9th Iranian Joint Congress on Fuzzy and Intelligent Systems
March 2-4, 2022, Bam, Kerman, Iran

# A New Edge and Pixel-Based Image Quality Assessment Metric for Colour and Depth Images

Seyed Muhammad Hossein Mousavi
Independent Scientist
Tehran, Iran
mosavi.a.i.buali@gmail.com

S. Muhammad Hassan Mosavi
Independent Scientist
Tehran, Iran
se.mu.ha.mousavi@gmail.com

***Abstract*— Measuring the quality of digital image is a complicated task and has high of importance in the field of image processing. Using edge quality metrics, it is possible to assess the power of edge detectors. Also, pixel and edge-based metrics are so crucial in dealing with a digital image. So, combination of edge and pixel features could handle not all but, almost all aspects of an image. Most recently using edge-based image quality metrics are popular, due to weakness of traditional image quality assessment metrics such as Mean Squared Error (MSE), Peak Signal-to Noise Ratio (PSNR) and Structural similarity (SSIM). Also, most of the image quality metrics are belonged to color images, but recently new metrics for depth images are emerged. This paper proposes a new Full-Reference (FR) image quality assessment metric for color and depth images, which works based on edge and pixel features. Proposed method is a combination of improved Edge Based Image Quality Assessments (EBIQA) and Peak Signal-to Noise Ratio (PSNR) methods. Proposed method is called Edge and Pixel-based Image Quality Assessment Metric (EPIQA). The system is validated using famous and benchmark performance metrics or quality measures such as Spearman Rank-Order Correlation Coefficient (SROCC) and Kendall Rank-Order Correlation Coefficient (KROCC), along with comparison with other similar methods on well-known related databases. Color databases have proper and diverse number of noises, but there is no proper depth noisy database, which it is decided to make one. Proposed method returned promising and satisfactory results in different tests.**

*Keywords- Image Quality Assessment (IQA) metric; Full-Reference (FR) metric; Edge and Pixel-based Image Quality Assessment (EPIQA); Depth Image Quality Metric*

## I. INTRODUCTION

Fast technology progress, especially in computer vision field, increased the need for measuring techniques of images quality. The quality is the root characteristic of image, which assesses the amount of degradation and distortion such as blurriness, colour shift and all types of noises in a digital image. With having these values, it is possible to compare it with a perfect or reference image. It is applicable especially when for instance, a sensor wants to normalize the receiving image quality to an ideal version for human eyes. Furthermore, using an Image Quality Assessment (IQA) metric, helps to control and minimize the amount of receiving or transmitting distortions on images. The paper presents a new IQA metric to cover almost all-important aspects of a digital image and for different purposes. IQA methods play crucial roles in different applications of digital image processing like system performance validation, image improvement and restoration, noise removal, image compressions algorithms performance comparison, restoration, packet loss, and digital watermarking [1, 2, 3, 4, 5]. The paper is consisted of V sections, which are as follow. Section I, demonstrate fundamentals of research. Section II, pays to prior related researches on the subject. Section III, explains the proposed IQA metric in details. Section IV covers all experiments, comparisons and validation results. Also, section V includes, conclusion, dissuasion and future works.

### A. Importance of the research

As it mentioned earlier, with increasing technological power of software and hardware in recent devices, it is needed to make more precise algorithms to deal with them. Digital image is not an exception in this area. Assessing the quality of digital image is important in three main fields. Initially, it can be employed for image quality monitoring for systems in quality control. For instance, a video acquisition system could get help from IQA metrics and adjust itself with the best receiving image possible. Second, IQA metric could be used in benchmark and famous systems, that are employed in image processing usages. Finally, this system can be mixed into an image processing machine to improve the algorithms' structure and the parameters.

A lot of traditional IQA metrics such as Mean squared error (MSE), Peak signal-to noise ratio (PSNR) and Structural similarity (SSIM) are completely use visual data inside an image matrix to calculate its quality. This information and data are too much as long as the Human Visual System (HVS) gets an image-based type data on its low-level information like the edges, zero-crossing data, lines and corners [6, 7]. So, changes in low-level image features are image degradations.

As it mentioned, most of the existing IQA metrics are made for specific condition, but this paper proposed a modified and combined version of an IQA metric which could overcome different image degradations. Also, proposed method could also be employed for depth or range images.

### *B. Human Visual System*

As it is clear, Human Visual System (HVS) is an appropriate tool to assess the quality of an image, but it is not possible to be integrated with an electronic device and includes a lot of error due to changing from one person to another. Moreover, each person has different opinion and taste about visual quality, which makes it harder have a standard assessment based on human opinion. So, it is needed to have an electronic standard system based on HVS structure and even more [6, 7].

### *C. Image Quality Assessments (IQA) factor or metric*

Any digital images could be polluted by a wide variety of noises and distortions in data transmission, data receiving, data computation, data compression, saving and this may result in quality prolusion and distortion. Image quality can refer to the level of accuracy that all imaging systems and sensors capture, process, store, compress, transmit and display [8]. As it mentioned, HVS is a fine tool for IQA, but it is not still a device to use. Automatic IQA metrics are made for these purposes. IQA metrics fall into two main categories. Subjective and objective IQA metrics.

#### *C.1 Subjective*

Image quality validation based on HVS is called subjective IQA. Some of the current subjective tests for IQA are done based on to Rec. ITUR BT.500-10 [9]. Plus, it is the best way to image quality assessment. The Mean Opinion Score or MOS is acquired by set of standard's results averaging. Actually, human eye observers rate the quality in terms of how bad the human thinks that the distortions are. Difference Mean Opinion Score (DMOS) is exactly as its name represents. Even if MOS can recognize Image Quality (IQ) better, it is time consuming process in real world usage [10, 11, 12]. Main point sizes of subjective validations are illustrated in Table I.

TABLE I. SUBJECTIVE ASSESSMENT OF IMAGES MAIN POINTS

| **MOS** | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| **Quality** | Bad | Poor | Fair | Nice | Perfect |
| **Deficiency** | Very Bad | Bad | Slightly Bad | Sensible | Imperceptible |

#### *C.2 Objective*

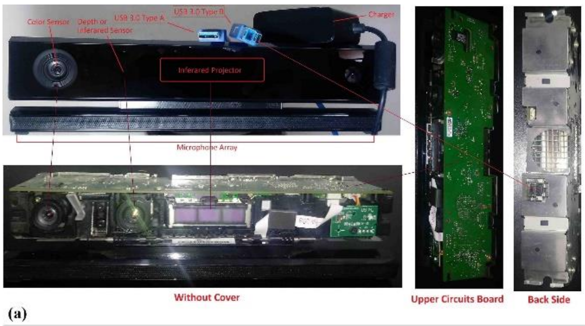
Makin an automatic IQA system which works as HVS is called objective IQA system. Many researches have been done during years to make a universal IQA system, but it was not so successful. These types of systems fall into three main categories. Full-Reference (FR), Non- Reference (NR) and Reduced-Reference (RR) [13]. This method is the most, accurate and famous one, which needs both reference and distorted images. Proposed method is belonged to this category [14]. It is called blind IQA. In this approach reference image is not exist. For instance, and in photography, mostly a NR algorithm is used to apprise the final user which a low-quality or high-quality image has been achieved. HVS could determine the quality in this condition, but it is so difficult to make an automatic system with same characteristics as human eyes [15]. This method is between FR and NR and just parts of the image is available. There are advantages like fewer data transmission, higher reliability, and wide usage range [16].

#### *C.2.1 FR measures type*

Full- Referenced (FR) IQA method could be classified into five classes of Pixel difference-based, Correlation-based, Edge-based, Context -based and Human Visual System-based systems as follow [17, 18]. As its name represents, every calculation is based on pixel by pixel. The mean square error (MSE), signal-to-noise ratio (SNR) and peak signal-to-noise ratio (PSNR) are in this class. Correlation is for comparing the difference of images. In IQA, pixels correlation is used as a scale of the quality. Here, original and the distorted images edges are found, then a displacement scale of edge parts is used to find the IQ in the entire digital image. In lieu of measuring pixels in both original and distorted data, pixel neighborhoods are compared by finding the probability to employee it for assessing the IQ.

### *D. Depth sensors and images*

Depth sensors are made to calculate the distance between sensor and the object. Also, they can be employed to make 3-dimentional model of the object. With these abilities, they can be useful to increase the accuracy of the final recognition. Kinect is one of the most practical depth sensors to have. It is so much cheaper than other depth sensors and efficient. It can be used on Microsoft Xbox 360 (Kinect V.1) or Xbox one (Kinect V.2) consoles or be used as a developer device. Kinect 2.0 was released with Xbox One on November 22, 2013. Because of the lower price and high power to use, a lot of developers and researcher use it as a main depth device. It could record RGB and Depth video frames with 1920*1080 resolution for RGB images and 512*424 for Depth images on 30 fps. Moreover, it is also capable to of working between 0.8-5.0-meter ranges [19]. An RGB-D image is simply a combination of an RGB (color) image and its corresponding depth image. A depth image is an image channel in which each pixel relates to a distance between the image plane and the corresponding object in the RGB image. It is also termed as 2.5D or Range image [20, 46]. Fig 1 shows Kinect V.2 internal structure (a) and Kinect V.1 VS V.2 specifications (b). Fig 2 represents difference between color and depth images by a visual example (sample from proposed NDDB).




(a)

| | Kinect V.1 | Kinect V.2 |
|---|---|---|
| **Color** | 640*480 in 30 fps | 920*1080 in 30 fps |
| **Depth** | 320*240 in 30 fps | 512*424 in 30 fps |
| **Sensor** | Structured Light | Time of Flight |
| **Range** | 0.8 till 0.4 m | 0.5 till 4.5 m |
| **Angle of View** | 57/43 degree | 70/60 degree |
| **Microphone Array** | yes | yes |

(b)

Figure 1. Kinect V.2 internal structure (a) and Kinect V.1 VS V.2 specifications (b)

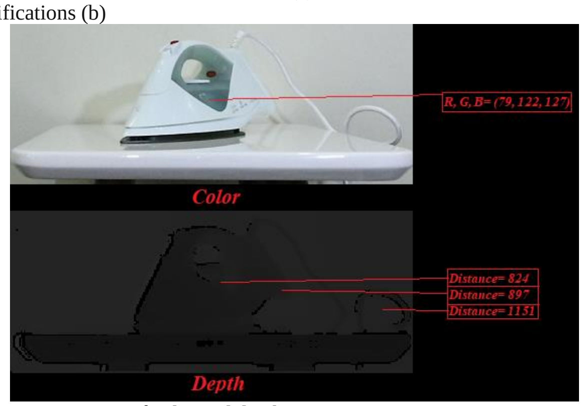



Figure 2. Structures of color and depth images

## II. Prior Related Researches

In this section, some of the most famous and new IQA metrics (color and depth) are explained [47]. This metrics are mostly pixel and edge based.

### *A. PSNR*

PSNR represents the unity level of signals. The Peak Signal to Noise Ratio or (PSNR) (1) matrix calculates the peak signal to noise ratio, in decibels unit, for images. The value is sometimes used as a IQ measurment for the original (input) and compressed (target) digital images. As higher the PSNR value, the better the compressing quality [12, 21].

$$\mathrm{PSNR} = 10\log_{10}\frac{L^2}{\mathrm{MSE}} \tag{1}$$

In that L is the range of value of pixel. It has unit of Dbwith limitation of 50. A good value is among 20 till 50.

### *B. MSE*

The Mean Square Error or MSE (2) and PSNR are the two best error metrics in order to measure image compression quality. MSE shows the cumulative squared error for the compressed and the original image, which PSNR represents a measure for peak error. The lower value of MSE is better [21].

$$\mathrm{MSE} = \frac{1}{M \times N}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}[X(i,j) - Y(i,j)]^2 \tag{2}$$

That, X and Y are arrays with size of M*N.

### *C. Structural Similarity (SSIM)*

The calculation of SSIM is well documented in the references [22, 23], so here, paper gives only a brief review in order to be able to present the proposal performed and evaluated in this paper. Structural similarity can be achived by comparing pixel intensities' local template which is normalized for luminance and contrast factors. If f is the original image and g the distorted one, the SSIM (f, g) is calculated through the (3) equation.

$$\mathrm{SSIM} = \left(\frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}\right)^{\alpha}\left(\frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_y^2 + C_2}\right)^{\beta}\left(\frac{\sigma_{fg} + C_3}{\sigma_f\sigma_g + C_3}\right)^{\gamma} \tag{3}$$

Being $\mu_i$ the mean intensity of the i image, $\sigma_i$ their standard deviation and $\sigma_{fg}$ is the covariance between the images.

### *D. Gradient Conduction Mean Square Error (GCMSE)*

GCMSE is an edge aware metric based on MSE. Here, weighted sum of gradients (distance pixels) is considered [24]. GCMSE always brings better performance and result than MSE and SSIM. The steps are:

- Gradient directions are estimated in four aims using (4), and then value of $G_p$ is found. Finally, The results are optimized by the k coefficient:

$$G = \frac{(I_2 - I_1)^2}{(I_2 - I_1)^2 + k^2} \tag{4}$$

The GCMSE is estimated based on (5):

$$\mathrm{GCMSE} = \frac{\sum_{x=1}^{m}\sum_{y=1}^{n}[(I_2(x,y) - I_1(x,y))G_p]^2}{C1 + \sum_{x=1}^{m}\sum_{y=1}^{n} G_p} \tag{5}$$

### *E. EBIQA*

Actually, edge preservation is one of the most important aspects during the human visual assessment. Edge Based Image Quality Assessment (EBIQA) method is aim to act on the human understanding of the reciving features [25]. Steps are:

- Edge's locations are detect using Sobel's edge detection method in both images (original and destorted).

- A pixel window in size of 16 × 16 vectors are formed at each image based on (6) and (7), where $I_1$ is the reference image and $I_2$ is the test image.

$$I_1 = (O, AL, PL, N, VHO) \tag{6}$$
$$I_2 = (O, AL, PL, N, VHO) \tag{7}$$

In that 'O' presents orientation of edges in the digital image, that actually is the total number of edges. 'AL' is the length average of whole edges. 'PL' computes the total number of pixels which have similar intensity level. 'N' presents sum of all pixels, which shape all edges. 'VH' shows edges in either vertically or horizontally directions as sum of pixels of them. Finally, we estimate EBIQA by (8), which mean Euclidean distance of vectors is estimated. This paper going to improve this method and combine it with PSNR.

$$\mathrm{EBIQA} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{(I_1 - I_2)^2} \tag{8}$$

### F. NSER

NSER is a zero-crossings method based [26]. Steps are as follow:

- Using gaussian kernel in the images on different Standard Deviation (SD) gauge to detect all edges.
- Then procedure is done for both digital images. Common edge ratio number placed by first edge number is found by (9):

$$p_i = ||I_1 \cap I_2|| / ||I_1|| \tag{9}$$

- The result is normalized by log to optimize correlation (10):

$$NSER(I_1, I_2) = -\sum_{i=1}^{N} \log 10(1 - p_i) \tag{10}$$

### G. Measure of Enhancement or Enhancement Measure (EME)

This metric is based on the concepts of the Webers Low of the HVS. It helps to select the best parameters [27].

$$EME = \max_{\Phi \in \{\Phi\}} \chi(EME(\Phi)), \text{which} \max_{\Phi \in \{\Phi\}}$$

$$= \chi\left(\frac{1}{k_1 k_2}\sum_{l=1}^{k_2}\sum_{k=1}^{k_1} 20 \log \frac{I^{\omega}_{max;k;l}}{I^{\omega}_{min;k;l}}\right) \tag{11}$$

Let an image x(n, m) be splitted into $k_1k_2$ blocks $w_{k,l}$ (i, j) of size $l_1 \times l_2$ and { Φ} be a provided class of orthogonal alter used for image improvement. Also $I^{\omega}_{max;k;l}$ and $I^{\omega}_{min;k;l}$ are respectively min and max of x (m, m) matrix inside the main block $w_{k,l.}$ The function χ is the sign one.

### H. Full reference Feature Similarity Measure (FSIM)

Full reference feature similarity or (FSIM) index is to asses IQ [28]. FSIM can achive vital aspects of IQ by using two low-level characterstics. Those are phase and gradient magnitudes.

Le, Thanh-Ha, Seung-Won Jung, and Chee Sun Won proposed a new depth image quality metric which demands only a single pair of color and depth images in 2017. Their method closely estimates the depth quality metrics that use the ground truth depth or stereo color image pair [29].

Also, in 2013, Tsai, Chang-Ting, and Hsueh-Ming Hang, proposed a novel 3D IQ metric to assess the quality of stereo images that may contain artifacts introduced by the rendering process due to depth map errors [30].

Another valuable work is Tian, Shishun, et al research in 2018. They proposed a full-reference metric to assess the quality of Depth-Image-Based-Rendering (DIBR) synthesized views, as they believed 2D quality metrics may fail to evaluate the quality of the synthesized views [31]. Fig 3 represents the workflow of objective IQA metrics which proposed method is full reference version of it.

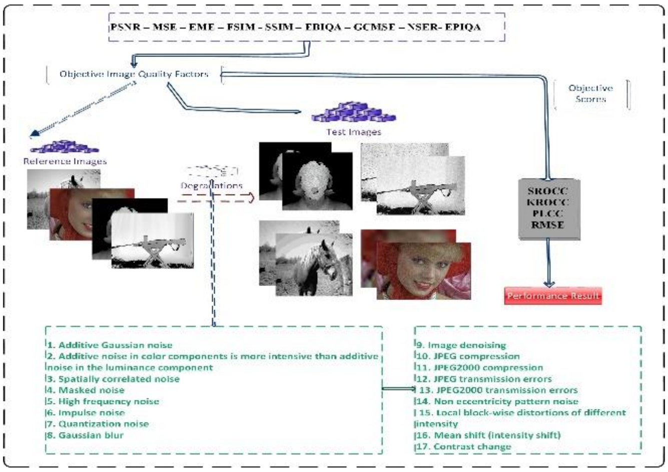


Figure 3. The workflow of objective Image Quality Assessment metrics

## III. PROPOSED IQA FACTOR OR METRIC

As edges play important roles in detecting and distinguishing objects in HVS, and human eye has this ability or IQA power, it is decided to make an automatic IQA metric based on edge and pixel features, which works similar to human eyes. Also, for covering all details, a pixel-based approach is combined with the proposed approach. In other hands, using such technique on depth images is almost unique. Proposed metric is called Edge and Pixel-based Image Quality Assessment metric (EPIQA).

In order to do that, after image acquisition process from input, pre-processing stage starts. Pre-processing is consisting of low pass and high pass filtering. First, median filter (12) applies on both reference and distorted images, then un-sharp masking (13) effect takes place on them. This makes image, smooth from inside and sharp from outside, which help to decrease most type of noises effects. In this stage and at the same time, PSNR (1) value between distorted and reference images will be calculated. As it is obvious, PSNR value is in the range of 20-50, which for finale combination, this value normalizes between ranges of 0-1. Based on repeated experiments, using a custom filter for edge detection performs better than traditional edge detection algorithms in this research. So, edge detection stage, will be done using proposed filter presented in Table II to extract the edges. Next stage is belonging to dividing both distorted and reference images to 8*8 blocks. Using larger blocks, loses some of the information and smaller blocks make the system so slow. So, using 8*8 blocks is so rational. For each block in both images, two vectors of (14) and (15) are defined.

TABLE II. PROPOSED EDGE DETECTION FILTER [32]

| | | | | | | |
|---|---|---|---|---|---|---|
| -1.2 | -0.8 | -0.6 | | 1.2 | 0.8 | 0.6 |
| 0 | 0 | 0 | | 0 | 0 | 0 |
| 1.2 | 0.8 | 0.6 | | -1.2 | -0.8 | -0.6 |
| | | | | | | |
| 1.2 | 0 | -1.2 | | -1.2 | 0 | 1.2 |
| 0.8 | 0 | -0.8 | | -0.8 | 0 | 0.8 |
| 0.6 | 0 | -0.6 | | -0.6 | 0 | 0.6 |

- *Median filter*

Low-pass Median filter (12) is a nonlinear filtering technique, normally employed for noise removal in an image or a signal. The fundamental idea is to convolve on the image block by block, replacing each block with the median of neighboring entries [33].

$$y[m,n] = \mathrm{median}\{x[i,j], (i,j) \in w \tag{12}$$

In which w shows a neighborhood, centered on location [m, n] in image.

- *Un-sharp masking*

A perfect sharpening technique which effects like a high-pass filter on digital image. Here opposite of un-sharp or not un-sharp filter which is masked (6) applies on image which means sharpening of edges. Parameters are amount, radius and threshold [34].

Un-sharp masking produce and image g(x, y) from an image f(x, y) via:

$$g(x,y) = f(x,y) - f_{smooth}(x,y) \tag{13}$$

Where $f_{smooth}(x,y)$ is smoother version of $f(x,y)$

$O_{i,j}$=(ED, ELA, GLR, NEP, EO (14)

$D_{i,j}$=(ED, ELA, GLR, NEP, EO) (15)

Which O and D are original and distorted images in i and j positions. Edge features are Edge Density (ED), Edge Length Average (ELA), Gray Level Region (GLR), Number of Edge Pixels (NEP) and Edge Orientation (EO).

Edge Density (ED): The number of edges in each 8*8 block of an image.

Edge Length Average (ELA): Average Length of each block. First, all edges' lengths are calculated, and a simple average makes it as a decimal number.

Gray Level Region (GLR): Number of regions with the same gray level in each block.

Number of Edge Pixels (NEP): Number of pixels for each edge in each block.

Edge Orientation (EO): Number of edges with vertical or horizontal orientation in each block.

For example, for $O_{i,j}$ and in Fig 4 (a) there are (6, 2, 3, 7, 5) values. Fig 4 (b) shows a sample from CSIQ database [35] in the edge detected form. In this Fig, right eye is shown in an 8*8 block. Such a block is used in each process of this paper.

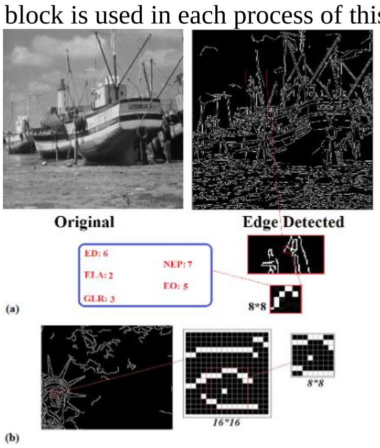



Figure 4. (a) US 1 Dollar (edge detection, final 8*8 sub block and extracting features), (b) a sample from CSIQ database (35) in the edge detected form. A sample of 8*8 block of reference image.

Now it is time to compute Euclidean distance $d_{i,j}$ between two corresponding blocks of original and distorted images, according to (16).

$$di,j = \left[\left(ED_{O_{i,j}} - ED_{D_{i,j}}\right)2 + \left(ELAO_{i,j} - ELAD_{i,j}\right)2 + \left(GLRO_{i,j} - GLRD_{i,j}\right)2 + \left(NEPO_{i,j} - NEPD_{i,j}\right)2 + \left(EOO_{i,j} - EOD_{i,j}\right)2\right]^{\left(\frac{1}{2}\right)} \tag{16}$$

And average distance is computes by final EPIQA (17) in range of 0-1:

$$EPIQA{:} = 1 - \left(\frac{1}{M \times N \times MAX(d_{i,j})} \sum_{i=1}^{M} \sum_{j=1}^{N} d_{i,j}\right) \tag{17}$$

Final step is computing the average for EPIQA and PSNR values acquired from both images. Final value is a decimal number in range of 0 and 1. As it is clear, as the number is closer to 1, the quality is higher and vice versa. Fig 5 represents proposed metric's flowchart. Also, Fig 6 shows the proposed metric's steps in visual form using one of the samples from the proposed Noisy Depth database (NDDB) database. Proposed NDDB database is explained in the next section.

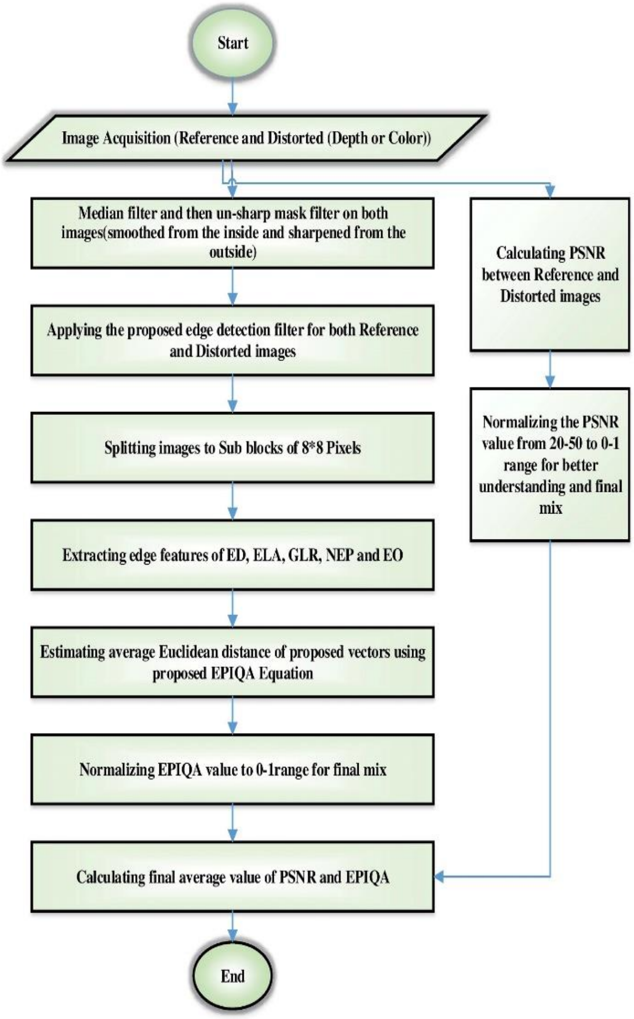



Figure 5. Proposed metric's flowchart

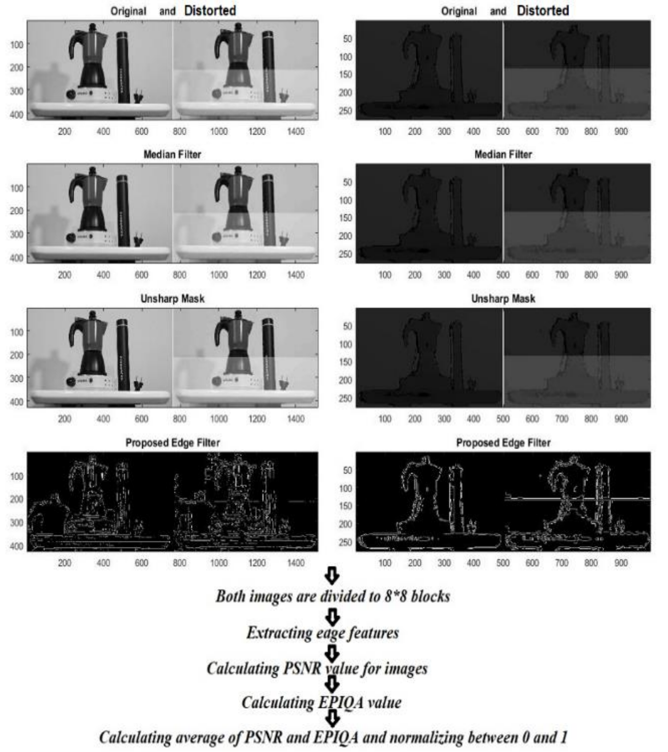



Figure 6. Proposed metric's steps in visual form (sample from proposed NDDB)

## IV. VALIDATING EXPERIMENTAL RESULTS

System is evaluated using four databases. Two color and two depth databases, which one of them is proposed in this paper, that explained in follow. For depth databases, just depth data is used to show the efficiency of the system just on depth data. Furthermore, four validation metrics are used for final measurement, which are explained in follow. For this purpose, a system with following specifications is used: Intel(R) Core (TM) i7-4790K CPU @ 4.00GHz, 16GB RAM, NVIDIA GeForce GTX 1050, 233GB SSD.

### *A. Databases*

There are different color-based databases for IQA, which we use two of them in this research. Databases such as TID2008 [36], CSIQ [35], LIVE [37], IVC [38], MICT [39], A57 [1] are available, that A57 and TID2008 are used in this research. Also, for depth data sample from Eurecom Kinect [40] are used. In addition, a database which is consists of 30 depth images is made using Kinect version 2 sensor and proposed in this research. Proposed database is called Noisy Depth database (NDDB) and includes small home objects.

For A57 database, pollutions are: FLT, JPG, JP2, DCQ, BLR and NOZ. For TID 2008 database, distortion types are 17 in 4 levels. Which makes 68 distorted images based on [36]. For depth images from depth databases, noises are added manually using seven noises of Gaussian [41, 42], Salt and pepper [41, 42], Poisson [43], Speckle [44], quantization [48], additive white gaussian [49] and block-wised [36] noises in just one level. These noises are added to Eurecom and proposed NDDB databases. Table III, presents the specifications of used databases. Additionally, some test samples from these databases are shown in Fig 7. Fig 8 shows recording environment for proposed NDDB database.

TABLE III. USED DATABASES' SPECIFICATIONS

| | A57 | TID2008 | Eurecom | NDDB |
|---|---|---|---|---|
| **Source images** | 3 | 25 | 52 | 30 |
| **Distorted images** | 54 | 1700 | 364 | 210 |
| **Distortion types** | 6 | 68 (17*4 levels) | 7 | 7 |
| **Image type** | Gray | Color | Depth | Depth |
| **Size** | 512*512 | 512*384 | 256*256 | 512*424 |
| **Cite** | [1] | [36] | [40] | - |

Figure 7. Samples from used databases


Figure 8. Recording environment for NDDB database

### *B. Noises*

Gaussian, salt and pepper, poison and speckle noises are explained in the paper and quantization, additive white gaussian and block-wised are referred to [48, 49] and [36] respectively and in order to save more space.

Gaussian noise is statistical noise with Probability Density Function (PDF) of the normal distribution. The PDF P of a Gaussian variable Z is given by (18):

$$P_g(Z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(Z-\mu)^2}{2\sigma^2}} \tag{18}$$

In which Z shows the grey level, μ the mean value and σ represents standard deviation value.

Gaussian noise happens during data acquisition sensor noise caused by weak illumination and high environment temperature. Gaussian noise can be reduced using a spatial smoothing filter on the image [41, 42].

Salt-and-pepper or impulse noise (19) is a type of noise which rarely seen on digital images. This one can be caused by sharp and fast transmission in the image which shows itself as white and black pixels. Median filtering could fix this noise [41, 42].

$$P(Z) = \begin{cases} P_a & \text{for } z = a \\ P_b & \text{for } z = b \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

Poisson noise (sometimes calls poisson noise) (20) is a type of noise which happens during electrical malfunctions, that can be defined by a Poisson formulation. In electronics shot noise originates from the discrete nature of electric charge [43].

$$\Pr(N = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} \tag{20}$$

Where $\lambda$ is the number of photons per unit time interval.

Finally, Speckle noise (21) in conventional radar outcome from random sway in the return data of an object that is no bigger than a data image-processing element.

$$g(m,n) = f(m,n) * u(m,n) * \eta(m,n) \tag{21}$$

In which g (m, n) is distorted image, u (m, n) is multiplicative component and $\eta$ (m, n) is extra component [44]. Effect of different noises on different images from used databases with different parameters is represented in Fig 9.

*C. Validation metrics*

For evaluating proposed EPIQA metric, four commonly performance metrics had been used. Used metrics are SROCC [25, 45], KROCC [25, 45], PLCC [25, 45] and RMSE [25, 45].

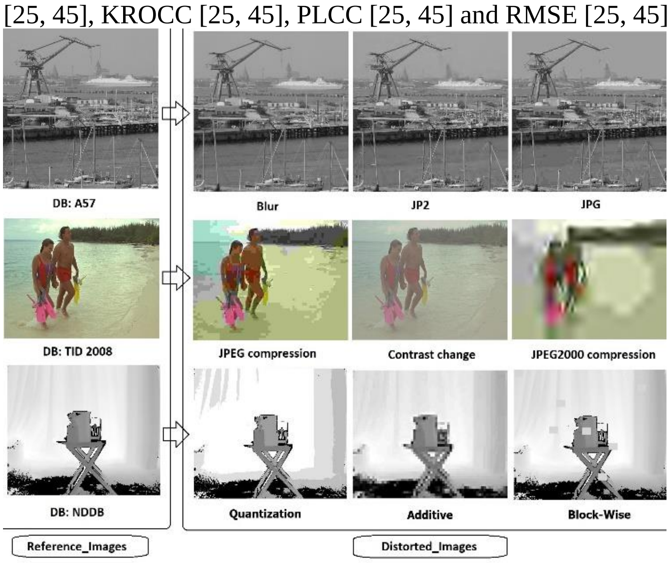



Figure 9. Effect of different noises on different images from used databases

*C.1 Spearman rank-order correlation coefficient (SROCC)*

SROCC shows the statistical association among the rankings of two variables. It evaluates how well the relationship between variables can be defined using a following function [25, 45].

$$\text{SROCC} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{22}$$

Where $d_j$ is the difference between the $i^{th}$ image's ranks in the subjective and objective evaluations.

*C.2 Kendall rank-order correlation coefficient (KROCC)*

This measure used to assess the serial dependence between two assessed quantities. Intuitively, the Kendall correlation between two variables will be high when observations have a homogeneous rank [25, 45].

$$\text{KROCC} = \frac{n_c - n_d}{0.5n(n-1)} \tag{23}$$

In which $n_c$ and $n_d$ are the number of compliant pairs and dissonant pairs in the data set.

*C.4 Pearson Linear Correlation Coefficient (PLCC)*

This one is the variables covariance which is divided by the product of their Standard Deviations (SD). In the following, it is supposed that $s_j$ is the subjective score ofthe $i^{th}$ image and $x_j$ is the objective score ofthe $i^{th}$ image. For the nonlinear regression analysis, first map $x_i$ to $q_i$ by the mapping function in (24) [25, 45].

$$q(x) = \beta_1\left(\frac{1}{2} - \frac{1}{1+e^{\beta_2(x-\beta_3)}}\right) + \beta_4 x \tag{24}$$

Where $\beta_j$ are parameters which have to be fitted. Third one is PLCC between MOS and the objective scores after nonlinear regression.

$$\text{PLCC} = \frac{\sum_{i=1}^{n}(s_{i-\bar{s}})(q_{i-\bar{q}})}{\sqrt{\sum_{i=1}^{n}(s_{i-\bar{s}})^2}\sqrt{\sum_{i--1}^{n}(q_{i-\bar{q}})^2}} \tag{25}$$

*C.5 Root-Mean-Square Error (RMSE)*

Final metric is RMSE between MOS and the objective scores after nonlinear regression, which can be explained as (26) [25, 45].

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(s_i - q_i)^2} \tag{26}$$

*D. Experimental results*

System is validated using four common metrics of SROCC, KROCC, PLCC and RMSE on four databases of A57, TID2008, Eurecom and proposed NDDB, and compared with some of the traditional and new IQA metrics. Table IV to VII shows acquired results on each database. As it is clear in all four databases, proposed EPIQA method could achieve better results in most cases. Also, the power of metrics could be categories as follow (from worst to best): SSIM, PSNR, [29], NSER, [30], FSIM, SC-IQA, EME, GCMSE, EBIQA and EPIQA. The neural linear regression with 70 hidden layers between proposed EPIQA and other methods and MOS is presented in Fig 10. These results are achieved for all databases and performance quality measures. Also, the results of performance metrics on all databases are represented and compared in Figs 11, 12, 13 and 14.

TABLE IV. COMPARISON RESULTS ON A57 DATABASE

| | PSNR | SSIM | EBIQA | GCMSE | MSE | FSIM | EME | [29] | [30] | SC-IQA | NSER | EPIQA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SROCC** | 0.618 | 0.806 | 0.860 | 0.851 | 0.651 | 0.829 | 0.846 | 0.790 | 0.665 | 0.837 | 0.823 | **0.874** |
| **KROCC** | 0.530 | 0.605 | **0.691** | 0.620 | 0.563 | 0.628 | 0.632 | 0.636 | 0.577 | 0.673 | 0.609 | 0.666 |
| **PLCC** | 0.707 | 0.801 | 0.878 | 0.867 | 0.740 | 0.814 | 0.833 | 0.857 | 0.754 | 0.844 | 0.810 | **0.901** |
| **RMSE** | 0.173 | 0.146 | 0.119 | 0.135 | 0.145 | 0.139 | 0.133 | 0.137 | 0.149 | 0.104 | 0.120 | **0.102** |

TABLE V. COMPARISON RESULTS ON TID2008 DATABASE

| | PSNR | SSIM | EBIQA | GCMSE | MSE | FSIM | EME | [29] | [30] | SC-IQA | NSER | EPIQA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SROCC** | 0.712 | 0.819 | 0.890 | 0.911 | 0.745 | 0.842 | 0.876 | 0.890 | 0.759 | 0.904 | 0.883 | **0.916** |
| **KROCC** | 0.690 | 0.688 | 0.763 | 0.727 | 0.703 | 0.711 | 0.728 | 0.742 | 0.717 | 0.770 | 0.725 | **0.777** |
| **PLCC** | 0.733 | 0.764 | 0.798 | 0.712 | 0.766 | 0.737 | 0.723 | 0.737 | 0.780 | 0.812 | 0.700 | **0.823** |
| **RMSE** | 0.127 | 0.126 | 0.114 | 0.120 | 0.125 | 0.129 | 0.138 | 0.112 | 0.139 | **0.100** | 0.125 | 0.106 |

TABLE VI. COMPARISON RESULTS ON EURECOM DATABASE

| | PSNR | SSIM | EBIQA | GCMSE | MSE | FSIM | EME | [29] | [30] | SC-IQA | NSER | EPIQA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SROCC** | 0.699 | 0.709 | **0.755** | 0.720 | 0.712 | 0.712 | 0.730 | 0.734 | 0.716 | 0.749 | 0.727 | 0.742 |
| **KROCC** | 0.710 | 0.694 | 0.770 | 0.766 | 0.723 | 0.707 | 0.754 | 0.758 | 0.727 | 0.764 | 0.731 | **0.775** |
| **PLCC** | 0.728 | 0.717 | 0.749 | 0.727 | 0.711 | 0.730 | 0.743 | 0.737 | 0.725 | 0.763 | 0.730 | **0.781** |
| **RMSE** | 0.118 | 0.112 | 0.110 | 0.114 | 0.111 | 0.115 | 0.127 | 0.121 | 0.125 | 0.114 | 0.114 | **0.106** |

TABLE VII. COMPARISON RESULTS ON PROPOSED NDDB DATABASE

| | PSNR | SSIM | EBIQA | GCMSE | MSE | FSIM | EME | [29] | [30] | SC-IQA | NSER | EPIQA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **SROCC** | 0.751 | 0.740 | 0.766 | **0.798** | 0.744 | 0.743 | 0.759 | 0.763 | 0.758 | 0.780 | 0.756 | 0.783 |
| **KROCC** | 0.697 | 0.705 | 0.732 | 0.729 | 0.710 | 0.728 | 0.727 | 0.731 | 0.714 | 0.746 | 0.724 | **0.750** |
| **PLCC** | 0.730 | 0.704 | 0.735 | 0.731 | 0.723 | 0.727 | 0.737 | 0.738 | 0.747 | 0.740 | 0.734 | **0.748** |
| **RMSE** | 0.111 | 0.124 | 0.109 | 0.117 | 0.124 | 0.127 | 0.122 | 0.136 | 0.148 | 0.113 | 0.129 | **0.101** |

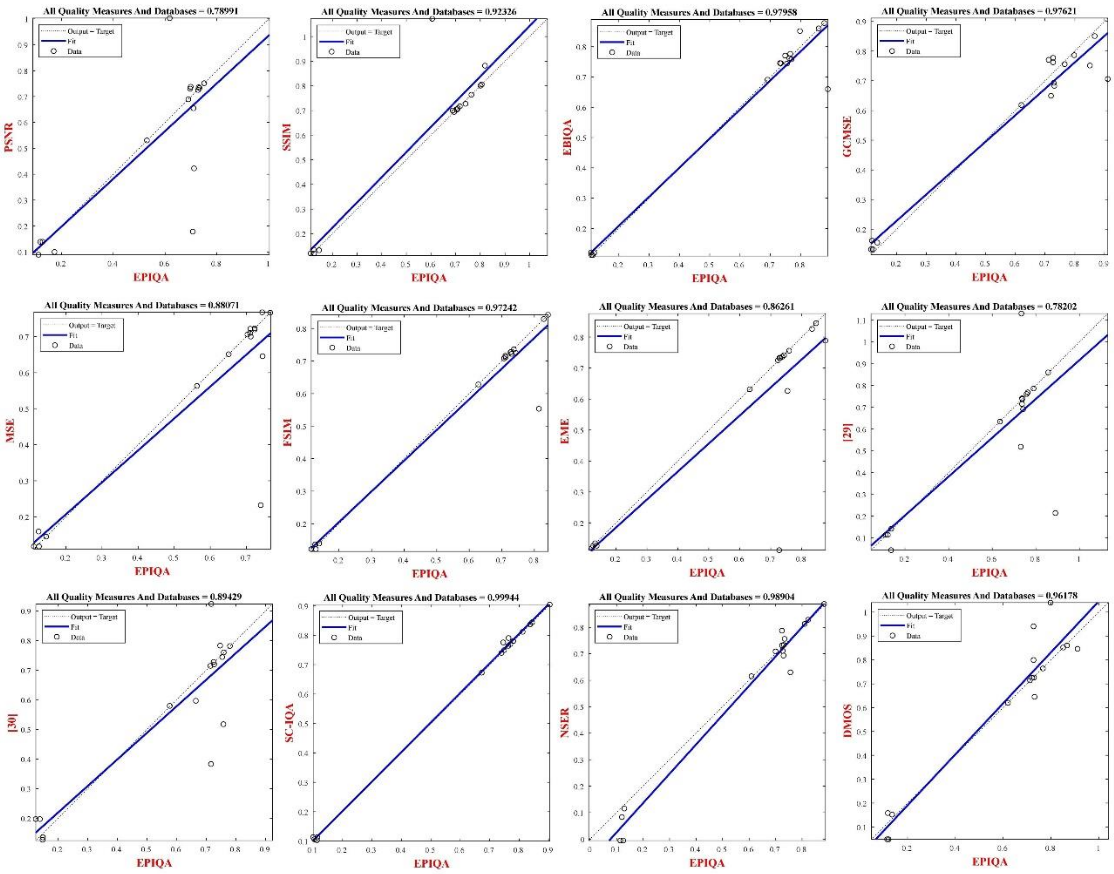



Figure 10. Neural linear regression for proposed method versus other methods (all databases and quality measures)

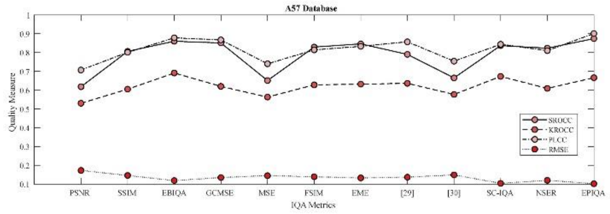



Figure 11. Performance result on A57 database

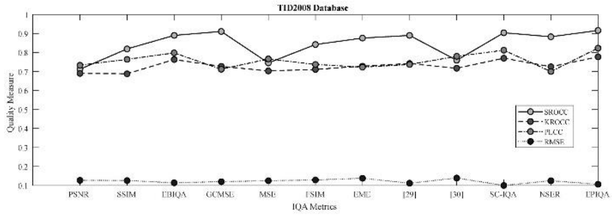



Figure 12. Performance result on TID2008 database

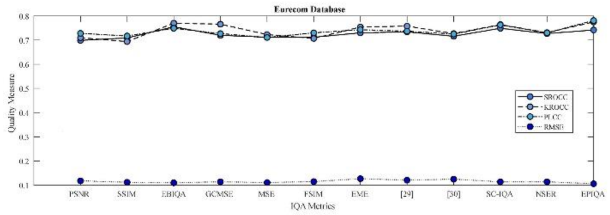



Figure 13. Performance result on Eurecom database

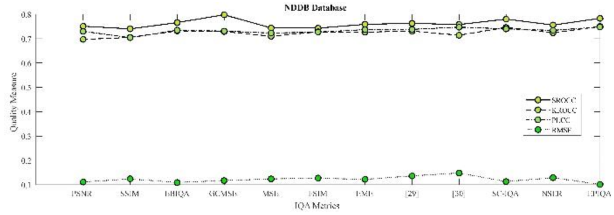



Figure 14. Performance result on NDDB database

## V. CONCLUSION AND FUTURE WORKS

The paper proposed an edge and pixel based IQA metric for color and depth images. Proposed method was tested using commonly used performance metrics on well-known color and depth databases. Also due to lack of distorted depth database, a new database of NDDB is made for this purpose. Also, at comparison stage, proposed EPIQA metric is compared with some traditional and new metrics, which in both color and depth experiments, satisfactory results were achieved. Methods such as EME, SC-IQA, EBIQA and GCMSE had good performance after proposed method. Future work is consisting of adding new structure such as context-based metrics to exist version; and also making system compatible with NR and RR systems. It is suggested to make a new color and depth distorted database with more distortion types.

## REFERENCES


[1] Chandler, Damon M., and Sheila S. Hemami. "VSNR: A wavelet-based visual signal-to-noise ratio for natural images." IEEE transactions on image processing 16.9 (2007): 2284-2298.

[2] Chandler, Damon M., Kenny H. Lim, and Sheila S. Hemami. "Effects of spatial correlations and global precedence on the visual fidelity of distorted images." Human Vision and Electronic Imaging XI. Vol. 6057. International Society for Optics and Photonics, 2006.

[3] Li, Xin. "Blind image quality assessment." IEEE International Conference on Image Proceedings. 2002. Vol. 1, 2002.

[4] Liu, Anmin, Weisi Lin, and Manish Narwaria. "Image quality assessment based on gradient similarity." IEEE Transactions on Image Processing 21.4 (2012): 1500-1512.

[5] Sheikh, Hamid R., Alan C. Bovik, and Gustavo De Veciana. "An information fidelity criterion for image quality assessment using natural scene statistics." IEEE Transactions on image processing 14.12 (2005): 2117-2128.

[6] Marr D (1982) "Vision: a computational investigation into the human rrepresentation and processing of visual information". W.H. Freeman and Company.

[7] Zhang, Min, et al. "Blind image quality assessment using the joint statistics of generalized local binary pattern." IEEE Signal Processing Letters 22.2 (2015): 207-210.

[8] Burningham, Norman, Zygmunt Pizlo, and Jan P. Allebach. "Image quality metrics." Encyclopedia of imaging science and technology (2002).

[9] BT, RECOMMENDATION ITU-R. "Methodology for the subjective assessment of the quality of television pictures." International Telecommunication Union (2002).

[10] Blasch, Erik, et al. "Image quality assessment for performance evaluation of image fusion." IEEE International Conference on Information Fusion, 2008.

[11] Navas, K. A., M. L. Aravind, and M. Sasikumar. "A novel quality measure for information hiding in images." 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2008.

[12] Wang, Zhou, et al. "Image quality assessment: from error visibility to structural similarity." IEEE transactions on image processing 13.4 (2004): 600-612.

[13] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment." The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003. Vol. 2. Ieee, 2003.

[14] Sheikh, Hamid R., Muhammad F. Sabir, and Alan C. Bovik. "A statistical evaluation of recent full reference image quality assessment algorithms." IEEE Transactions on image processing 15.11 (2006): 3440-3451.

[15] Wang, Zhengyou, et al. "No-reference hybrid video quality assessment based on partial least squares regression." Multimedia Tools and Applications 74.23 (2015): 10277-10290.

[16] Yu, Ming, et al. "A method for reduced-reference color image quality assessment." 2009 2nd International Congress on Image and Signal Processing. IEEE, 2009.

[17] Sheikh, Hamid R., Muhammad F. Sabir, and Alan C. Bovik. "A statistical evaluation of recent full reference image qu ality assessment algorithms." IEEE Transactions on image processing 15.11 (2006): 3440-3451.

[18] Al-Najjar, Yusra AY, and D. Chen Soong. "Comparison of image quality assessment: PSNR, HVS, SSIM, UIQI." International Journal of Scientific & Engineering Research 3.8 (2012): 1.

[19] Sarbolandi, Hamed, Damien Lefloch, and Andreas Kolb. "Kinect range sensing: Structured-light versus time-of-flight kinect." Computer Vision and Image Understanding 139 (2015): 1-20.

[20] Ganguly, Suranjan, Debotosh Bhattacharjee, and Mita Nasipuri. "3D Image Acquisition and Analysis of Range Face Images for Registration and Recognition." Handbook of Research on Emerging Perspectives in Intelligent Pattern Recognition, Analysis, and Image Processing (2015): 185.

[21] ECE, CSE, and M. M. U. Mullana. "Image quality assessment techniques pn spatial domain." IJCST 2.3 (2011).

[22] Wang, Zhou, and Alan C. Bovik. "Modern image quality assessment." Synthesis Lectures on Image, Video, and Multimedia Processing 2.1 (2006): 1-156.

[23] Silvestre-Blanes, Javier, and Rubén Pérez-Lloréns. "SSIM and their dynamic range for image quality assessment." ELMAR, 2011 Proceedings. IEEE, 2011.

[24] López-Randulfe, Javier, et al. "A quantitative method for selecting denoising filters, based on a new edge-sensitive metric." 2017 IEEE International Conference on Industrial Technology (ICIT). IEEE, 2017.

[25] Attar, Abdolrahman, Asadollah Shahbahrami, and Reza Moradi Rad. "Image quality assessment using edge based features." Multimedia Tools and Applications 75.12 (2016): 7407-7422.

[26] Zhang, Min, Xuanqin Mou, and Lei Zhang. "Non-shift edge based ratio (NSER): An image quality assessment metric based on early vision features." IEEE Signal Processing Letters 18.5 (2011): 315-318.

[27] Agaian, Sos S., Karen Panetta, and Artyom M. Grigoryan. "A new measure of image enhancement." IASTED International Conference on Signal Processing & Communication. Citeseer, 2000.

[28] Zhang, Lin, et al. "Training quality-aware filters for no-reference image quality assessment." IEEE MultiMedia 21.4 (2014): 67-75.

[29] Le, Thanh-Ha, Seung-Won Jung, and Chee Sun Won. "A new depth image quality metric using a pair of color and depth images." Multimedia Tools and Applications 76.9 (2017): 11285-11303.

[30] Tsai, Chang-Ting, and Hsueh-Ming Hang. "Quality assessment of 3D synthesized views with depth map distortion." Visual Communications and Image Processing (VCIP), 2013. IEEE, 2013.

[31] Tian, Shishun, et al. "SC-IQA: Shift compensation based image quality assessment for DIBR-synthesized views." IEEE International Conference on Visual Communications and Image Processing. 2018.

[32] Mousavi, Seyed Muhammad Hossein and Marwa Kharazi. "An Edge Detection System for Polluted Images by Gaussian, Salt and pepper, Poisson and Speckle Noises.", 4th National Conference on Information Technology,Computer & TeleCommunication, At Iran-Torbat-e Heydarieh, 2017.

[33] Huang, T., G. J. T. G. Y. Yang, and G. Tang. "A fast two-dimensional median filtering algorithm." IEEE Transactions on Acoustics, Speech, and Signal Processing 27.1 (1979): 13-18.

[34] Polesel, Andrea, Giovanni Ramponi, and V. John Mathews. "Image enhancement via adaptive unsharp masking." IEEE transactions on image processing 9.3 (2000): 505-510.

[35] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," Journal of Electronic Imaging, 19 (1), March 2010.

[36] N. Ponomarenko, V. Luki n, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for Evaluation of Full-Reference Visual Quality Assessment Metrics", Advances of Modern Radioelectronics, Vol. 10, pp. 30-45, 2009.

[37] Sheikh HR, Sabir MF, Bovik AC (2006) "A statistical evaluation of recent full reference image quality assessment algorithms". IEEE Trans IP 15:3440–3451

[38] Subjective quality assessment IRCCyN/IVC database, http://www2.irccyn.ec-nantes.fr/ivcdb.

[39] MICT Image Quality Evaluation Database, http://mict.eng.utoyama.ac.jp/mictdb.html.

[40] Rui Min, Neslihan Kose, Jean-Luc Dugelay, "KinectFaceDB: A Kinect Database for Face Recognition," IEEE Transactions on Systems, Man, and Cybernetics: Systems 44.11 (2014): 1534-1548.

[41] Gonzales, Rafael C., and Richard E. Woods. "Digital Image Processing", Addison & Wesley Publishing Company. Reading, MA (1992).

[42] Jain, Anil K. "Fundamentals of digital image processing". Englewood Cliffs, NJ: Prentice Hall, 1989.

[43] Hasinoff, Samuel W. "Photon, poisson noise." Computer Vision. Springer US, 2014. 608-610.

[44] Jaybhay, Jyoti, and Rajveer Shastri. "A study of speckle noise reduction Filters." signal & image processing: An international Journal (SIPIJ) 6.3 (2015): 71-80.

[45] Wang, Zhou, and Qiang Li. "Information content weighting for perceptual image quality assessment." IEEE Transactions on Image Processing 20.5 (2011): 1185-1198.

[46] Mousavi, Seyed Muhammad Hossein. "A New Way to Age Estimation for RGB-D Images, based on a New Face Detection and Extraction Method for Depth Images." International Journal of Image, Graphics and Signal Processing 10.11 (2018): 10.

[47] Mousavi, Seyed Muhammad Hossein, Vyacheslav Lyashenko, and Surya Prasath. "Analysis of a robust edge detection system in different color spaces using color and depth images." Компьютерная оптика 43.4 (2019).

[48] Robertson, Mark A., and Robert L. Stevenson. "DCT quantization noise in compressed images." IEEE Transactions on Circuits and Systems for Video Technology 15.1 (2005): 27-38.

[49] Damera-Venkata, Niranjan, et al. "Image quality assessment based on a degradation model." IEEE transactions on image processing 9.4 (2000): 636-650.