# Perceived image quality prediction based on image quality attributes extraction

Pinchas Zorea

**Abstract -** As the smartphones and tablet computers with embedded cameras and high definition resolution became an essential part in our life. A great deal of effort has been made by the smartphones vendors in recent years to develop an "objective" image quality metrics in order to predict the how the image quality is perceived by the consumers. Unfortunately, only limited success has been achieved. This paper describes the evaluation of a new "objective" image quality model that predicts the perceived image quality. The new model is based on the standard image quality attribute extraction to the VQEG image quality evaluation tool criteria. This improves and reduces the process and cost by providing a new quantitative method to evaluate perceived image quality of color images on smartphone displays. Four image quality factors: brightness, contrast, color saturation and sharpness, were chosen to represent perceived image quality. This new image quality assessment model is based on results of human visual tests that compared with image analysis by the software application VIQET. During the research, the VIQET tool was calibrated based on results from human visual experiments. This paper describes the evaluation of the new model according to the VQEG recommendations.

**Index Terms -** Perceived IQ (Image Quality), Human Visual Test (HVT), objective image quality assessment, subjective image quality, image quality attributes, VIQET (VQEG Image Quality Evaluation Tool), mean opinion score (MOS).

## 1 INTRODUCTION

This paper describes the evaluation of proposed new model which includes a framework and the VIQET for smartphones perceived image quality prediction. The model is composed of a HVTs (Human Visual Tests) procedure and an evaluation by the VIQEG. The VIQET is an objective, no-reference photo quality evaluation tool. The correlations between the metrical and perceptual results indicated that MOS, MSE, PSNR metrics give excellent prediction performance in most cases.

The statistical analyses were conducted to check whether the increase of the image quality attributes would lead to improvement in user's perceived image quality. The finding is useful for the smartphones industry to have a better understanding of the concrete benefit of enhancing the image quality attributes. The proposed quality assessment model is useful also for image quality assessment of any mobile or desktop displays.

One unique feature of the new model was the capability of incorporating existing full reference image quality metrics without modifying them. This research implemented the framework for smartphones displays and used the framework to evaluate the prediction performance of state-of-the-art image quality metrics regarding the most important image quality attributes for projection displays.

*Pinchas Zorea is currently pursuing P.hD. degree program in Department of Mathematics and Informatics, State University of Moldova, Chisinau. Email: pini.zorea@gmail.com*

The evaluated image quality attributes were brightness, contrast, color saturation and sharpness, however the proposed framework was not bound by the possibilities. All the metric evaluations were supported by the correlation of objective and subjective experimental results. The proposed image quality assessment framework was originally designed for smartphones displays, but could be easily adapted to other types of displays with limited modifications. In conclusion, with the results that obtained in this paper, the new approach provided by the new model can be a good process for perceived image quality prediction.

The new model flow chart in Figure 1, presents the method used during the new model development process which includes subjective IQ assessment via HVTs and objective IQ assessment with VIQET.

## 2 NEW MODEL FOR PERCEIVED IQ PREDICTION

The new model development process is illustrated in Figure 1. The flow chart describes the IQ subjective IQ assessment through HVTs and the objective IQ assessment using the VIQET.

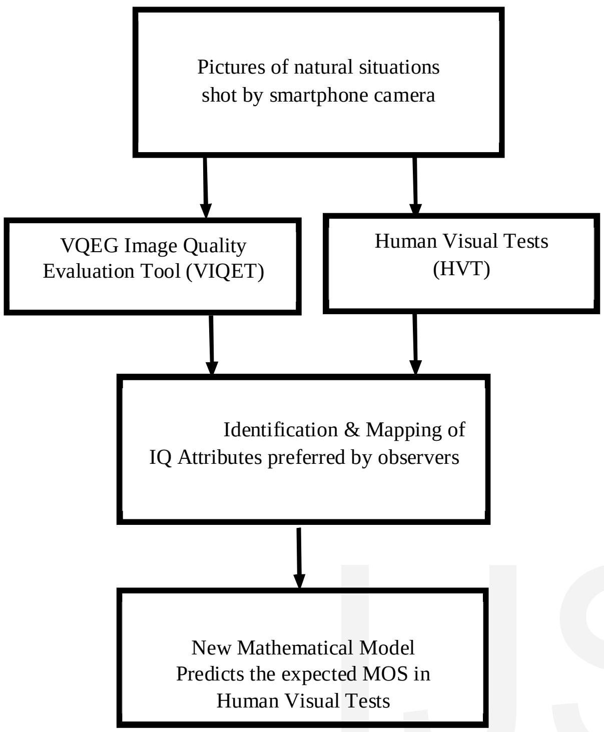



Fig. 1. Flow chart of the new model development process

The VQEG Image Quality Evaluation Tool (VIQET) is an objective, no-reference photo quality evaluation tool. VIQET is an open source tool designed to evaluate quality of consumer photos. In order to perform photo quality evaluation, VIQET requires a set of photos from the test device. It estimates an overall Mean Opinion Score (MOS) for a device based on the individual image MOS scores in the set.

- VIQET is an open source project that is available at www.GitHub.com/VIQET.
- The desktop tool installer can be downloaded at: https://github.com/VIQET/VIQET-Desktop/releases
- The source code can be found at: https://github.com/VIQET/VIQET-Desktop

In order to perform photo quality evaluation, VIQET requires a set of photos from the test device. It estimates an overall MOS of a device based on the individual image MOS scores in the set. The estimated MOS of each photo is based on a number of image quality features and statistics extracted from the test photo. The mapping from extracted features to MOS is based on psychophysics studies that were conducted to create a large dataset of photos and associated subjective MOS ratings.
The studies were used to learn a mapping from quantitative image features to MOS.

The estimated MOS by VIQET falls in a range of 1 to 5, where 1 corresponds to a low quality rating and 5 corresponds to excellent quality. Figure 2. demonstrates an example of VIQET RGB histogram and Figure 3. demonstrates VIQET Sharpness map.

Fig. 2. An example of VIQET RGB histogram

Fig. 3. An example of VIQET Sharpness map

Multi-scale edge acutance: refers to how sharp the outline of objects in an image are and how many edges were detected in the scene.
The sharper the image, the higher the multi-scale edge acutance feature.
Noise signature index:
refers to how noisy or grainy an image is. This feature value ranges from 0 to 589.
The higher the index, the grainier the image.
Saturation:refers to how vivid and intense a color is.
An image with poor color saturation will look washed out or faded.
When a color's saturation level is reduced to 0, it becomes a shade of gray.
Illumination: refers to how well lit an image is.
An image is considered well-lit if it is bright and has a sufficient amount of detail. Its values ranges from 0-255.
Dynamic Range: is the range between the lightest and darkest regions in an image while maintaining details of an image in both the lightest and darkest spots (represented in shades of grey).

### 2.1 Image quality analysis by VIQET

The VIQET is an objective, no reference photo quality Evaluation tool.

VIQET is a free and open source tool designed to evaluate quality of consumer photos. In order to perform photo quality evaluation, VIQET requires a set of photos from the test device.
It estimates an overall MOS for a device based on the individual image MOS scores in the set.
The estimated MOS for each photo is based on a number of image quality features and statistics extracted from the test photo. The mapping from extracted features to MOS is based on psychophysics studies that were conducted to Create a large dataset of photos and associated subjective MOS ratings. The studies were used to learn a mapping from quantitative image features to MOS.
The estimated MOS by VIQET falls in a range of 1 to 5, where 1 corresponds to a low quality rating and 5 corresponds to excellent quality.

The same images used for rating IQ by HVTs (Human Visual Test) were required for IQ rating by VIQET to analyze each individual image and get its IQ scores.
This paper describes the performance evaluation and validation of the proposed new model. Comparing the perceived image quality scores given by observers during the HVTs with predicted scores by VIQET as the outcomes of the new model, as demonstrated in Figure 4. , showing scatter plots of the perceived image quality versus the predicted image quality.



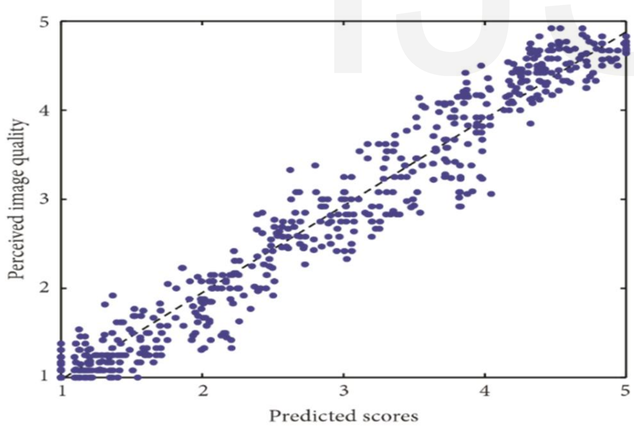



Fig. 4. Perceived image quality and predicted scores [1]

It seems that there is no literature aiming at further estimating the perceived image quality of smartphones where the image quality attributes have been evaluated. Hence, this study can only benchmark the performance of the proposed method with the image IQ attributes.

The VIQET correlation coefficients will be used to prediction the perceived smartphone image quality.

To evaluate the performance of the proposed quality assessment model, this study followed the standard performance evaluation procedures of VQEG [2].

The standard was developed for calculating the prediction error between a mathematical model and subjective scores (i.e., human viewers' opinion).

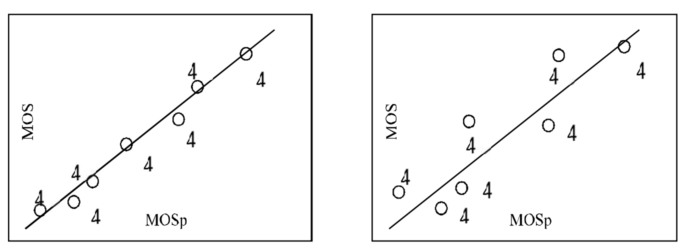
## 2.2 The new model performance evaluation procedures

According to the VQEG [2], the performance of an objective quality model is characterized by three prediction attributes:

Accuracy — is the ability to predict the distortions between MOS and MOSp. In an ideal case, the relationship between the MOS and MOSp is expected to be linear. Figure 5. illustrates the hypothetical relationships between the MOS and the MOSp of two models. Model-I is more accurate than the Model-II because most of the images evaluations are reasonably closer to the straight line.

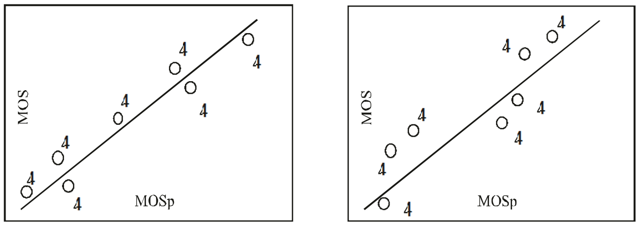
Monotonicity — is the degree to which the model's predictions agree with the relative magnitudes of subjective quality ratings. The prediction monotonicity is the extent of agreement between the subjective test and the objective model of variations in picture quality. As an example, viewers rank image A for many different levels of compressions where it implies the picture quality gets better when the level of compression is minimal.

A monotonic objective model should give the same result, but it does not follow the trend even though they are mathematically equivalent. Figure 3.3 illustrates the hypothetical relationships between the MOS and the MOSp of two models. Model-I has a better Pearson correlation than model-II, but it falsely predicts degradation in picture quality in two events when the assessors actually see an improvement in picture quality. Therefore, in terms of monotonicity, model-II is better than model-I.




Model-I (accurate) Model-II (not accurate)
Fig. 5. Hypothetical models with different prediction accuracy [3]

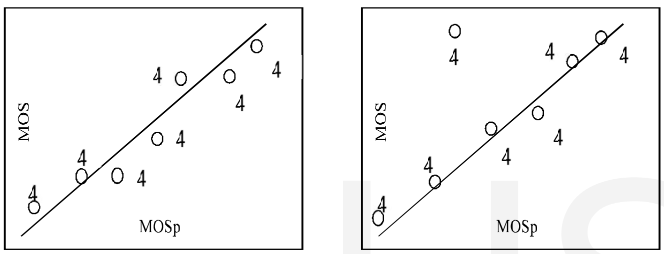
ISSN 2229-5518




Model-I (monotonic) Model-II(non monotonic)

Fig. 6. Different prediction monotonicity [3]

Consistency — is the degree to which the model maintains prediction accuracy over the range of all types of images or for a subset of images. An objective model should perform well over a wide range of test images with minimum prediction error. Figure 7. shows two hypothetical models with MOS and the MOSp, and in terms of consistency, model-I is more consistent than model-II.




Model-I (consistent) Model-II (inconsistent)

Fig. 7. Different prediction consistency [3]

The followings are the performance evaluation metrics recommended by VQEG for objective quality assessment model:

## 2.3 Model evaluation by Objective image quality assessment with VIQET

To measure the prediction performance of the objective model qualitatively, following the standard performance evaluation procedure recommended in VQEG [2].
Where mainly linear correlation coefficient, average absolute prediction error (AAE), RMSE, and OR between predicted objective scores MOSp and subjective scores MOS were used for evaluation. In order to verify the permanence of the new model, using the VIQET for the Objective image quality assessment.

By considering the fifty images of natural scenes database. The database is divided into four categories for training and testing (in Appendix A):

- Outdoor daylight
- Indoor arrangements
- Indoor wall hang
- Outdoor night

Images loaded into the VIQET per categories and were analyzed with the tool.

## 2.4 VIQET image quality analysis scores and predicted MOSp

The VIQET generated image quality scores of five image quality attributes of the fifty images loaded into the tool. MOSp is the predicted MOS of each image for Subjective image quality assessment by human.

Metric 1: Pearson Correlation Coefficients of outdoor day

The Pearson correlation coefficient is a very helpful statistical formula that measures the strength between variables and relationships. When conducting a statistical test between two variables, it is a good idea to conduct a Pearson correlation coefficient value to determine just how strong that relationship is between those two variables.

In this section the strength and relationship between MOSp and MOS of outdoor day images is measured. The values of MOS and MOSp of the outdoor day images are presented in Table 1. and the MOSp relation to MOS is illustrated in Figure 8.

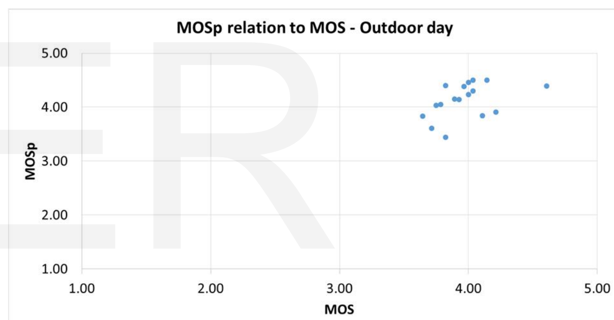



Fig. 8. MOSp relation to MOS

$$CC = \frac{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)\left(MOS_p(i) - \overline{MOS_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)^2}\sqrt{\sum_{i=1}^{N}\left(MOS_p(i) - \overline{MOS_p}\right)^2}}$$

(3.1)

where the index *i* denotes the image sample and N denotes the total number of samples.

Table 1. Pearson correlation coefficients

| MOS | MOSp | Pearson correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 0.44 |

It has also been observed from the Table 1. that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of outdoor day

The calculation of Pearson's correlation for this data gives a value of 0.44 which does not reflect that there is indeed a perfect relationship between the data. Spearman's correlation for this data however is 1, reflecting the perfect monotonic relationship. Spearman's correlation works by calculating Pearson's correlation on the ranked values of this data. Ranking (from low to high) is obtained by assigning a rank of 1 to the lowest value.

If we look at the plot of the ranked data in Figure 9. then we see that they are linearly related.

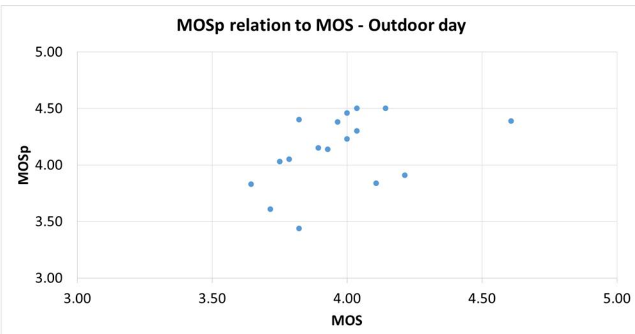



Fig. 9. Ranked data of MOSp relation to MOS

$$SROCC = 1 - \frac{6\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2}{N(N^2-1)} \quad (3.2)$$

where 6 is a constant (it is always used in the formula).

Table 2. SROCC

| MOS | MOSp | Spearman rank order correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 1.00 |

The significant Spearman correlation coefficient value of 1.00 confirms what was apparent from the graph, there appears to be a strong positive correlation between the two variables MOSp and MOS.

Metric 3: Outlier ratio of outdoor day

This metric evaluates an objective model's ability to provide consistently accurate predictions for all types of video sequences and not fail excessively for a subset of sequences, i.e., prediction consistency. The model's prediction consistency can be measured by the number of outlier points (defined as having an error greater than some threshold as a fraction of the total number of points). Figure 10. presents the Outlier Ration values, a smaller outlier fraction means the model's predictions are more consistent.

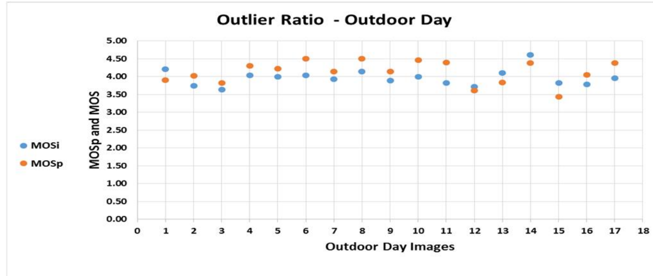



Fig. 10. Outlier ratio values

The objective test plan specifies this metric as follows:

Outlier Ratio = outliers / N

$$OR = \frac{(total\ number\ of\ outliers)}{N} \quad (3.3)$$

where an outlier is a point for: $|MOS(i) - MOS_p(i)| > 2\times\sigma(MOS(i))$, where $\sigma(MOS(i))$ represents the standard deviation of the individual scores associated with the image sample *i*.

Table 0. Outlier ratio

| MOS | MOSp | Outlier ratio |
|---|---|---|
| 3.97 | 4.13 | 0.036 |

The smallest Outlier Ratio is better. Table 3. shows the outlier ratio (OR = 0.036) for the new model calculated over the main partitions of the subjective data.

Metric 4: Average/Mean absolute prediction error of outdoor day

The MAE measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables. Expressed in words, the MAE is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. The MAE is a linear score which means that all the individual differences are weighted equally in the average.

$$AAE = \frac{1}{N}\sum_{i=1}^{N}|MOS(i) - MOS_p(i)| \quad (3.4)$$

Table 4. Average absolute prediction Error

| MOS | MOSp | Average absolute prediction Error |
|---|---|---|
| 3.97 | 4.13 | 1.00 |

Metric 5: Root mean square error of outdoor day

The RMSE is a quadratic scoring rule which measures the average magnitude of the error. The equation for the RMSE is given in both of the references. Expressing the formula in words, the difference between forecast and corresponding observed values are each squared and then averaged over the sample. Finally, the square root of the average is taken. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means the RMSE is most useful when large errors are particularly undesirable.

The MAE and the RMSE can be used together to diagnose the variation in the errors in a set of forecasts. The RMSE will always be larger or equal to the MAE; the greater difference between them, the greater the *variance* in the individual errors in the sample. If the RMSE=MAE, then all the errors are of the same magnitude

Both the MAE and RMSE can range from 0 to ∞. They are negatively-oriented scores: Lower values are better.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2} \tag{3.5}$$

Table 5. RMSE

| MOS | MOSp | Root mean square error |
|---|---|---|
| 3.97 | 4.13 | 0.16 |

Lower value of RMSE is better, Table 5. presents a very low value of RMSE (0.16) for the indoor images, which strongly supports the good performance of new model prediction accuracy.

Metric 1: Pearson correlation coefficients of indoor

In this section the strength and relationship between MOSp and MOS of indoor images is measured.
The values of MOS and MOSp of the indoor images are presented in Table 5. and the MOSp relation to MOS is illustrated in Figure 11.

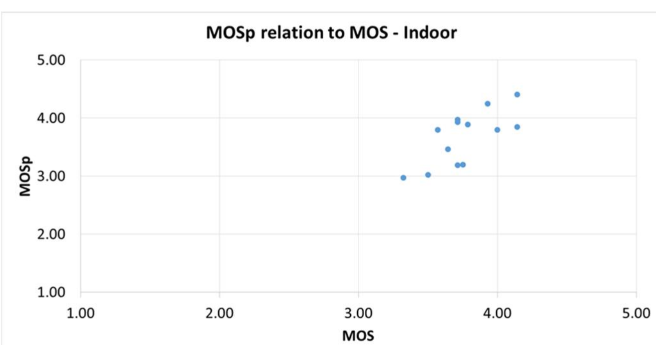



Fig. 11. MOSp relation to MOS

$$CC = \frac{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)\left(MOS_p(i) - \overline{MOS_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)^2}\sqrt{\sum_{i=1}^{N}\left(MOS_p(i) - \overline{MOS_p}\right)^2}} \tag{3.6}$$

Where the index *i* denotes the image sample and N denotes the total number of samples.

Table 6. Pearson correlation coefficients

| MOS | MOSp | Pearson correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 0.72 |

Table 6. shows that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of indoor

The calculation of Pearson's correlation for this data gives a value of 0.72, which reflect that there is a very good relationship between the data. Spearman's correlation for this data however is 1, reflecting the perfect monotonic relationship. Spearman's correlation works by calculating Pearson's correlation on the ranked values of this data. Ranking (from low to high) is obtained by assigning a rank of 1 to the lowest value.

If we look at the plot of the ranked data in Figure 12. then we see that they are linearly related.

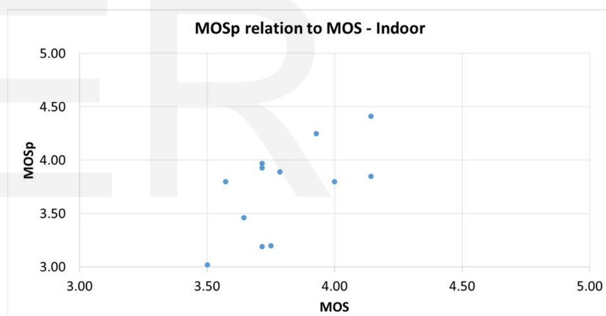



Fig. 12. Ranked data of MOSp relation to MOS

$$SROCC = 1 - \frac{6\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2}{N(N^2-1)} \tag{3.7}$$

where 6 is a constant (it is always used in the formula).

Table 7. SROCC

| MOS | MOSp | Spearman rank order correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 1.00 |

The significant Spearman correlation coefficient value of 1.00 confirms what was apparent from the graph, there appears to be a strong positive correlation between the two variables MOSp and MOS.

Metric 3: Outlier ratio of indoor

This metric evaluates an objective model's ability to provide consistently accurate predictions for all types of image sequences and not fail excessively for a subset of sequences, i.e., prediction consistency. The model's prediction consistency can be measured by the number of outlier points (defined as having an error greater than some threshold as a fraction of the total number of points). Figure 13. presents the Outlier Ration values, a smaller outlier fraction means the model's predictions are more consistent.

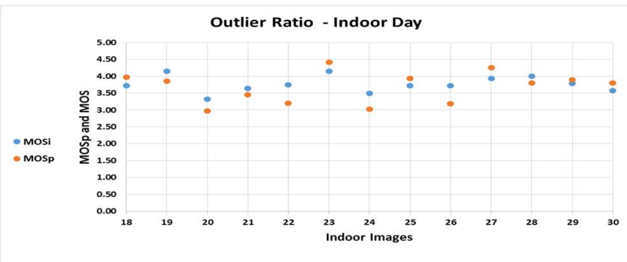



Fig. 13. Outlier ratio values

The objective test plan specifies this metric as follows:
Outlier Ratio = outliers / N

The smallest Outlier Ratio is better. Table 8. shows the outlier ratio (OR = 0.11) for the new model calculated over the main partitions of the subjective data.

$$OR = \frac{(total\ number\ of\ outliers)}{N} \qquad (3.8)$$

where an outlier is a point for: $|MOS(i) - MOS_p(i)| > 2\times\sigma(MOS(i))$, where $\sigma(MOS(i))$ represents the standard deviation of the individual scores associated with the image sample *i*.

Table 8. Outlier ratio

| MOS | MOSp | Outlier ratio |
|---|---|---|
| 3.97 | 4.13 | 0.11 |

Metric 4: Average/Mean absolute prediction error of indoor

The MAE measures the average magnitude of the errors in a set of forecasts, without considering their direction. It measures accuracy for continuous variables of the indoor images. Expressed in words, the MAE is the average over the verification sample of the absolute values of the differences between forecast and the corresponding observation. The MAE is a linear score which means that all the individual differences are weighted equally in the average.

$$AAE = \frac{1}{N}\sum_{i=1}^{N}|MOS(i) - MOS_p(i)| \qquad (3.9)$$

Table 9. Average absolute prediction error

| MOS | MOSp | Average absolute prediction error |
|---|---|---|
| 3.97 | 4.13 | 0.31 |

Metric 5: Root mean square error of indoor

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2} \qquad (3.10)$$

Table 10. Root mean square error

| MOS | MOSp | Root mean square error |
|---|---|---|
| 3.97 | 4.13 | 0.09 |

Lower value of RMSE is better, Table 10. presents a very low value of RMSE (0.09) for the indoor images, which strongly supports the good performance of new model prediction accuracy.

Metrics 4 and Metrics 5 are considered as a measure of prediction accuracy. The values in above exhibit good accuracy, monotonicity, and consistency in predictions. The measurement of prediction accuracy and monotonicity can be measured by Pearson correlation and Spearman rank order correlation metrics, whereas the consistency can be evaluated by the number of outlier points.

Metric 1: Pearson correlation coefficients of outdoor night

In this section the strength and relationship between MOSp and MOS of outdoor Night images is measured. The values of MOS and MOSp of the indoor images are presented in Table 11. and the MOSp relation to MOS is illustrated in Figure 14.

Table 11. MOSi and MOSp values

| Image | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|
| MOSi | 2.50 | 3.57 | 3.64 | 3.54 | 3.32 | 3.25 | 3.29 | 3.07 | 3.32 | 1.25 |
| MOS | 2.01 | 3.45 | 3.44 | 3.04 | 2.79 | 2.83 | 3.02 | 2.10 | 3.84 | 1.46 |

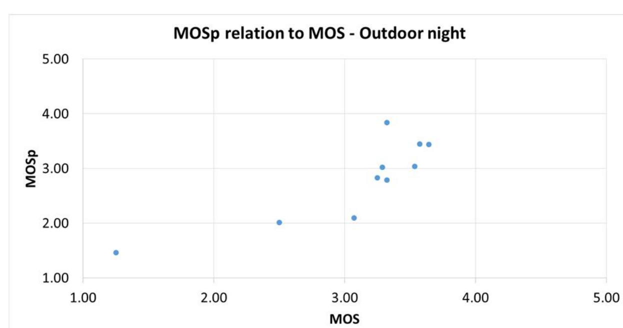



Fig. 14. MOSp relation to MOS

International Journal of Scientific & Engineering Research, Volume 8, Issue 2, February-2017
ISSN 2229-5518


$$CC = \frac{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)\left(MOS_p(i) - \overline{MOS_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)^2}\sqrt{\sum_{i=1}^{N}\left(MOS_p(i) - \overline{MOS_p}\right)^2}} \quad (3.11)$$

where the index $i$ denotes the image sample and N denotes the total number of samples.

Table 12. Pearson correlation coefficients

| MOS | MOSp | Pearson correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 0.84 |

The high Pearson Correlation Coefficient value (0.84) which observed in Table 12. is strongly support that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of outdoor night

The calculation of Pearson's correlation for this data gives a value of 0.84, which reflect that there is a very good relationship between the data. Spearman's correlation for this data however is 1, reflecting the perfect monotonic relationship. Spearman's correlation works by calculating Pearson's correlation on the ranked values of this data. Ranking (from low to high) is obtained by assigning a rank of 1 to the lowest value.

If we look at the plot of the ranked data in Figure 15. then we see that they are linearly related.

$$SROCC = 1 - \frac{6\sum_{i=1}^{N}\left(MOS(i) - MOS_p(i)\right)^2}{N(N^2 - 1)} \quad (3.12)$$

Where 6 is a constant (it is always used in the formula).

Table 13. Spearman rank order correlation coefficient

| MOS | MOSp | Spearman rank order correlation coefficient |
|---|---|---|
| 3.97 | 4.13 | 1.00 |

The significant Spearman correlation coefficient value of 1.00 confirms what was apparent from the graph, there appears to be a strong positive correlation between the two variables MOSp and MOS.

Metric 3: Outlier ratio of outdoor night

This metric evaluates an objective model's ability to provide consistently accurate predictions for all types of image sequences and not fail excessively for a subset of sequences, i.e., prediction consistency. The model's prediction consistency can be measured by the number of outlier points (defined as having an error greater than some threshold as a fraction of the total number of points). Figure 3.14 presents the Outlier Ration values, a smaller outlier fraction means the model's predictions are more consistent.

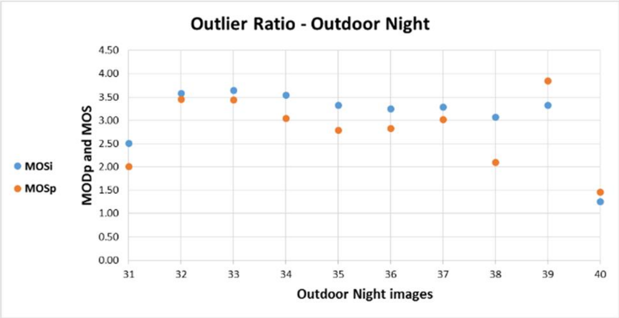



Fig. 15. Outlier ration

The objective test plan specifies this metric as follows:
Outlier Ratio = outliers / N

$$OR = \frac{(total\ number\ of\ outliers)}{N} \quad (3.13)$$

where an outlier is a point for: $|MOS(i) - MOS_p(i)| > 2\times\sigma(MOS(i))$, where $\sigma(MOS(i))$ represents the standard deviation of the individual scores associated with the image sample $i$.

Table 14. Outlier ratio

| MOS | MOSp | Outlier ratio |
|---|---|---|
| 3.97 | 4.13 | 0 |

The smallest OR is better. Table 14. shows the outlier ratio (OR = 0) for the new model calculated over the main partitions of the subjective data.

Metric 4: Average/Mean absolute error of outdoor night

The Average/Mean absolute error is a quantity used to measure how close forecasts or predictions are to the eventual outcomes.

MAE between objective MOSp and subjective MOS scores is defined by:

$$AAE = \frac{1}{N}\sum_{i=1}^{N}\left|MOS(i) - MOS_p(i)\right| \quad (3.14)$$

Table 15. Average absolute prediction error

| MOS | MOSp | Average absolute prediction error |
|---|---|---|
| 3.97 | 4.13 | 0.42 |

Metric 5: Root mean square error of outdoor night

The MAE and the RMSE were used together to diagnose the variation in the errors in a set of forecasts. The RMSE will always be larger or equal to the MAE; the greater difference between them, the greater the *variance* in the individual errors in the sample. If the RMSE=MAE, then all the errors are of the same magnitude

Both the MAE and RMSE can range from 0 to ∞. They are negatively oriented scores: Lower values are better.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2} \quad (3.15)$$

Table 16. RMSE

| MOS | MOSp | Root mean square error |
|---|---|---|
| 3.97 | 4.13 | 0.28 |

Lower value of RMSE is better, Table 16. presents a very low value of RMSE (0.28) for the indoor images, which strongly supports the good performance of new model prediction accuracy.

Metrics 4 and Metrics 5 are considered as a measure of *prediction* accuracy.

The values in above exhibit good accuracy, monotonicity, and consistency in predictions. The measurement of prediction accuracy and monotonicity can be measured by Pearson correlation and Spearman rank order correlation metrics, whereas the consistency can be evaluated by the number of outlier points.

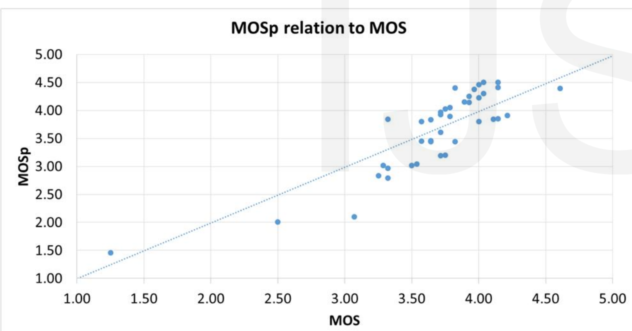



Fig. 16. MOSp relation to MOS of all images

## 3 New Model and New Frame Work Implementation

The framework describes how to extract the IQ parameters measures and provided by VIQET to MOSp (predicted MOS).

IQ attributes of VIQET and their value range

- Multi-scale Edge Acutance (MsEA) – rang: higher is better
- Noise Signature Index (NSI) – range: 0 – 590
- Color saturation (CS) – range: 0 -255
- Illumination (IL) – range: 0 – 255
- Dynamic Range (DR) – 0 – 255

### 3.1 RMS calculations of IQ attributes values of VIQET

Once the new model was evaluated (in Chapter III) against MOS received from humans through HVTs, the RMS (Root Mean Square) value of each VIQET IQAa values was calculated as follows:

$$MsEA_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N} \boldsymbol{MsEAi}^2} \quad (3.16)$$

While $MsEA_i$ is the value measured by VIQET of each image under IQ evaluation.

$$DR_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N} \boldsymbol{DRi}^2} \quad (3.17)$$

While DRi is the value measured by VIQET of each image under IQ evaluation.

$$CS_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N} \boldsymbol{CSi}^2} \quad (3.18)$$

While CSi is the value measured by VIQET of each image under IQ evaluation.

$$NSI_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N} \boldsymbol{NSIi}^2} \quad (3.19)$$

While NSIi is the value measured by VIQET of each image under IQ evaluation.

$$IL_{RMS} = \sqrt{\frac{1}{N}\sum_{i=0}^{N} \boldsymbol{ILi}^2} \quad (3.20)$$

While ILi is the value measured by VIQET of each image under IQ evaluation.

MOSp should be calculated by Formula (3.21).

The MOSp =

$$\left(\frac{\text{MsEA}}{\text{MsEArms}} * 0.25\right) + \left(\frac{\text{NSI}}{\text{NSImax}} * 0.167\right) + \left(\frac{\text{CS}}{\text{CSmax}} * 0.167\right) + \left(\frac{\text{IL}}{\text{ILmax}} * 0.167\right) + \left(\frac{\text{DR}}{\text{DRrms}} * 0.25\right) \quad (3.21)$$

While the MsEA, NSI, CS, IL, DR are the immediate IQ attributes measured by VIQET, divided by the corresponding RMS values. The contribution of each IQAs is weighted according to the IQA effect on the overall perceived IQ.

Instruction for the new model implementation

a. Install the VIQET application.
b. Upload images for IQ evaluation to VIQET.
c. Run the IQ measurement.
d. Use Formula (xx) for predicted MOS.

## 4 General Conclusions and Recommendations

This research proposes a new model consists of a framework and computer based application, the VIQET for smartphones perceived image quality prediction. The framework is composed of a HVTs procedure and an evaluation by the VIQEG.

The VIQET is an objective, no-reference photo quality evaluation tool. VIQET is an open source tool designed to evaluate quality of consumer photos. In order to perform photo quality evaluation, VIQET requires a set of photos from the test device. It estimates an overall MOS for a device based on the individual image MOS scores in the set. This thesis provides a detailed description and analysis of subjective image quality assessment through HVT and objective image quality assessment based on VIQET analysis.

The correlations between the metrical and perceptual results indicated that MOS, MSE, PSNR metrics give excellent prediction performance in most cases in terms of both correlation and its variance. According to the group comparison had comparatively better prediction performance than no reference metrics.

The statistical analyses were conducted to check whether the increase of the image quality attributes would lead to improvement in user's perceived image quality.

The finding is useful for the mobile phone industry to have a better understanding of the concrete benefit of enhancing the image quality attributes. The proposed quality assessment model is useful also for image quality assessment of any mobile or desktop displays.

One unique feature of this proposed framework was the capability of incorporating existing full reference image quality metrics without modifying them. This research implemented the framework for smartphones displays, and used the framework to evaluate the prediction performance of state-of-the-art image quality metrics regarding the most important image quality attributes for projection displays. The evaluated image quality attributes were brightness, contrast, color saturation and sharpness, however the proposed framework was not bound by the possibilities. All the metric evaluations were supported by the correlation of objective and subjective experimental results

In addition, this study also investigated the strategies to extend subjective experiments with baseline adjustment method, which is expected to save quite a lot of time and resources for subjective experiments. In a broader point of view, the originally defined research scope have been fully covered by the research presented in this thesis, all research goals have been successfully achieved, and the corresponding research questions have been answered. The proposed image quality assessment framework was originally designed for smartphones displays, but could be easily adapted to other types of displays with limited modifications.

In conclusion, with the results that obtained in this study, that the framework and the new approach provided by this research can be a good process for perceived image quality prediction.

## Acknowledgement

Special thanks to my supervisors for great suport and advice to:
Prof. Dr.habil. Florentin PALADI
Vice-rector for research, Moldova State University.
Prof. Dr. Tudor Bragaru
Associate Professor, Mathematics and Informatics Department, Moldova State University.

## REFERENCES

[1] SHEIKH H. R., BOVIK A. C. AND CORMACK L. K., "No-Reference Quality Assessment Using Natural Scene Statistics: JPEG2000," IEEE Transactions on Image Processing, vol. 14, no. 11, pp. 1918-1927, December 2005.

[2] VIDEO QUALITY EXPERTS GROUP (VQEG). Final report from the VQEG on the validation of objective models of video quality assessment, FR-TV Phase II. http://www.vqeg.org/.

[3] PARVEZ SAZZAD Z. M., No-Reference image quality Assessments for JPEG and JPEG2000 Coded Images. PhD thesis, University of Toyama, Department of Systems Science and Engineering, March 2008.

[4] ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures, International Telecommunication Union, Geneva, Switzerland, 2002.

[5] VQEG: Final report from the video quality experts group on the validation of objective models of video quality assessment. Available: http:/ / www.vqeg.org.

[6] ITU-T, "Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment," Recommendation P.913, ITU-T, 2014.

[7] ITU, "Methodology for the subjective assessment of the quality of television pictures," ITU-R Recommendation BT 500-13, ITU Radio communication Sector, 2012.

[8] ITU-T, "Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment," Recommendation P.913, ITU-T, 2014.