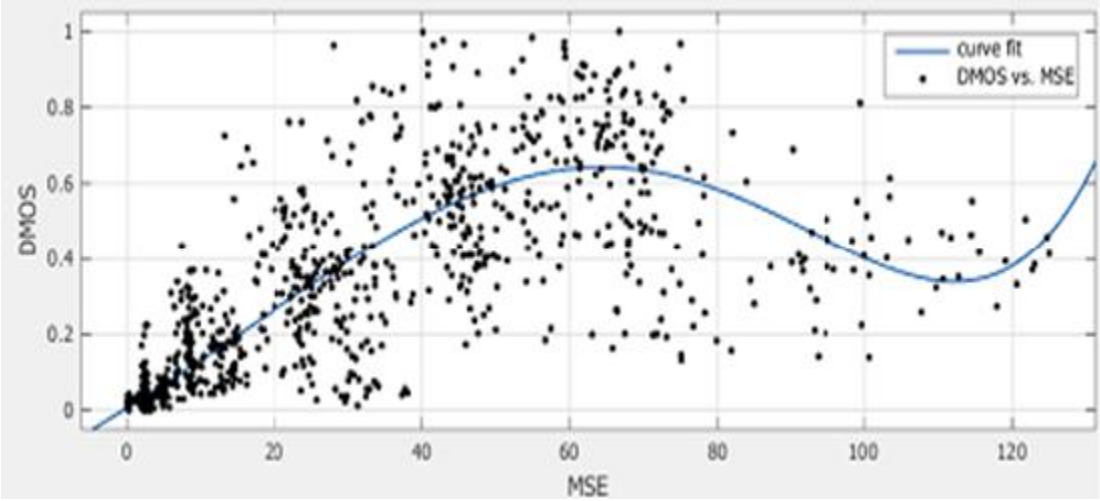


Fig.4.2. Plot of DMOS VS MSE

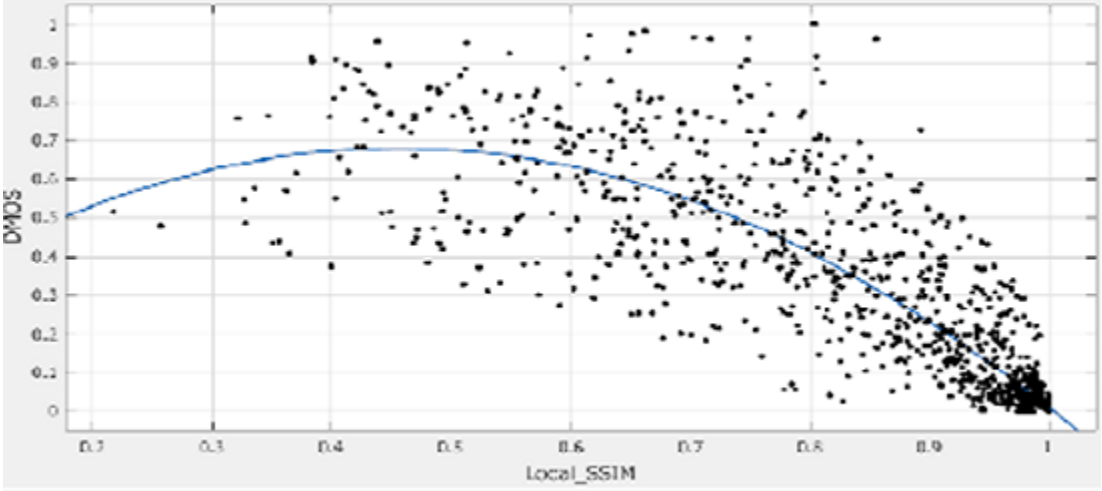


Fig.4.3. Plot of DMOS VS SSIM

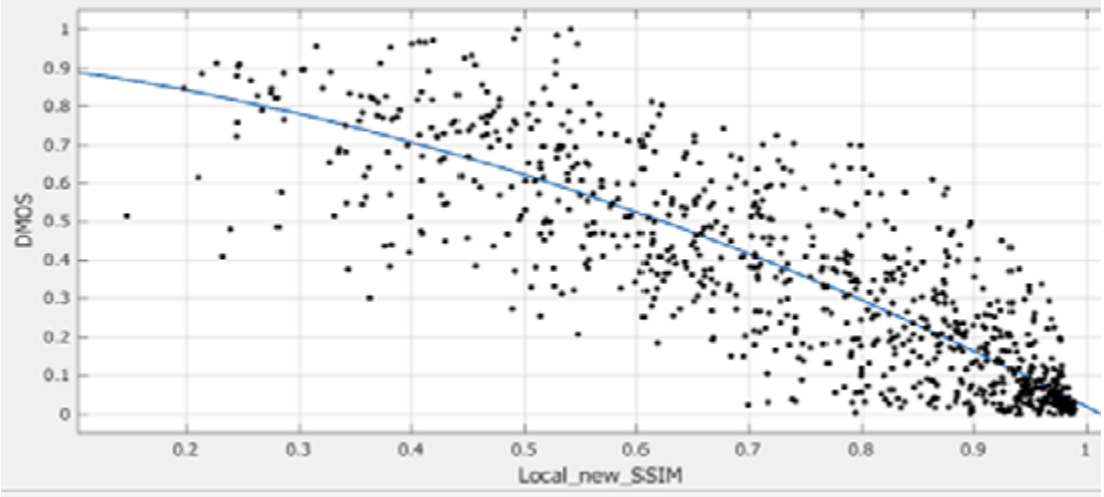


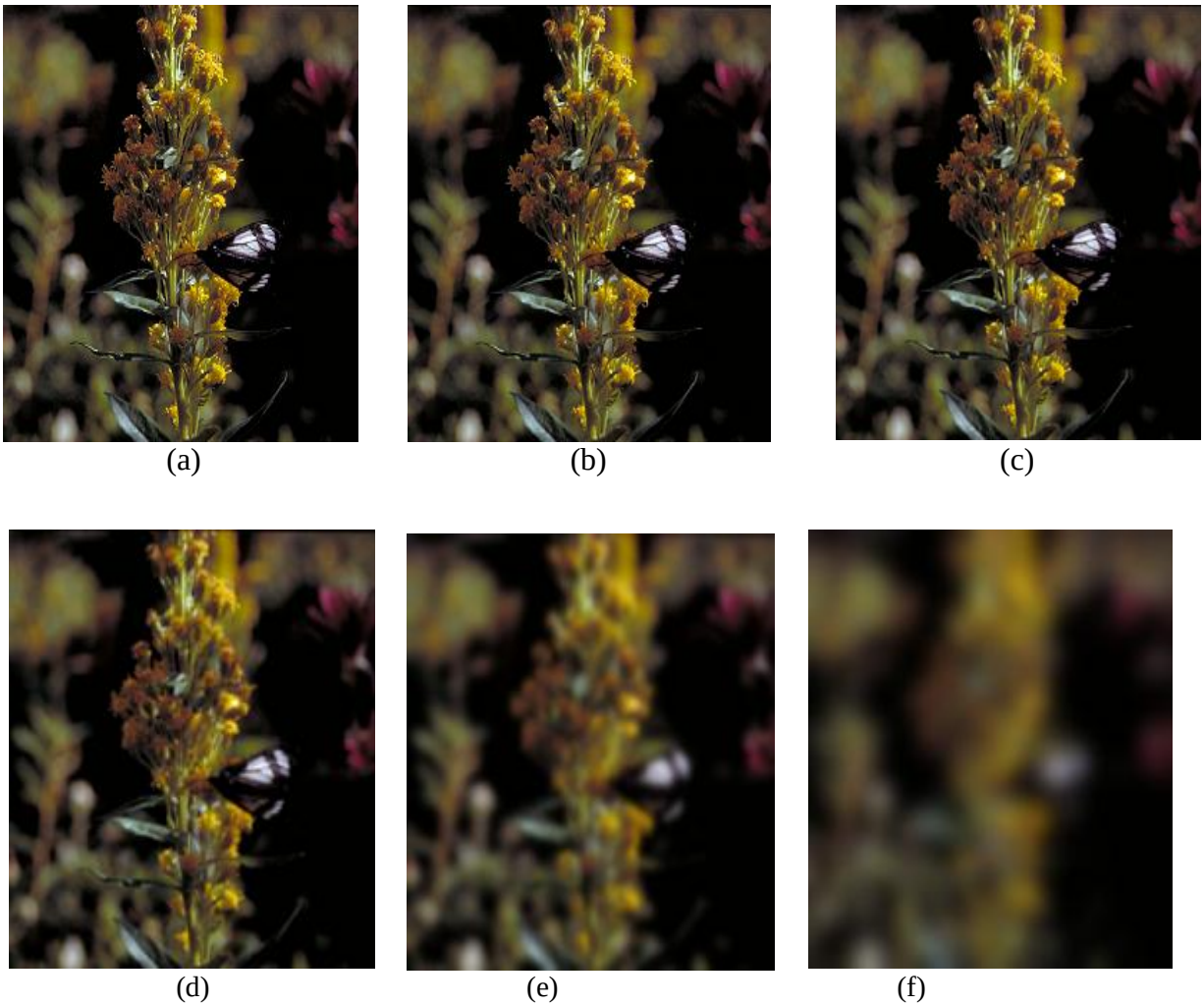
Fig.4.4. Plot of DMOS VS proposed SSIM

In the plot of Fig.4.2, Fig.4.3 and Fig.4.4, each black point corresponds to image quality measured by the objective approach versus subjective rating of that image. Plots show the relation between DMOS and objective approach. It is seen that the proposed SSIM goes better with DMOS values than other methods.

4.2. Experiment2:

Results are taken on images with different extents of various types of distortions to find how well the proposed method agrees with human perception. The CSIQ dataset, which contains 900 images, is considered. The dataset contains undistorted images and corresponding set of multiple distorted images. The distortions considered belong to six different categories, namely, due to additive white Gaussian and pink noise, blur, contrast, JPEG and JPEG 2000 compression. Different extents of these distortions are also considered. The proposed SSIM values are compared with the subjective DMOS measure for four categories of distortions, owing to similarity in observations within images with noise and compression distortions.

4.2.1. Distortion type- Blur



(a) (b) (c)

(d) (e) (f)

Fig.4.5. Image (a) is original image and (b), (c), (d) and (e) are images with different level of blurriness present [10].

Proposed SSI M is applied on images with various levels of blur, and the values are shown in Table VII in comparison to the corresponding DMOS values. As can be seen, the proposed measure is in agreement with human perception.

TABLE VII. PROPOSED STRUCTURAL SIMILARITY INDEX VALUES OF IMAGES SHOWN IN FIG.4.5

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.9124	0.078
SSIM((a),(c))	0.7765	0.202
SSIM((a),(d))	0.6075	0.482
SSIM((a),(e))	0.4181	0.779
SSIM((a),(f))	0.2451	0.953

In order to compare the proposed SSIM measure to the existing SSIM and MSE measures, scatter plots of the subjective ratings (DMOS) versus (a) MSE (b) Existing SSIM (c) proposed SSIM is shown respectively in Fig. 4.6, 4.7 and 4.8. The SSIM measures are obtained by computing it locally on the window size 8*8 over 150 **blurred** images of CSIQ database. As can be seen, the linear correlation between human percept ion and the proposed SSIM is the highest.

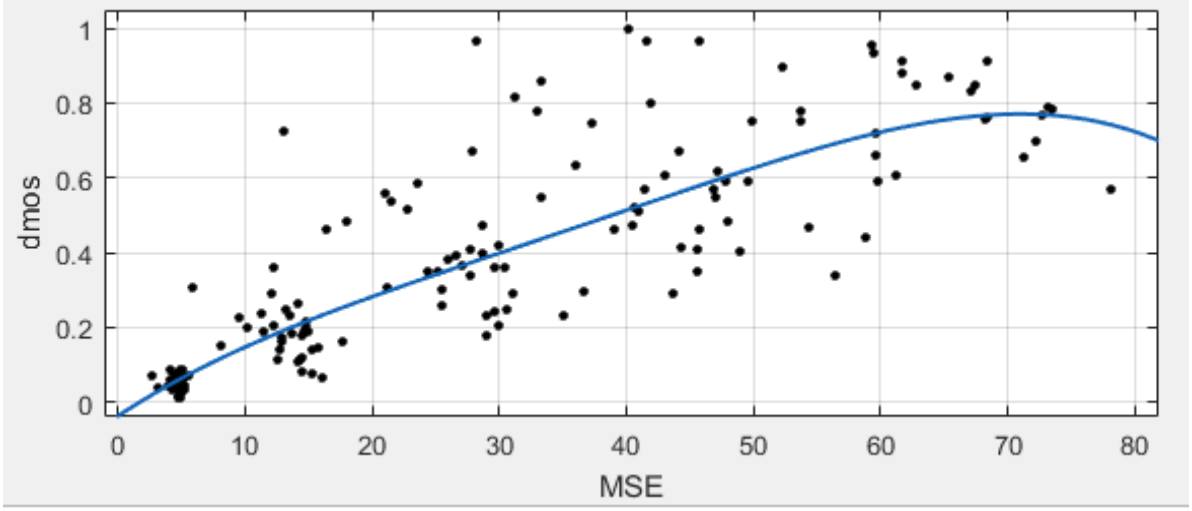


Fig.4.6. Plot of DMOS VS MSE for blurred images

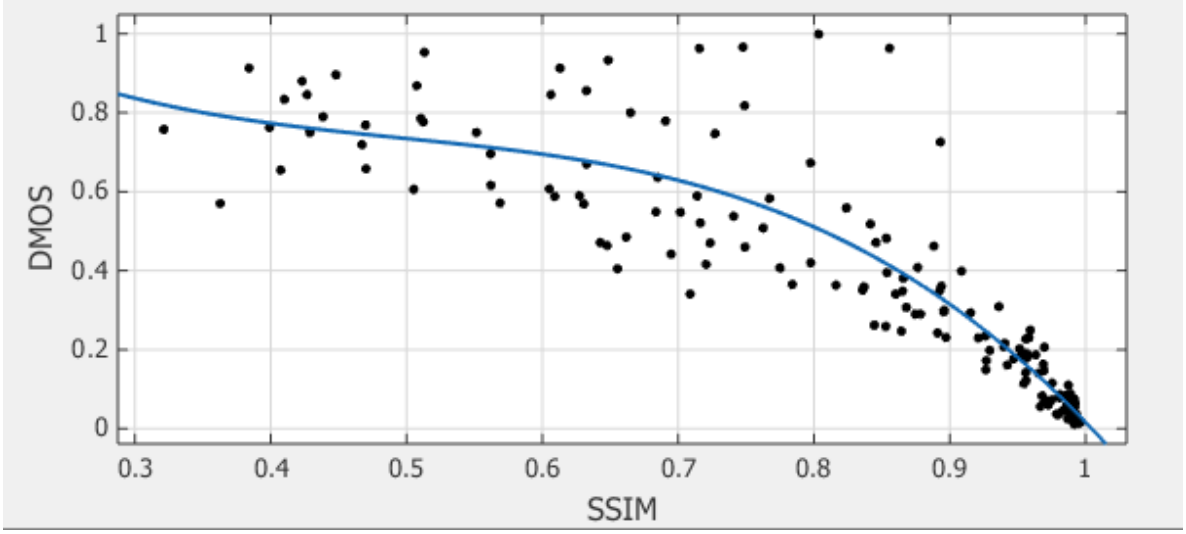


Fig.4.7. Plot of DMOS VS SSIM for blurred images

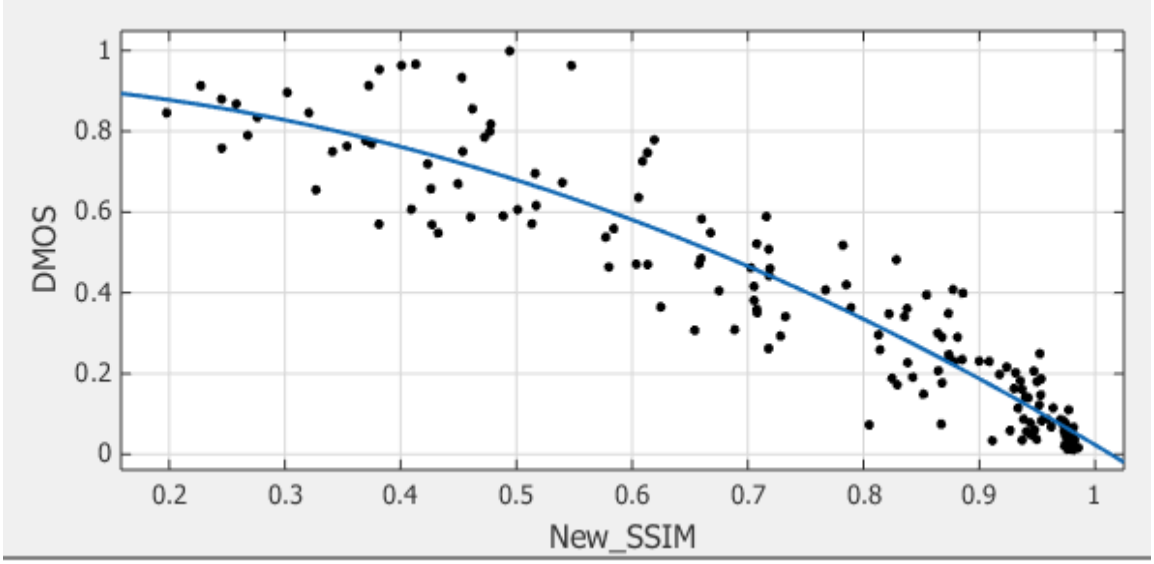


Fig.4.8. Plot of DMOS VS proposed SSIM for blurred images

4.2.2. Distortion type- contrast

Proposed SSIM is applied on images with various levels of contrast, and the values are shown in Table VIII in comparison to the corresponding DMOS values. As can be seen, the proposed measure is almost in agreement with human perception.

(a) (b) (c)

(d) (e) (f)

Fig.4.9. Image (a) is original image and (b), (c), (d) and (e) are images with different level of contrast present [10].

TABLE VIII. PROPOSED STRUCTURAL SIMILARITY INDEX VALUES OF IMAGES SHOWN IN FIG.4.9

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.9368	0.090
SSIM((a),(c))	0.8525	0.212
SSIM((a),(d))	0.5806	0.342
SSIM((a),(e))	0.4635	0.389
SSIM((a),(f))	0.4521	0.398

In order to compare the proposed SSIM measure to the existing SSIM and MSE measures, scatter plots of the subjective ratings (DMOS) versus (a) MSE (b) Existing SSIM (c) proposed SSIM is shown respectively in Fig. 4.10, 4.11 and 4.12. The SSIM measures are obtained by computing it locally on the window size 8*8 over 150 **contrast altered** images of CSIQ database. As can be seen, the linear correlation between human perception and the proposed SSIM is the highest.

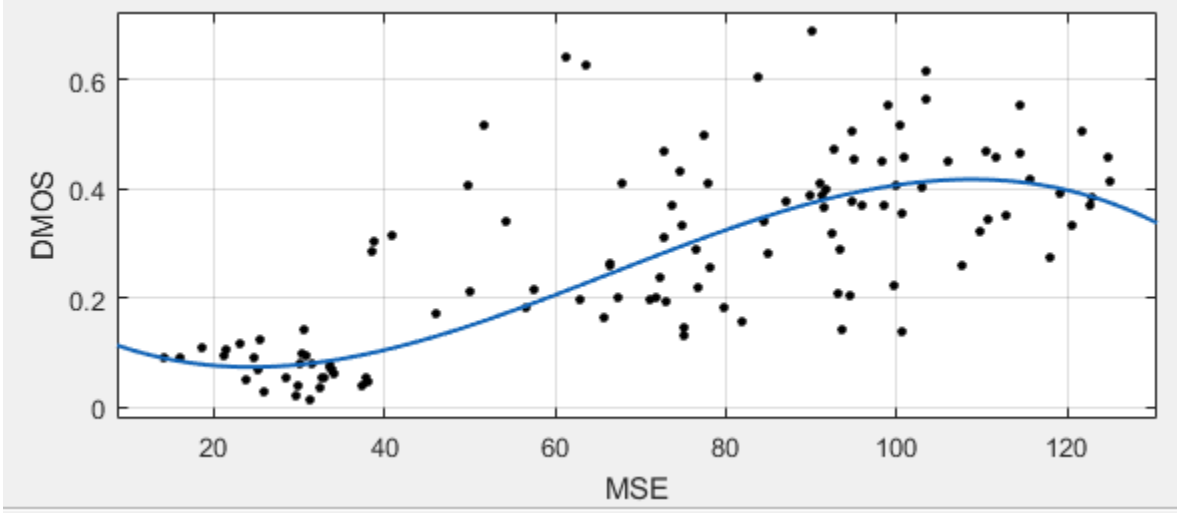


Fig.4.10. Plot of DMOS VS MSE for contrast altered images

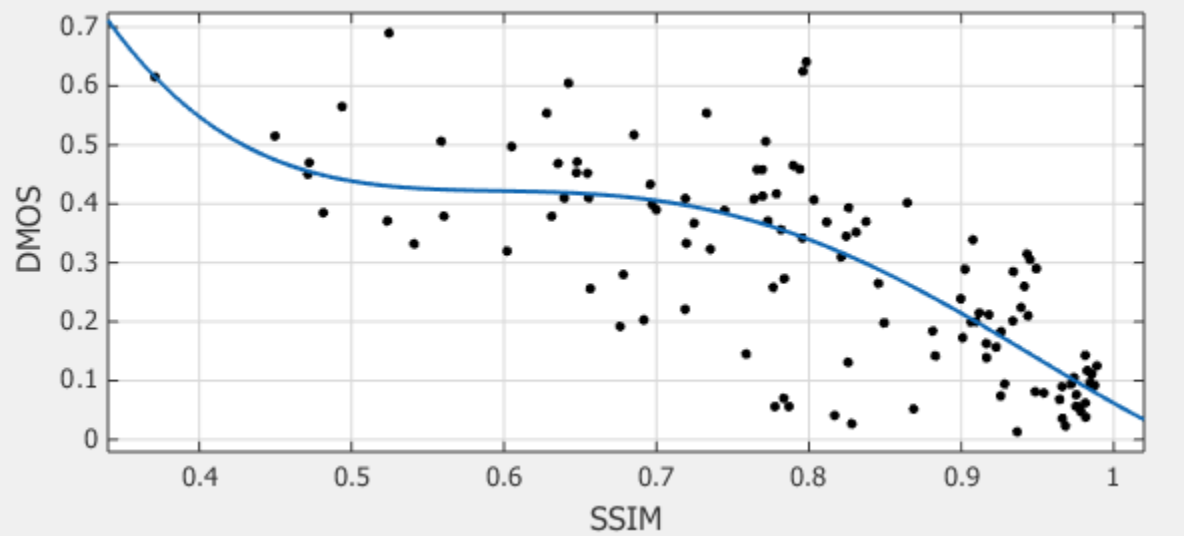


Fig.4.11. Plot of DMOS VS SSIM for contrast altered images

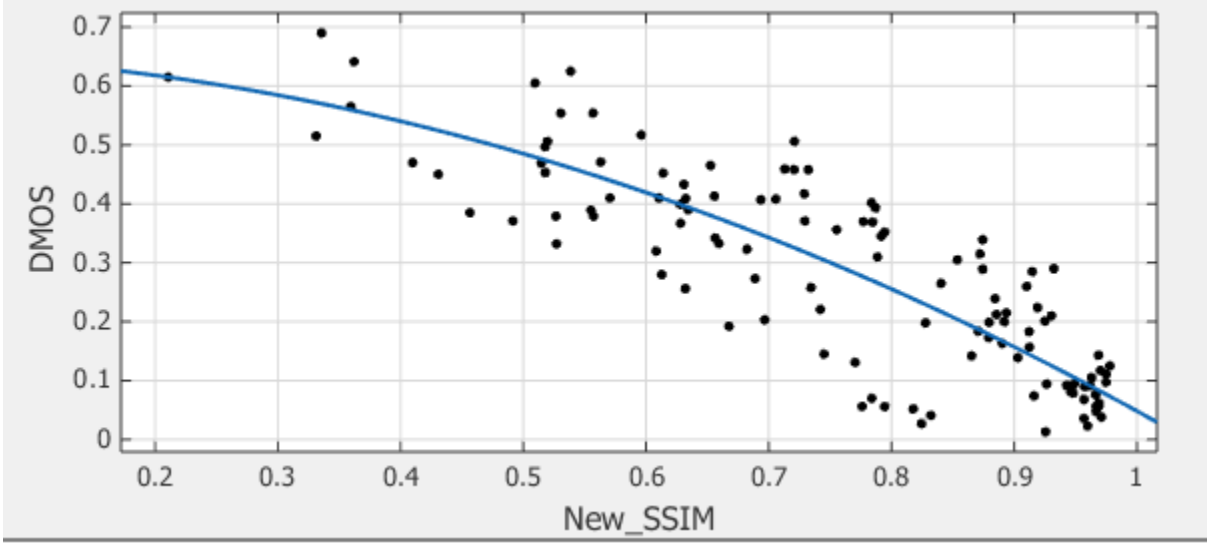


Fig.4.12. Plot of DMOS VS proposed SSIM for contrast stretch images

4.2.3. Distortion type- additive pink noise

Proposed SSIM is applied on images with various levels of additive pink noise, and the values are shown in Table IX in comparison to the corresponding DMOS values. As can be seen, the proposed measure is in agreement with human perception.

(a) (b) (c)

(d) (e) (f)

Fig.4.13. Image (a) is original image and (b), (c), (d) and (e) are images containing different level of additive pink noise [10].

TABLE IX. PROPOSED STRUCTURAL SIMILARITY INDEX VALUES OF IMAGES SHOWN IN FIG.4.13

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	0.7582	0.032
SSIM((a),(c))	0.5948	0.193
SSIM((a),(d))	0.4759	0.352
SSIM((a),(e))	0.3537	0.505
SSIM((a),(f))	0.2813	0.642

Scatter plots of the subjective ratings (D MOS) versus (a) MSE (b) Existing SSIM (c) proposed SSIM is shown respectively in Fig.4.14, 4.15 and 4.16 As can be seen, both existing SSIM and proposed S SIM are not in linear correlation with human perception. Structural similarity looks less suitable to evaluate images corrupted with pink noise. Similar observations are made on images with additive white Gaussian noise.

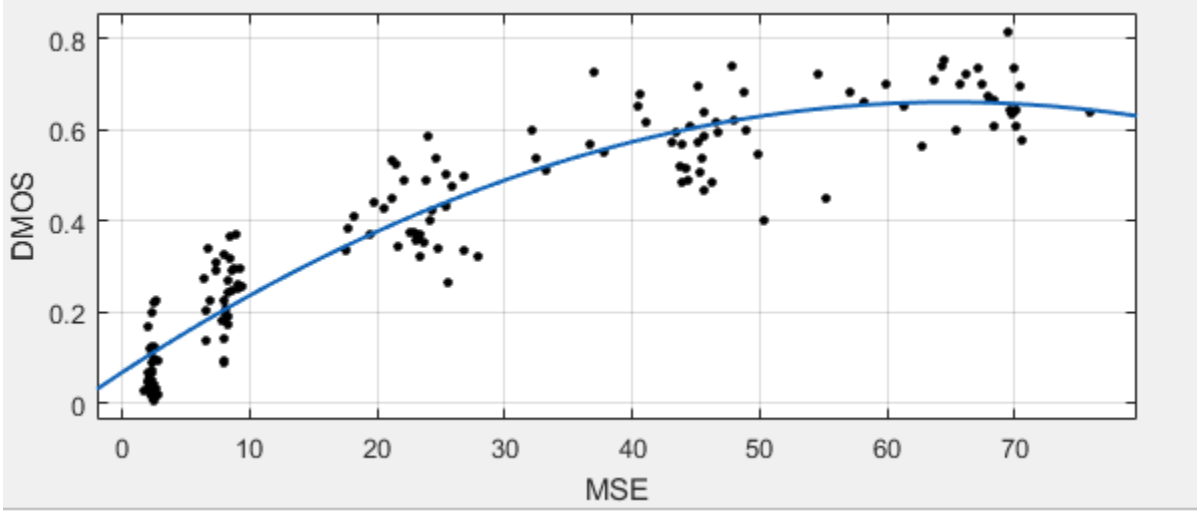


Fig.4.14. Plot of DMOS VS MSE for images having additive pink noise

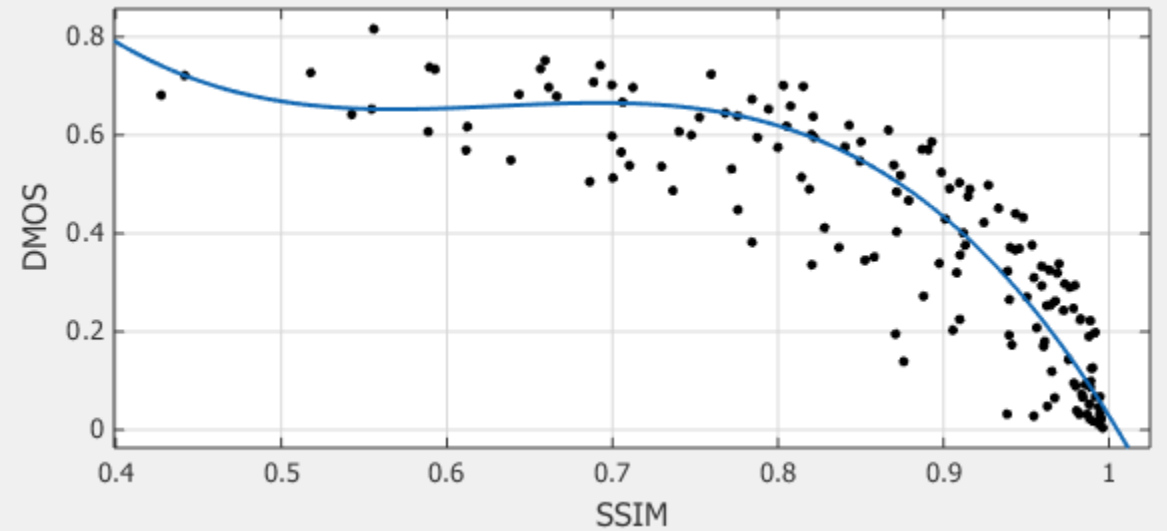


Fig.4.15. Plot of DMOS VS SSIM for images having additive pink noise

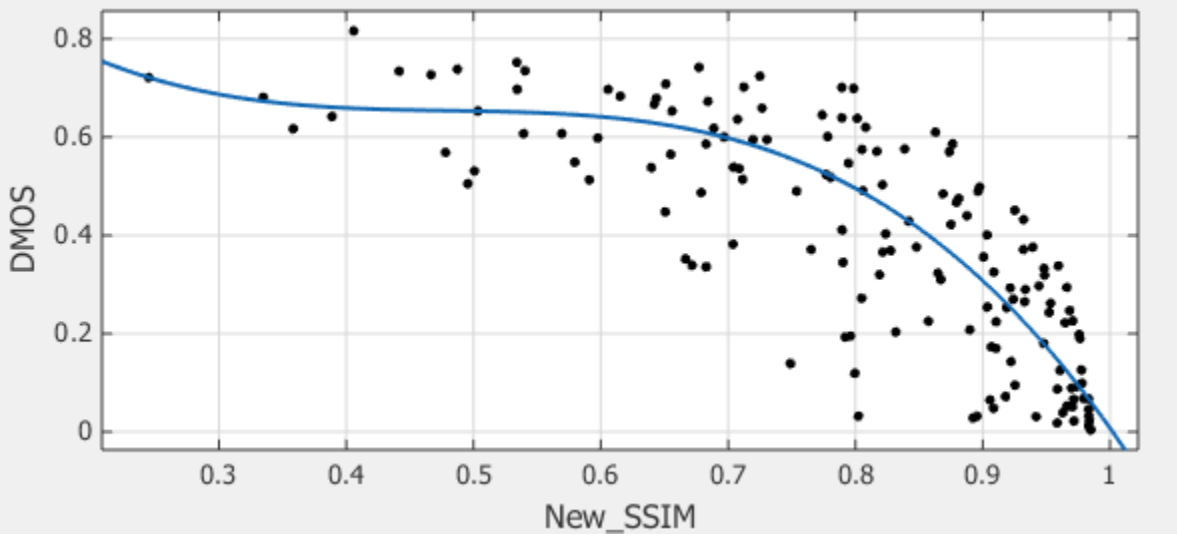


Fig.4.16. Plot of DMOS VS proposed SSIM for images having additive pink noise

4.2.4. Distortion type- compressed by JPEG 2000

Similarly, Propsed SSIM is applied on image of sky, compressed with different level of compression, values are shown in Table X.

(a) (b) (c)

(d) (e) (f)

Fig.4.17. Image (a) is original image and (b), (c), (d) and (e) are images containing different level of JPEG 2000 compressed [10].

TABLE X. PROPOSED STRUCTURAL SIMILARITY INDEX VALUES OF IMAGES SHOWN IN FIG.4.17

Comparison between two images	**Proposed SSIM Value**	**DMOS (Subjective rating)**
SSIM((a),(b))	*0.7582*	0.024
SSIM((a),(c))	*0.5948*	0.301
SSIM((a),(d))	*0.4759*	0.655
SSIM((a),(e))	*0.3537*	0.851
SSIM((a),(f))	*0.2813*	1.000

It is clearly seen from the results, the proposed SSIM measure quantitatively reflects in more degradation in quality. Scatter plots of the subjective ratings (DMOS) versus (a) Existing SSIM (b) proposed SSIM applied locally on the window size 8*8 over 150 images of CSIQ database, **compressed by JPEG compression,** are shown respectively in Fig. 4.18, 4.19 and 4.20 to differentiate the performance of proposed SSIM with already existing SSIM and MSE.

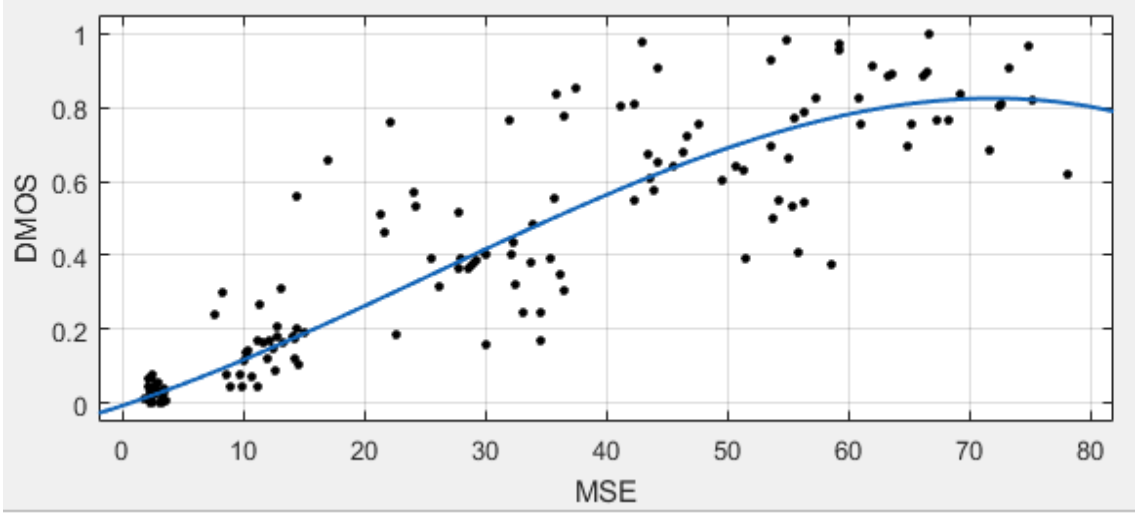


Fig.4.18. Plot of DMOS VS MSE for compressed images by JPEG 2000

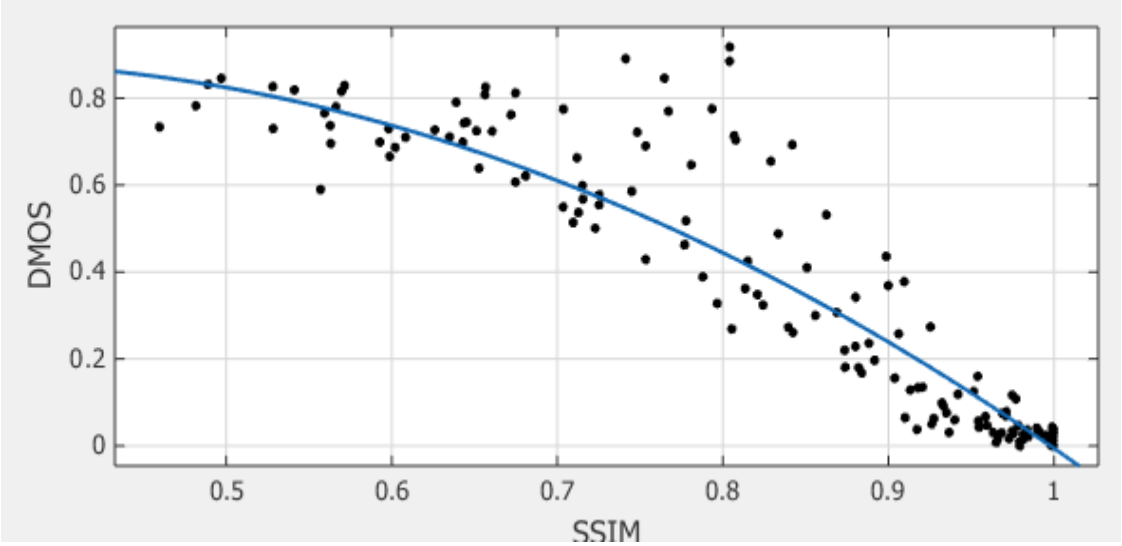


Fig.4.19. Plot of DMOS VS SSIM for compressed images by JPEG 2000

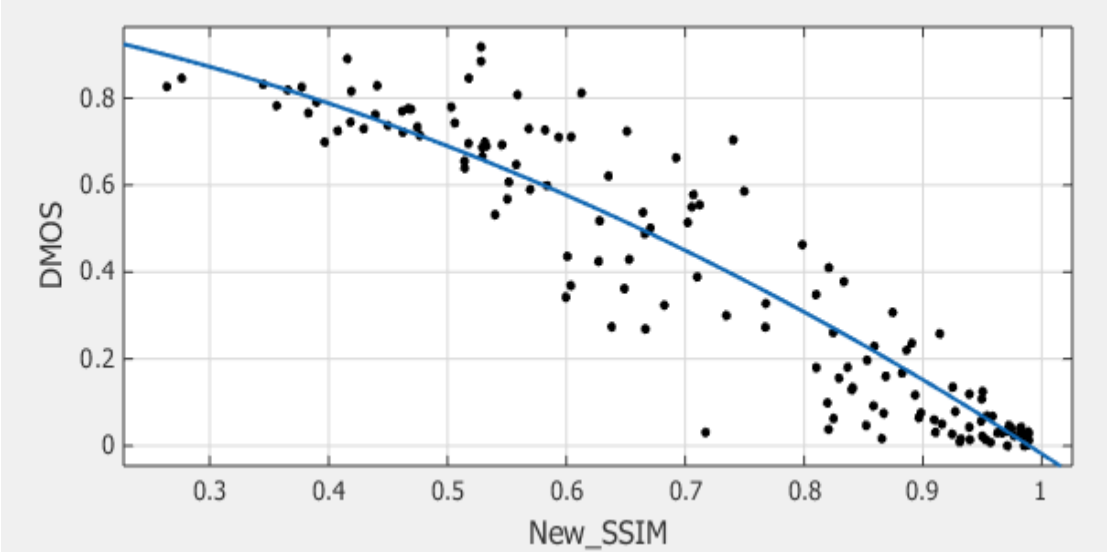


Fig.4.20. Plot of DMOS VS proposed SSIM for compressed images by JPEG 2000

It is clearly seen from the plots, the proposed SSIM measure goes with human perception better than other methods.

4.3. Experiment 3

The performance of the proposed SSIM is compared with existing SSIM and MSE by using Pearson linear correlation coefficient (CC), Kendall rank correlation coefficient (KROCC) and Spearman's rank correlation coefficient (SROCC). The CC criteria help to define for prediction accuracy, whereas, KROCC and SROCC provides prediction monotonicity. Results are taken on the different type of distorted images of CSIQ database. The value of these performance measures needs to be near -1 for SSIM and proposed SSIM, near 1 for MSE.

TABLE XI. CORRELATION METRICS FOR MSE, AND PROPOSED AND EXISTING SSIM WITH DMOS

Type Of Distort-ion	SROCC			KROCC			CC		
	NEW SSIM	SSIM	MSE	NEW SSIM	SSIM	MSE	NEW SSIM	SSIM	MSE
Gaussian Noise	-0.8453	-0.9255	0.9239	-0.6435	-0.7544	0.7493	-0.8219	-0.8941	0.9208
Blur	-0.9350	-0.9245	0.8622	-0.7720	-0.7665	0.6598	-0.9264	-0.8504	0.8243
Contrast	-0.8263	-0.7393	0.7125	-0.6258	-0.5322	0.4981	-0.8338	-0.7221	0.7095
Pink noise	-0.8045	-0.8925	0.9264	-0.6024	-0.7075	0.7528	-0.7489	-0.8117	0.9001
JPEG	-0.9305	-0.9205	0.8867	-0.7629	-0.7526	0.6934	-0.9356	-0.9019	0.8640
JPEG 2000	-0.9100	-0.9172	0.8841	-0.7260	-0.7497	0.7141	-0.9119	-0.8749	0.8841

A better IQA measure should have higher negative CC, SROCC, and KROCC values. It is seen that the proposed SSIM does better than existing SSIM and MSE in all the cases except the cases of noises. In fact, surprisingly, for noises, MSE seems to be doing the best, essentially pointing out that here structural similarity becomes less important than pixel-wise similarity.

CHAPTER 5

CONCLUSION AND FUTURE WORK

IN this thesis, we have extended the SSIM measure by introducing distribution shape parameters in the existing measure. Higher order moments are used as the distribution shape descriptors. It is shown that use of higher order moments adds useful information to SSIM relevant to structural similarity between image regions. Experimental results are taken on various types of distorted images from the publicly available CSIQ dataset which contains more than 900 images. Subjective rating by users is used as ground truth to compare the proposed SSIM to MSE and existing SSIM. From the outcome of the comparison, it is found that the proposed technique is an effective method of evaluating the image quality with better correspondence with human perception.

Following are the areas where proposed work can be extended.

- As said earlier, image quality measures can be used in many applications. One such application is image and video coding. Usually full reference quality of the image is checked after decoding the image. In future, it is intended that the encoding itself is adapted to both no-reference and full reference quality of the image. The goal is especially to design such a technique for region-of interest (ROI) based image coding, such that the quality at non-ROI areas does not get degraded to unacceptable amounts. The intension is to test the new method on many images and videos with different characteristics and carry out statistical analysis of performance.
- As shown in the results, the proposed SSIM does better than existing SSIM and MSE in all the cases except the cases of noises. So more work has to be done regarding this.

BIBLIOGRAPHY

[1]. Shyamprasad Chikkerur and Vijay Sundaram, "Objective video quality assessment methods : a classification, review and performance comparison", IEEE Trans. on broadcasting , vol. 57(II) , 2011.

[2]. A. G. George and A. K. Prabavathy, "A survey on different approaches used in image quality assessment," international journal of emerging technology and advanced engineering, vol.3, february 2013.

[3]. B. Girod, "What's wrong with mean-squared error," in Digital Images and Human Vision, A. B. Watson, Ed. Cambridge, MA: MIT Press,1993, pp. 207–220.

[4]. S. Winkler and P. Mohandas, "The evolution of video quality measurement: from PSNR to hybrid metrics", IEEE Trans . on broadcasting , vol. 54(3) , 2008.

[5]. Zhou Wang , Bovik, A.C. ,Sheikh, H.R. and Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Trans . on image processing, Vol. 13(4) , 2004 .

[6]. A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance,"IEEE Trans. Commun., Vol. 43, pp. 2959–2965, Dec. 1995.

[7]. Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," IEEE, asilomar Conf. signals, pp. 1398–1402, Nov. 2003.

[8]. C. Li and A. C. Bovik, "Three-component weighted structural similarity index," SPIE, vol. 7242, pp. 72420q-1–72420q-9, 2009.

[9]. M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "Complex wavelet structural similarity: a new image similarity index," IEEE Trans. image processing, vol. 18(11), pp. 2385–2401, nov.2009.

[10]. E. C. Larson and D. M. Chandler, "Categorical image quality (csiq) database," [online] http://vision. okstate. edu/csiq, 2010.

[11]. Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," IEEE Trans. Image Processing, vol. 20(5), pp. 1185–1198, May 2011.

[12]. AM Rohaly, J Libert, PCorriveau, A Webster "Final report from the video quality experts group on the validation of objective models of video quality assessment 2000." IEEE Trans. Image Processing, vol. 20(5), pp. 1185–1198, 2011.

[13]. Y. Luo and X. Tang, "Photo and video quality evaluation: focusing on the subject", 16th international conf. on multimedia, 2008.

[14]. Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," IEEE Trans. signal processing.: image communication, Vol. 19(2), 2004.

[15]. L. Zhang, D. Zhang, X. Mou, and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," IEEE Trans. Image Processing, Vol. 20(8), pp. 2378–2386, 2011.

[16]. K. Egiazarian, J. Astola, V. Lukin, F. Battisti, and M. Carli, "New full-reference quality metrics based on HVS," 2006.

[17]. F. Zhang, S. Li, L. Ma, and K. N. Ngan, "Limitation and challenges of image quality measurement," Vol. 7744, pp. 774402–774408, 2010.

[18]. K. Seshadrinathan and A. Bovik, "Automatic prediction of perceptual quality of multimedia signals—a survey," Multimedia Tools Appl., Vol. 51(1), pp. 163–186, 2011.

[19]. W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," J. Visual Communication Image Represent., Vol. 22(4), pp. 297–312, May 2011.