UNIVERSITATEA DE STAT DIN MOLDOVA

Cu titlu de manuscris

C.Z.U.: 004.382.76: 004.932:77(043)

ZOREA Pinchas

EVALUAREA CALITĂȚII PERCEPUTE A IMAGINII PENTRU SMARTFONURI CU CAMERA FOTO ÎNCORPORATĂ

Specialitatea 232.02 – Tehnologii, produse și sisteme informaționale

Teză de doctor în tehnică

Conducători științifici:

BRAGARU Tudor
doctor în științe economice,
conferențiar universitar

PALADI Florentin
doctor habilitat în științe fizico-matematice, profesor universitar

Autor:

CHIȘINĂU, 2018

MOLDOVA STATE UNIVERSITY

Presented as manuscript

U.D.C.: 004.382.76: 004.932:77(043)

ZOREA Pinchas

ASSESSMENT OF PERCEIVED IMAGE QUALITY FOR SMARTPHONES WITH EMBEDDED CAMERA

Specialty 232.02 – Information technologies, products and systems

PhD thesis in Technic

Scientific supervisors:



BRAGARU Tudor

Dr. in economic sciences,

associate professor



PALADI Florentin

Dr. habilitate in physics and

mathematics, professor

Author: 

CHISINAU, 2018

© **Zorea Pinchas, 2018**

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisors: prof. dr. habilitate Florentin PALADI and associated prof. dr. Tudor BRAGARU. Their initial encouragement, continuous support and educating advice were essential to the coming about of this thesis.

I would also like to thank my students in “Nofe-Golan” high school and in “Ort Braude” engineering college for their support during the human visual tests. Furthermore, I would like to thank to lots of people that were kindly and peacefully participating in my image quality experiments.

Special thanks to my family, my beloved wife Anat and my beloved children Reut, Or, Noa, Stav, Shai, Nofar and my first Granddaughter Lia.

TABLE OF CONTENTS

ANNOTATION 7

LIST OF TERMS AND ABBREVIATIONS 10

INTRODUCTION 12

1. ANALYSIS OF SITUATION IN THE SMARTPHONES IMAGE QUALITY 18

1.1. Image quality assessment background 18

1.2. The concept of image quality assessment 19

1.2.1. The image quality circle 22

1.3. Smartphones vendors in-house image quality assessment 23

1.4. Image quality attributes 25

1.4.1. Relationships between image quality attributes 28

1.4.2. Summary of image quality attributes 29

1.5. Subjective image quality assessment 30

1.6. Objective image quality assessment 30

1.7. The challenge of non-reference image quality assessment 34

1.8. Research on perceived image quality arena 35

1.9. Summary 41

2. PERCEIVED IMAGE QUALITY ASSESSMENT 43

2.1. Image quality attributes of VIQET 44

2.2. The image quality assessment method 45

2.3. Subjective image quality metrics 46

2.4. HVS-based feature metrics 47

2.5. Implementation of subjective image quality assessment 48

2.6. Implementation for perceived image quality prediction 51

2.7. Human visual test and VIQET image analysis 53

2.7.1. Subjective IQ assessment for IQAs rating 54

2.7.2. Extraction of IQAs to VIQET image quality categories 58

2.7.3. Image quality analysis by VIQET 58

2.8. Examination of the new model 61

2.8.1. Image quality valuation with VIQET 68

2.8.2. The impact of different IQ attributes 75

2.8.3. Evaluation of smartphone perceived image quality 81

2.8.4. Perceived image quality based on IQ attributes 82

2.9. Summary 85

3. NEW MODEL PERFORMANCE EVALUATION ____ 87

3.1. VQEG's standard performance evaluation procedures 88

3.2. Model evaluation by objective image quality assessment with VIQET 92

3.2.1. VIQET image quality analysis scores and predicted MOSp 93

3.3. The new image quality assessment model evaluation 95

3.4. Model for perceived image quality evaluation 107

3.5. The study novelty and practical implementation 111

3.6. Summary 113

GENERAL CONCLUSIONS AND RECOMMENDATIONS ____ 115

BIBLIOGRAPHY ____ 117

LIST OF FIGURES ____ 126

Appendix 1. Natural scenes images of test content ____ 129

Appendix 2. VIQET users' guide ____ 134

Appendix 3. Ishihara color plates ____ 147

Appendix 4. Voting card for IQ assessment ____ 148

Appendix 5. Ort Braude Engineering College certification ____ 149

DECLARATION OF GUARANTOR ____ 150

CURRICULUM VITAE ____ 151

ANNOTATION

of PhD thesis "*Assessment of perceived image quality for smartphones with embedded camera*"
Specialty 232.02 – *Information Technologies, Products and Systems,*
presented by Pinchas ZOREA, Moldova State University, Chisinau, 2018
to obtain the title of doctor in the Technical Sciences

Thesis structure: The thesis contains Introduction, 3 chapters, general conclusions and recommendations, bibliography of 101 titles. The main text amounts to 107 pages, includes 66 figures, 30 tables, 37 formulas, and 5 appendixes. The obtained results of the thesis were published in 7 scientific papers.

Keywords: human visual test (HVT), perceived image quality (PIQ), image quality attributes (IQAs), mean opinion score (MOS), objective/subjective image quality assessment, video quality experts group (VQEG), VQEG image quality evaluation tool (VIQET).

Aim of research is to develop a new no or zero (NR) reference subjective quality assessment model and framework that would enable the smartphones industry prediction of the PIQ.

The objectives of thesis include the design of NR subjective and objective image quality metrics based on extensive visual tests experiments and evaluation by SW tool VIQET that measure the perceived image quality of smartphones users.

Scientific novelty and originality of the obtained results is reflected in a new approach which predicts the PIQ by extraction of the classic, objective IQAs *(brightness, contrast, color saturation and sharpness)* to new subjective IQ criteria that evaluated by VIQET. This is a new assessment process based on a new diploid model, which aims to predict how the smartphones users perceived image quality.

Important scientific solved problem consists in elaborating a new diploid model for PIQ assessment that describes how to extract the standard IQAs which used in large screens TVs into new IQAs of SW tool VIQET. This can be used by smartphone producers and vendors in order to shorten „time to market", or by any image quality experts, including in the academia in order to reduce time and cost of PIQ assessment process of smartphones with small HD displays in comparison to the process that based on many human physical tests.

Theoretical significance is supported by analyzing, specifying and defining the theoretical principles and new PIQ criteria which implemented in the SW tool VIQET in order to predict the expected *MOS* in HVTs. The new model determines the relationship between the standard IQAs values and the VIQET PIQ criteria.

Applicative value of the work is determined by the developed model, which has huge potential for smartphone industry and users in reducing significantly the time and cost of PIQ assessment.

Implementation of results: the obtained results are used in „Ort Braude College of Engineering" in Israel, and can further be used by students and researchers in image processing. Also, an Israeli company (ASI – Applied Spectral Imaging Ltd.) evaluated the PIQ framework and found the benefits of PIQ framework for IQ improvement of images taken by medical spectral imaging system.

ADNOTAREA

tezei de doctorat

„Evaluarea calității percepute a imaginii pentru smartfonuri cu camera foto încorporată”

Specialitatea 232.02 – *Tehnologii informaționale, produse și sisteme,*

prezentată de Pinchas ZOREA, Universitatea de Stat din Moldova, Chișinău, 2018

pentru obținerea titlului de doctor în științe tehnice

Structura tezei: Teza constă din introducere, 3 capitole, concluzii generale și recomandări, bibliografie (101 titluri). Textul de bază este expus pe 107 pagini, incluzând 66 figuri, 30 tabele, 37 formule și 5 anexe. Rezultatele obținute au fost publicate în 7 lucrări științifice.

Cuvinte-cheie: test vizual uman (HVT), calitate percepută a imaginii (PIQ), atribute de calitate a imaginii (IQAs), media scorurilor de opinie (MOS), evaluare obiectivă/subiectivă a calității imaginii, grup de experți în domeniul calității imaginii (VQEG), instrumentar VQEG de evaluare a calității imaginii (VIQET).

Scopul cercetării constă în dezvoltarea unui nou cadru și model de evaluare subiectivă a calității imaginii fără referință la original (NR), care ar permite industriei de smartfonuri predicția PIQ.

Obiectivele tezei includ proiectarea metricilor subiective și obiective de calitate a imaginii fără referință la original NR, în baza unor multiple teste vizuale experimentale și evaluării cu instrumentarul VIQET, care permite măsurarea PIQ de către utilizatorii de smartfonuri.

Noutatea științifică și originalitatea rezultatelor se reflectă în noua abordare a predicției PIQ prin extragerea noilor criterii subiective de calitate evaluate cu instrumentarul VIQET din atributele clasice de calitate IQAs *(luminozitate, contrast, saturație de culoare* și *claritate)*. Acesta reprezintă un nou proces de evaluare bazat pe un nou model diploid, care urmărește să prezică modul în care utilizatorii de smartfonuri percep calitatea imaginilor.

Problema științifică importantă soluționată constă în elaborarea unui nou model diploid de evaluare a PIQ, care descrie cum pot fi extrase noile IQAs, folosite de instrumentarul VIQET din IQAs standarde, folosite pentru televizoare cu ecrane mari. Acest lucru poate fi utilizat de către producătorii și furnizorii de smartfonuri pentru a scurta „timpul ieșirii pe piață” sau de către orice alți experți și cercetători, inclusiv din mediul academic, pentru a reduce timpul și costul procesului de evaluare a calității percepute pentru smartfonuri cu ecrane mici de înaltă rezoluție (HD) comparativ cu multiplele teste fizice.

Semnificația teoretică este susținută de analiza, specificarea, definirea principiilor teoretice și a noilor criterii PIQ, implementate în VIQET pentru previziunea valorii medii așteptate MOS în HVT. Noul model stabilește relația dintre valorile IQAs standarde și criteriile utilizate în VIQET.

Valoarea aplicativă este determinată de noul model elaborat, ce are un potențial solid pentru industria și utilizatorii de smartfonuri în ce privește reducerea semnificativă a timpului și a costurilor de evaluare a PIQ.

Implementarea rezultatelor: rezultatele obținute sunt utilizate în Ort Braude College of Engineering (Israel) și pot fi utilizate cu succes de către studenți și cercetători în domeniul procesării imaginilor. Compania israeliană (ASI-Applied Spectral Imaging Ltd.), de asemenea a evaluat noua metodă PIQ și a constatat avantajele acesteia în îmbunătățirea imaginii calității în sistemele medicale spectrale.

АННОТАЦИЯ

Диссертационной работы *«Оценка воспринимаемого качества изображения для мобильных устройств со встроенной камерой»,*
Специальность 232.02 – *Информационные технологии, системы и продукты*
представленной господином Пинхас ЗОРЯ,
Государственный Университет Молдавии, Кишинев, 2018
на соискание учёной степени кандидата технических наук.

Структура работы: Диссертация состоит из введения, трех глав, общих выводов и рекомендаций, библиографии, включающей 101 название. Основной текст представлен на 107 страниц, включая 66 рисунков, 30 таблиц, 37 формул и 5 приложений. Полученные результаты были опубликованы в 7 научных статьях.

Ключевые слова: визуальный тест, осуществлённый человеком (HVT), воспринимаемое качество изображения (PIQ), атрибуты качества изображения (IQAs), среднее значение оценок (MOS), объективная/субъективная оценка качества изображения, экспертная группа по качеству видео (VQEG), инструмент VQEG для оценки качества изображения (VIQET).

Цель исследования заключается в разработке новой модели субъективной оценки качества, без ссылки на оригинал NR, которая позволит производителям смартфонов прогнозировать PIQ.

Задачи исследования включают разработку субъективных и объективных NR показателей качества на основе множества визуальных экспериментальных тестов и оценки при помощи программного приложения VIQET, позволяющий измерять воспринимаемое пользователями смартфонов качество изображения.

Научная новизна и оригинальность результатов состоит в разработке нового подхода к прогнозированию PIQ, который позволяет извлечь классические объективные IQAs (*яркость, контраст, насыщенность цвета* и *резкость),* в новых IQAs критериях, используемых в VIQET. Это новый процесс оценки, основанный на новой диплоидной модели, целью которой является предсказание того как пользователи смартфонов воспринимают качество изображения.

Важной научной задачей решенной в исследовании является разработка новой диплоидной модели оценки PIQ, описывающей как извлечь новые IQAs инструмента VIQET из стандартных IQAs, используемые для больших телевизионных экранов. Это может использоваться производителями и поставщиками смартфонов для сокращения времени выхода на рынок, или любыми экспертами PIQ, в том числе в академических кругах, для сокращения времени и снижения стоимости оценки PIQ для смартфонов с малыми экранами высокой четкости (HD) в сравнение с ручным процессом на основе множества визуальных тестов, осуществляемых людьми.

Теоретическая значимость работы подтверждается анализом и определением теоретических принципов и новых критериев PIQ, внедрённых в программном приложении VIQET для прогнозирования ожидаемой MOS в HVT. Новая модель определяет взаимосвязь между стандартными значениями IQAs и критериями IQ VIQET.

Практическая применяемость работы определена разработанной моделью, которая имеет огромный потенциал для индустрии и пользователей смартфонов, заключающийся в сокращении времени и снижении стоимости оценки PIQ.

Внедрение результатов: полученные результаты используются в «Ort Braude College of Engineering» (Израиль) и может быть успешно использована студентами и исследователями в области обработки изображений. Израильская компания (ASI - Applied Spectral Imaging Ltd.) также оценила новый метод PIQ и нашла преимущества для улучшения IQ визуализации изображений в медицинских спектральных системах.

LIST OF TERMS AND ABBREVIATIONS

Color Accuracy	Colors captured by the camera compared to expected color on display
Color Space	A color model is an abstract mathematical model describing the way colors can be represented as tuples of numbers, typically as three or four values or color components (e.g. RGB and CMYK are color models).
Color Uniformity	Measurement of pixel's color throughout the full frame
CPIQ	Camera Phone Image Quality
DMOS	Difference Mean Opinion Score
DMOSp	Difference Mean Opinion Score, predicted
DR	Dynamic Range
FR	Full Reference
HD	High Definition
HDR	High Dynamic Range
HLS	Hue, Lightness, Saturation
HSV	Hue, Saturation, Value
HVS	Human Visual System
HVT	Human Visual Test
HW	Hard Ware
IEEE	Institute of Electrical and Electronics Engineers
IQ	Image Quality
IQA	Image Quality Assessment/Attribute
ISP	Image Signal Processing / Image Signal Processor
ITU	International Telecommunications Union
JND	Just Noticeable Difference
JPEG	Joint Picture Experts Group
LCD	Liquid Crystal Display
LPCC	Linear Pearson Coefficients Correlation
MOS	Mean Opinion Score
MOSp	Mean Opinion Score, predicted
MSE	Mean Squared Error
NR	No (or Zero) Reference
PSNR	Peak Signal-to-Noise Ratio

RGB	Red, Green, Blue
RMSE	Root Mean Squared Error
RR	Reduced Reference
SBS	Side By Side
SMIA	Standard Mobile Imaging Architecture
SNR	Signal-to-Noise Ratio
SW	Software
UIQM	Universal image quality Model
VIQET	VQEG Image Quality Evaluation Tool
VQEG	Video Quality Experts Group

INTRODUCTION

The aim of this research study is to provide ***a new framework*** to predict the perceived image quality by smartphones users. The motivation for this research rise due to the rapid technology development in the mobile devices with display and embedded camera.

Smartphones are increasingly becoming primary photos acquiring and display devices for consumers. Smartphones vendors are striving at improving camera capabilities and delivered quality of captured photos and videos. Smartphones camera image quality is driving consumer preference in smartphones purchasing and affects popular consumer usage models like image and video sharing (Instagram, Facebook, etc.).

Smartphones and tablet computers with embedded digital camera acquire and display high-definition images and videos continues to advance dramatically. It is estimated [1] that more than half of digital photographs and videos being taken by smartphones. Developing objective image quality assessment model to predict the perceptual image quality of smartphones, as digital visual media has become necessary, as the everyday photography by smartphones continues to increase exponentially. The most common way of determining the image quality of an image or video is to perform visual experiments in order to solicit opinion from human observers, which is time consuming, expensive and requires special infrastructure (visual labs).

This research presents the principles and methods of image quality assessment and provides new model for predicting the image quality of smartphones. There are several researches on NR subjective image quality assessment for PIQ which propose new models and new process based on: using natural scene statistics [2, 3], measure the impacts of appearance parameters on perceived image quality for mobile-phone displays [4] or measure the impact of three image quality attributes (colorfulness, contrast and brightness) [5]. None of the proposed existing models addresses the requirement for reliable objective PIQ prediction process.

The relevance and importance of the raised problem. Smartphones became an essential part in our life. People of all generations are making more and more use of mobile devices as multimedia entertainment device with high image quality delivered by the HD display and high resolution embedded cameras (back & front side camera).The image content processed and displayed by these digital imaging systems largely differs in perceived quality depending on the system and its applications. The displayed content is captured either by the embedded camera or from external source (e.g. other mobile device, Facebook, Instagram, You Tube etc.).

What is image quality testing in industry?

- Using full reference or reduced reference metrics;
- The process is time consuming and expensive;
- Need results immediately – "Time to market";

What is image quality testing in academia/research?

- Using full reference metric;
- Dealing with new algorithms compared to old versions;
- Dealing with known distortions (Noise, Gaussian blur, JPEG artifacts, etc.);
- Have availability of time, not marketing products;
- Tests done under controlled conditions (test targets, light, distance etc.);

If vendor A measures 5 and vendor B measures 2, who's to say which score is right?

An alternative process is needed in order to address:

- Use standard image quality assessment metric across the industry;
- Correlate objective results with perceived IQ by human;
- Meaningful consumer rating scale (in 1-5 "stars");

Mobile device producers are striving at differentiating through camera capabilities and delivered image quality of captured images. They measure the perceived IQ in long and expensive process. The perceived image quality is measured by outsource special labs which is expensive and time consuming. This study provides an ***efficient, reliable and quick evaluation solution that provides in real time clear answer to how the smartphone users perceive the image quality.*** This solution can be implemented in-house. Moreover, in-house process secures the intellectual property (IP) not to be disclosed outside the company. As the camera and display quality are driving consumer preference in mobile device purchasing and impacts popular applications like image and video sharing (Instagram, Facebook, etc.). In addition, the new model supports the smartphones vendors to be able to optimize the experience of smartphones users.

Although human beings judge image quality in a real-time without reference is a subjective image quality assessment. ***Developing a model to simulate this perception is still an industrial (by mobile devices vendors) and academic challenge***. The outcome of this research is a reliable benchmark and image quality metrics for no-reference objective image quality, which can quantitatively predict perceived image quality.

The image quality assessment is for smartphones. Smartphones with embedded camera and imaging systems already have millions of users and exist enormous growth potential.

Therefore, focusing research into technology and user interaction in image quality is increasingly important through combining technical parameters and human perception. The signals addressing these imaging systems originate from a diversity of sources. Image content can be either computer generated, or can be recorded with a simple image camera or with a more sophisticated video capturing device.

Research on modeling image quality in a FR framework where the original content can be used as a reference is well established in literature and in the industry in many current applications. However, with the latest advances in smartphones with embedded digital cameras and HD display and their vast use to capture images and videos, perceived image quality has raised another level of investigation. A new NR objective image quality assessment is required in order to measure and predict the smartphones users perceived image quality. What makes an image more pleasant than another one? What process is required to improve the perception of the image quality *?* To answer these questions, NR image quality assessment methods are required. However, so far, most of the existing methods and framework on objective NR perceptual image quality are focused on the impact of image compression and transmission.

The goal of this thesis is to ***develop a new model and framework of NR image quality assessment, which aims to predict the image quality perception by smartphone consumers.***

The objective of the new model includes the design of NR subjective and objective image quality metrics based on extensive visual tests experiments and evaluation, predicted of perceived IQ with VIQET. This includes a new image quality assessment model based on SW tool that measure the perceived image quality of smartphones users. The SW tool is the VIQET, which have image quality coefficients used for image quality assessment by that tool. The VIQET coefficients were calculated based on the new model and modified in the VIQET. The VIQET coefficients were modified while the standard IQAs (brightness, contrast, colors and sharpness) were extracted to the VIQET IQ criteria in order to be able to quantitatively predict perceived image quality.

At baseline, the research examines how the smartphone users perceive the image quality of their devices via visual tests in NR subjective image quality assessment. Scores given by the observers have been analyzed and compared with VIQET scores. As result of the analysis of the HVT and VIQET scores, the VIQET image quality attributes coefficients were recalculated and modified in the tool as essential part of the new model development. This procedure has been ran until the highest correlation between HVT scores and VIQET scores have been achieved.

The research objectives:

- Situation analysis of the NR IQ assessment methods;
- Identify the most important IQAs that make the image more "pleasant" to smartphones users;
- Investigate the relationship between standard IQAs and VIQET IQ criteria;
- Develop new model to extract the standard IQAs to VIQET IQ criteria;
- Improve the new model based on comparative analysis with HVT results;
- New model performance evaluation according to VQEG recommendations;
- Address the requirement for a reliable and efficient process for prediction of perceived IQ;

The scientific novelty of the obtained results. However, with the latest digital image processing advances in digital cameras and smartphones display, and their widespread use to capture photos and videos, image quality has opened another level of investigation. In particular, ***perceived image quality of photos captured by such devices cannot be evaluated due to missing a reference image to compare with***. Therefore a new reliable NR image quality assessment model which provides an immediate IQ scoring is required. In this research the scientific novelty is the new model for reliable prediction of perceived IQ, based on the VIQET image measurement based on predefined IQ criteria. This new model brings a new approach for NR image quality assessment to be used for perceived image quality prediction with immediate image scoring.

Objective image quality assessment were primarily designed to quantify the average image quality scores given by experts. The measurements are known to be accurate, but it may take a long time to collect a large number of subjective image quality scores running many HVT, which are typical time consuming and expensive. In addition, objective image quality assessment can't predict the perceived image quality by users which is not the real practice of image quality assessment. In contrast, we can take full advantage of a VIQET for image quality evaluation in order to predict what will be the image quality score given by non-experts viewers. In this case, using VIQET tool to measure the perceived image quality attributes can be a fast, inexpensive to the subjective image quality HVT based approach.

Theoretical importance and practical value of the work. This thesis describes how to extract the standard IQAs to new IQAs, which implemented in the VIQET in order to be able to estimate the expected MOS by HVT. The new model explains the relationship between the standard IQAs values and the VIQET IQAs values. It is also demonstrates how to transform one to other in order to get the most reliable MOS. The calculated MOS predicts the perceived IQ by humans.

The practical value of the new approach based on this model is improving or even replacing the existing IQ assessment procedures in the smartphones industry.

It is out of question that the subjective testing (i.e., human viewers ranks the quality of images) is the most accurate method for perceived image quality assessment. It reflects true human perception. However, these assessments are time consuming and expensive. Furthermore, they cannot be done in real time while using smartphones. Therefore, the goal of this research is to develop an objective image quality assessment method. A computational model based on VIQET that can predict perceived image quality. This correlating well with subjective prediction that is required in the image processing industry. This research proposes a NR objective quality assessment for smartphone images based on VIQET.

Results approval. Both of the objective and subjective results of many HVT experiments and images analysis with VIQET were studied with respect to the statistical analysis. In this context, both descriptive and analytical studies are performed to explore the absolute values and the distribution of numerical data. Based on the experiments results analysis the new proposed model was gradually tested and improved with respect to the proposed novel idea for perceived image quality prediction. To evaluate the performance of the proposed quality assessment model, this study followed the standard performance evaluation procedures of VQEG [6]. The standard was developed for calculating the prediction error between a mathematical model and subjective scores (i.e., human viewers' opinion).

According to the VQEG [6], the performance of an objective quality model is characterized by three prediction attributes: ***accuracy, monotonicity and consistency*.** The motivation was to correlate the objective results generated by the VIQET with subjective results (the outcome of the HVT) in order to evaluate selected image quality metrics, by assuming that the subjective results approximate the actual image quality perception of an overall average observer. In this context, the statistics based psychometric rating and scaling procedures were incorporated to minimize the impact of the variance of judgment criteria between different observers.

Summary of the doctoral thesis sections. In this work, the relationship between subjective perceived image quality assessment responses through HVT from a lab study and objective image quality assessment responses from the VIQET with controlled online of image quality attributes. The results analysis helped to understand the relationship between subjective IQ assessment and IQ assessment by VIQET. Finding the ***relationship between IQ attributes whit the most contribution to perceived IQ in visual experiments***, and IQ attributes parameters for VIQET. This research suggests a new model to maximize the validity of the data obtained and the correspondence with the lab.

This thesis is divided into three chapters.

Chapter I includes the motivation, goals, questions, and methodologies of this research. An overview of the history and definition of image quality, as well as the existing research methodologies regarding objective and subjective experiments, were presented. This chapter provides a review and analysis of various methods for objective and subjective image quality assessment. Find their limitations for not being consistent with human visual perception.

Chapter II presents the experimental study of perceived image quality of smartphones with controlled IQ attributes: brightness, contrast, color saturation and sharpness.

The perceptual results of HVT and VIQET scores were analyzed in order to identify good correlation between the perceived IQ by human's eye and VIQET analysis scores.

During the experimental process, the scores were analyzed and the VIQET was calibrated in order to improve the correlation between both processes. In the end of the experimental process, the VIQET was calibrated for perceived IQ prediction.

Chapter III describes the correlations between the metrical and perceptual results was analyzed and indicated that MOS, MSE, PSNR metrics give excellent prediction performance in most cases in terms of both correlation and its variance.

The statistical analyses and the research outcomes were summarized in this chapter. The outcomes strongly support the increase of the IQ attributes will lead to improvement in user's perceived image quality.

I. ANALYSIS OF SITUATION IN THE SMARTPHONES IMAGE QUALITY

Smartphones with embedded cameras and imaging systems already have billions (2.1 billion in 2016) of users and growth potential is enormous [1]. Most of smartphones have two cameras: rear camera (on the backside of the device) which is the main camera with the highest resolution and front camera (on the display side) with lower resolution. While the smartphone' cameras and displays resolution are increasing in every day the image quality became a major factor for the users. This puts the smartphones vendors into a race on developing and producing higher and higher resolutions.

Therefore, focusing research into technology and user interaction in image quality is increasingly important through combining technical parameters and human perception.

Using HVT as a subjective image quality in order to measure the perceived image quality by humans. Development of VIQET that calibrated according to the HVT scores in order to meet the results expected by humans.

The main goal of this research is to develop a new IQ assessment model using the VIQET evaluation results in order to improve the prediction reliability of perceived image quality by human in order to save time and resources required for physical HVT.

While looking at the IQ circle in Figure 1.1, there is no doubt that the process with HVT is much complicated, cost and time consuming for the smartphones vendors. Therefore, the process done in outsourcing in contract with special organizations experts in public surveys and subjective evaluation procedures.

1.1. Image quality assessment background

People of all generations are making more and more use of digital imaging systems in their daily lives. The image content rendered by these digital imaging systems largely differs in perceived quality depending on the system and its applications.

The idea of the quality of the image started with the invention of the earliest optical instruments, the optical telescope and microscope (1600-1620). Galileo was a key figure in both these inventions [2].

This concept appears again in the early days of photography, 1860-1930, during the development of television, 1935-1955, and continues with digital imaging to the present day. One might assume that with over four centuries of experience with the concept of image quality that we would be close to a complete understanding of the problem. One reason why we are still far from a complete understanding of image quality, and particularly image quality models, is because the perceived image quality is mostly subjective and we lack a structure or a framework. To

address the deficiency a concept called the image quality circle (IQC) was proposed in 1989 at the IS&T annual meeting [1].

With the development of imaging and multimedia technologies, visual information, recorded by images has become the main source for knowledge acquisition. In the process of visual information acquisition, processing, transmission, and storage, some artifacts or noise may be introduced to images, which degrade the visual quality. In a typical digital imaging system, the image is captured and transformed into digital signal by the sensor. This raw digital image signal is then processed to reduce the noise and is compressed for storage or transmission. When the image is finally displayed on the screen to the end user, it might not be same as the original version because it has been exposed to various kinds of distortions.

The sources of distortion could be ranged from motion blurring, Gaussian noise, sensor inadequacy, compression, error during transmission or the combination of many factors. To improve the performance of visual information acquisition, transmission, processing, and storage systems, it is essential to assess visual qualities of the images; so that it can maintain, control and possibly enhance the quality of the image before storage or transmission. The objective of image quality assessment is to provide computational models to measure the perceptual quality of a given image. Recently, a number of techniques have been designed to evaluate the quality of images and videos.

The accurate prediction of quality from an end-user perspective has received increased attention with the growing for compression and communication of digital image and video services over wired and wireless networks. Image quality methods can be categorized in two parts: subjective and objective. The subjective assessment of image is done on the bases of subjective experiments [2]. While objective image quality assessment methods were mainly based on some mathematical measures. The past five years have demonstrated and witnessed the tremendous and imminent demands of visual quality assessment metrics in various applications. Chapter 2 of this thesis describes both of the methods of image quality assessment.

1.2. The concept of image quality assessment

Image quality is not a new term. The earliest history regarding “the quality of an image” can be traced back to the beginning of 17th century, when optical instruments, the telescope and the microscope were invented [2]. At that moment, image quality was no more than an optical concept associated to the acquisition instruments. In recent years, thanks to the rapid advancement of imaging technologies and the tremendous growth in the use of digital media, the scope of image quality has been greatly extended to cover the entire imaging pipeline. For display image quality

assessment, it is important to understand what we going to measure before we actually perform the assessment. Therefore, a clear definition to image quality is required. However, there is no universal and comprehensive definition yet. This is mainly because the term image quality may have significantly different meanings to people from different perspectives with different concerns. In the existing literature, several definitions of image quality have been proposed:

- Jacobson [4] defined image quality as the subjective impression found in the mind of the observer relating to the degree of excellence exhibited by an image.
- Engeldrum [2] interpreted image quality as the integrated set of perceptions of the overall degree of excellence of the image. In his theory of image quality circle, the concept of image quality was associated with customer perceptual rating, customer perception, physical image parameters, and technology variables, of which the image quality assessment components formed a closed loop.
- Janssen [5] followed visual-cognitive processes to define image quality as the degree to which the image was both useful and natural. In this case, the usefulness of an image was defined as the precision of the internal representation of the image; and the naturalness of an image was defined as the degree of correspondence between the internal representation of the image and knowledge of reality as stored in memory.
- Ridder et al. [6] divided image quality into three categories: fidelity, usefulness, and naturalness. Among them, fidelity was referred to the reproduction accuracy of an observed image in comparison to the original, which was assumed to have perfect quality. Usefulness indicated image suitability for the designed task. Naturalness was defined as a match between a reproduced image and the mental impression of an observer, affected by memory traces.
- Fairchild [7] defined image quality as the perceptible visual differences from some ideal and the magnitude of such differences.
- Yendrikhovskij [8] suggested that image quality was understood as the subjective impression of how well image content was rendered or reproduced.
- Keelan [9] defined image quality as an impression of its merit or excellence, as perceived by an observer neither associated with the act of photography, nor closely involved with the subject matter depicted.
- The IIIA (International Imaging Industry Association) [10] defined image quality as the perceptually weighted combination of all visually significant attributes of an image when it was considered in its marketplace or application.

Based on the proposals presented above, it is not difficult to see that image quality is. commonly defined from the subjective perspective. This is mainly because humans are the ultimate visual information interpreter. The visual stimuli are acquired by the human visual system, and the corresponding signals are further decomposed and forwarded along millions of neuron pathways in parallel to the human brain in order to interpret.

Human interpretations are fuzzy in nature. One observer may make his/her own independent judgment regarding the quality of an image even without knowing what image quality actually is. From this point of view, image quality is a subjective and relative term, because one observer may have significantly different perception criteria regarding one or more specific image quality attributes. The underlying attributes are terms of visual perception in the current concerns.

These concerns may vary with respect to the imaging applications and their related contexts. For example, one picture about an extreme sports man climbing rocks was taken with a digital still camera; in this case, the motion blur for the movements, the details of the person's struggling face, and the image resolution for magazine level printing can be the main concerns.

The vivid colors of the background scene may not be because the trees and sky might be completely de-focused in order to feature the climber in the foreground.

Due to the probabilistic nature of the human brain and its high context dependence, ordinary people actually refer image quality as the overall quality of an image reproduction with respect to his/her own perceptual opinion regarding a set of weighted image quality attributes, while scientific researchers may refer image quality as the mean perceptual opinions with respect to one or more visually significant attributes in the concern of current application. In the qualitative approach, image quality can be defined explicitly with well written text statements, but the corresponding computation remains an open question for research discussions.

In an objective manner, ***image quality can be quantized as one or more numbers by applying image quality metrics to the captured images with or without referring to their originals***, which are assumed to have perfect image quality. Depending on the availability of the reference images, the metrics can be broadly divided into three groups: full reference, reduced reference and no reference.

In the full reference approach, image quality is defined as the magnitude of quality degradation from the original image to its reproductions with respect to one or more specific image quality attributes, and all pairs of pixels in the two images are used. In contrast, the reduced reference based metrics count only image features.

The no reference based metrics determine the quality of an image blindly, absolutely without its original. In all these cases, the scales of and the interpretations to the numeric quality indication

are totally different. In other words, image quality is defined implicitly with numbers in independent numeric spaces of such image quality metrics. Thus, the quality of an image is the predicted image quality and it needs to correlate well with the actual image quality in order to be claimed to be valid. In this case, the subjective image quality forms an approximation of the ground truth.

With respect to the discussions above, image quality should be defined as the subjective quality, which is an approximation of the actual image quality. The ***subjective image quality stands for the mean perceptual opinions obtained from the statistical regression analysis of subjective ratings,*** which are sampled from a specific human population with respect to a few visually significant attributes in the concern of current imaging application**.** This definition of image quality is used in the rest of this research.

1.2.1. The image quality circle

The image quality circle [3], which is shown in Figure 1.1, is briefly described. The goal of an imaging system designer is to relate the technology variables of the imaging system or technology to the customer quality preference.

Figure 1.1 shows this fundamental objective via the arrow. The link between customer quality preference and the imaging system and materials technology variables is typically determined by selecting a variable, printing images, and then asking customers to judge the quality of the printed image. This clearly works, but it is inefficient over time because a new data collection effort is required every time a parameter is changed.

The IQC breaks the relationship between technology variables and customer perceptions down into a series of definable and measurable steps. The four elements of the IQC approach are depicted in Figure 1.1 and are described in counter-clockwise order around the circle.

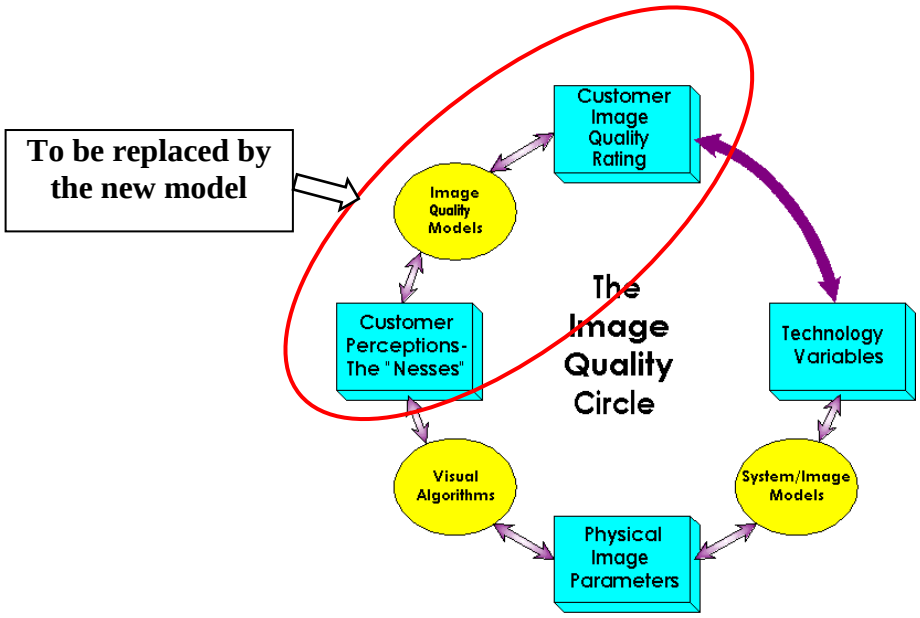


Fig.1.1. The image quality circle [3].

Customer image quality preference is the overall image quality rating as judged by customers. This is an interval scale of overall image quality that can be defined percent, say 0 to 100, or adjectivally, such as “poor” (0 -20), "bad"(20 - 40),”fair” (40-60), "good" (60-80), “very good” (80-90) or "excellent" (90-100).

Customer perceptions, the major customer perceptual attributes of image quality are such dimensions as colorfulness, lightness and sharpness. These are called the "nesses" to emphasize the perceptual, as opposed to the physical nature of these attributes. In this study the “nesses” attributes have been extracted to more reliable IQ attributes to be used by VIQET.

1.3. Smartphones vendors in-house image quality assessment

The flow chart in Figure 1.2 illustrates the metrics of perceptual image quality of smartphones. This process preformed in the vendors IQ labs by image quality experts, then, they perform subjective IQ assessment through visual experiments while the observers are “non-experts” in image quality.

This process was described very well in the “Image quality circle” [3]. While a new IQAs the IQ experts evaluate the IQ in labs through visual examinations in ***FR IQ assessment*** and comparing one image to other [11]. Once the IQ has approved the next phase will be subjective IQ assessment through many visual tests in order to measure the smartphone consumers “taste” of perceived IQ. Image quality requires a systematic assessment approach from both subjective and objective perspectives. In both cases, image quality is characterized based on a set of image quality attributes, which are terms of human perceptions [12]. The ultimate goal is to correlate the objective assessment results with the subjective assessment results, so that we can eventually eliminate the demand of observers.

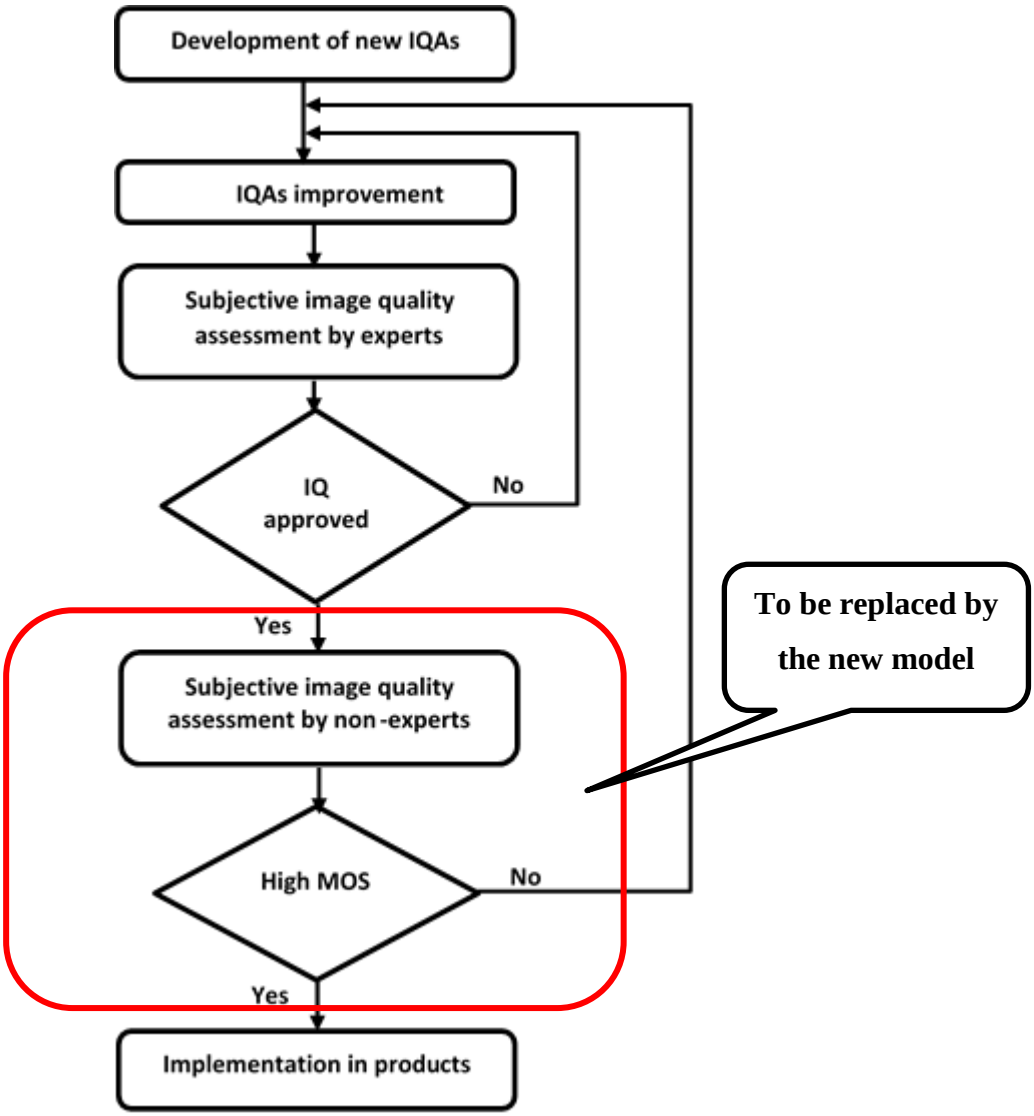


Fig.1.1. New IQAs image quality assessment process at smartphones vendors' labs [13].

Generally, image quality can be separated into two levels: low-level/concrete attributes, which can be measured directly with instruments or estimated based on the measurement results; high-level/abstract attributes, which are abstraction of low-level attributes but they are strongly associated with observer's expertise and experience regarding the underlying image quality attributes, such as naturalness and usefulness. The difference between the two levels is not limited to the abstraction level, but it seems that the importance of high level attributes lies in their ability to inform the observer of the meaning of low-level attributes for the general quality [13].

Among the ***low-level attributes***, physical properties such as screen dimension, display resolution, refreshing rate, have impacts on the perceived image quality, but in a typical work-flow of image quality assessment they can be assumed to be constants, since they are independent from the image content and normally do not vary over time. In this research, we used cameras as the

acquisition devices. For this reason, we only focused on the assessment of ***low-level perceptual image quality attributes***, which were image content dependent.

There is a lot of research related to characterizing electronic devices based on these image quality attributes. These devices were, but not limited to, printers, CRT monitors, LCD/LED monitors, projection displays. According to these studies, the image quality attributes can be generally divided into five groups: lightness, contrast, colorfulness, sharpness and noises. Each group may include several sub-groups with respect to various classification criteria. Most of the image quality attributes were studied with associations to many other image quality attributes.

1.4. Image quality attributes

IQAs that affect the perceived image quality in subjective IQ assessment and in vast use in the industry are: illumination, color saturation, contrast and sharpness.

Illumination

Illumination stands for the perceived intensity of light coming from the image itself, rather than any property of the portrayed scene [6]. It should be used only for non-quantitative reference to physiological sensations and perception of light, so it ranges from “light” to “dark” [12].

Lightness is a close concept to brightness, which is also a perceptual image quality attribute. In this case, lightness is defined as the brightness of an area relative to the brightness of a similarly illuminated area that appears white or highly transmitting. Lightness has a significant impact on the perceptual experience [14].

Color saturation

Color is a human sensation and it represents the perception of incidental light acquired by the human visual system. The accuracy of color reproduction in an image can be represented by the color distance between the image reproduction and its reference in a specific color space. In most cases, when people use the term of color, they actually exclude lightness and refer the term of color to colorfulness, which is a perceptual attribute that covers the aspects of hue, saturation and gamut.

Thus, **color saturation** can be defined as the attribute of a visual perception according to which an area appears to exhibit more of less of its hue. In this context, chroma is defined as the colorfulness of an area judged in proportion to the brightness of a similarly illuminated area that appears to be white or highly transmitting, while saturation is defined as the colorfulness of an area judged in proportion to its brightness [15].

In addition, hue is defined as an attribute of a visual perception according to which an area appears to be similar to one of the colors, red, yellow, green, and blue, or to a combination of

adjacent pairs of these colors considered in a closed ring [16]. For electronic devices, such as scanner, printer, and displays, the color gamut stands for the entire range of colors that the device can reproduce accurately in a specific color space. The color gamut is expected to be as large as possible, but none of the known devices can reproduce all colors [17]. Nevertheless, the most pleasing color might not necessarily be the most accurate color [18].

Contrast

In most literature, contrast for an image was defined as a measure of the luminance variation relative to the average luminance in the surrounding region, however no standard definition to contrast in a complex scene has been given. One most widely adopted definition for measuring image contrast is Michelson formula [19]:

$$C_M = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{1.1}$$

where I_{max} and I_{min} stand for the maximum and minimum value of lighting respectively. Another widely adopted contrast definition is the Weber fraction specially defined for simple test patterns [20]:

$$C_W = \frac{I_s - I_b}{I_b} \tag{1.2}$$

Where I_s and I_b stand for the foreground and background lightness respectively. In the research domain of tone reproduction, contrast is defined as the rate of change of the relative luminance of image elements of a reproduction, as a function of the relative luminance of the same image elements of the original image; on log-log coordinates, contrast is the slope of the relationship between the reproduction and original [20].

In the cases for color complex scene, we may define contrast approximately as a measurement of the luminance and/or chromatic variations in one region relative to the average variance in the surrounding region in the same scene. There are two important aspects in the contrast research. One of them is related to the contrast sensitivity function in Figure 1.3. DeValois et al. [21] indicated that the contrast sensitivity of human visual system followed a certain curve with respect to the current average luminance level, and the spatial frequency of luminance variations.

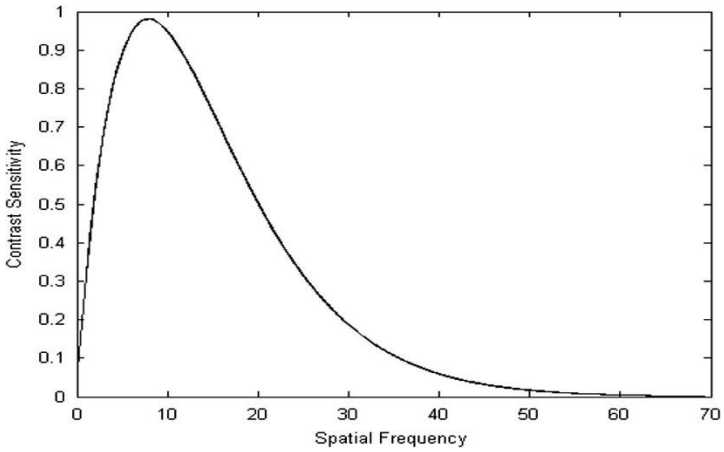


Fig.1.1. An illustration of contrast sensitivity function [22].

Thus, an optimization for the image content or size can be achieved accordingly by keeping the high spatial frequency components in the images, while the low frequency ones are being eliminated. The other research aspect of contrast is related to contrast masking, which is a visual phenomenon of human visual system.

The term is used commonly to refer to any destructive interaction or interference among transient stimuli that are closely coupled in space or time [23]. Thus, the masked signal shows different visual effect under the different contrast masking signal [24]. This effect is modeled either with a threshold elevation image, or with a contrast transducer function calculated from the masking curve of contrast discrimination experiments, given that the image is decomposed into the appropriate spatial frequency bands [25].

Sharpness is an attribute defining how abrupt the boundaries are between different tones and colors [26, 27]. It is commonly recognized to be an important image quality attribute for perceptual evaluation despite the technology used, and it is closely associated with other attributes, such as lightness, contrast, and blur. Since sharpness defines the amount of details the human can observe in image reproductions at a certain distance, it is commonly referred to as the counterpart of blur. The human visual system has a remarkable capability to detect image blur without seeing the original image, but unfortunately the underlying mechanism is not well understood.

One way to determine sharpness is measuring the rise distance of the slant edges, or calculate the density of line pairs with increasing spatial frequency, or do the corresponding analysis in the frequency domain, where frequency is measured in cycles or in line pairs per distance (millimeters, inches, pixels or degree). Specifically, the International Organization for Standardization defined ISO 12233 to standardize the procedure of measuring the resolution and spatial frequency responses of camera lens with a special test chart. The existing research regarding sharpness is

largely focused on the design and evaluation of reduced reference based and no reference based image quality metrics.

Noise is a random variation of image density, visible as grain in film and pixel level variations in digital images. It is a key image quality factor, nearly as important as sharpness.

In most cases noise is perceived as a degradation in quality. But some Black & White photographers like its graphic effect: Many favored 35mm Tri-X film. (Film grain has very different statistics from digital noise— it's multiplicative rather than additive, and its spectrum is dependent on density.) Pointillist painters, most notably George Seurat, created "noise" (specks of color) by hand; a task that can be accomplished in seconds today with Photoshop plugins. But by and large, the majority of photographers, especially color and large-format photographers, dislike noise with good reason.

Noise, such as speckles, spikes, reseals, missing data, marks, blemishes, banding and abnormalities, are created either in expectation or unexpectedly during the processes of acquisition, transmission, and processing of image data [28]. Similarly, artifact is a range of errors in the perception or representation of any visual information introduced to an image in the processes, such as optical acquisition, digital sampling, image compressing and signal processing [29, 30].

1.4.1. Relationships between image quality attributes

Many types of ***perceptual image quality attributes*** were introduced in the previous sections, however they are not completely independent from each other. One image quality attribute may have connections to one or more other attributes. For example, the estimation of illumination is based on the measurements of luminance, which form a foundation for the studies of many other image quality attributes. For example, in the study of perceptual contrast of projection displays, Majumder et al. [31] emphasized that luminance is more important for perception than chrominance.

In the White's illusion phenomenon, the relationship between the lightness of two gray regions was revealed to be the opposite of what is predicted by local edge ratios or contrasts. In other studies [32, 33, 34], Illumination was also integrated into the computation of image contrast. In the study of digital printing, the banding and contouring artifacts were found to have connections with lightness. Ridder [35] studied the naturalness of images with respect to the saturation and lightness variations. It was found that the difference between naturalness and quality diminished with decreasing lightness.

In addition, the evaluation of contrast attribute has a strong connection to the measurement of Illumination, as well as colorfulness, sharpness/blur, and artifacts as well. Several studies [36, 37, 38, 39] regarding contrast sensitivity of human eye and its effects on image quality were presented. In addition, several studies [40, 41] for determining structural similarity or degradation based on contrast measurements were presented. In these contrast studies focus on modeling, both luminance and chrominance attributes were used.

Regarding the research of colors, a huge amount of effort has been expended. One good example is the research related to color appearance modeling [42, 43, 44]. One color appearance model includes predictors of at least the relative color appearance attributes, such as lightness, Chroma and hue [45], and it can be used to predict image quality. Such a model addresses the perspectives of presented stimuli, viewing condition, colorimetry, color appearance phenomena, and chromatic adaption etc. Color appearance models incorporate chromatic adaptations as well as the predictors of brightness and colorfulness.

They also adopt the color adaptation model as a module in the initial step, so this module can be selected or replaced in preference. In the post-adaption step, adapted tri-stimulus data and other additional data, like absolute luminance level, colorimetric data on the proximal field, background and surround, are combined to provide higher level signals in order to produce predictors of color appearance attributes.

1.4.2. Summary of image quality attributes

Image quality can be characterized based on perceptual attributes from various perspectives, however the selection of the most important image quality attributes has different priorities in different research domains. For digital printing, the research [12] suggested that lightness, contrast, sharpness, artifacts, colors, and physical attributes are all important. Lindberg [47] evaluated many image quality attributes, such as color gamut, sharpness, contrast, tone quality, detail highlights, detail shadow, gloss level, gloss variation, color shift, patchiness, mottle, and ordered noise.

Among them, the print mottle and color gamut were found to account for most of the variations with respect to the factor analysis. Johnson [12] specially remarked colorfulness, sharpness, and contrast for printing. For mobile displays, Gong et al. [48] suggested that clearness was the most important, followed by naturalness, sharpness, colorfulness, contrast and brightness. In this context, clearness is a high-level attribute associated with other low-level attributes; however, the actual numeric relationship was not given in the research.

In contrary, Kim et al. [49] emphasized that naturalness had a high priority than clearness related to overall image quality. Thomas et al. [50] and Strand et al. [51] remarked lightness and

colorfulness for projection displays, while Majumder et al. [54] indicated that lightness is more important than the colorfulness. In this research, with respect to the literature, we can see that the image quality has strong connections to contrast and sharpness image quality attributes despite the actual display technology used. In addition, it was known that for projection displays spatial uniformity is an important image quality attribute [54]. Hence, in this research, we pay special attentions to contrast, sharpness and spatial uniformity by utilizing our proposed image quality assessment framework.

1.5. Subjective image quality assessment

Subjective image quality assessment has fundamental importance to the design and validation of objective metrics [55]. It provides a better understanding of how quality is assessed by the HVS and this understanding greatly helps for mapping objective quality prediction to subjective quality experience. In addition, quality scores resulting from subjective experiments are widely accepted as the benchmark for evaluating the performance of an objective metric and for comparing alternative metrics proposed in the literature.

To obtain useful and reliable results from subjective experiments, it is necessary to design an experimental protocol that best fits the goal of the image quality assessment problem at hand. the protocol is presented by the "International Telecommunication Union" in ITU-R Recommendation BT.500-11, Methodology for the subjective assessment of the quality of television pictures [55] and by VQEG in final report from the video quality experts group on the validation of objective models of video quality assessment [56]. In this protocol aspects related to viewing conditions test material and test methods have to be discussed and selected. Typical issues regarding subjective tests are documented in [56] and [55]. In recent years, the dramatic increase in research on objective quality metrics has pushed the need for public, freely available, databases of images/ videos and their corresponding subjective quality scores to the forefront.

Having these databases largely facilitates the development of new objective metrics and their performance evaluation in a comparative setting with existing metrics. A direct comparison on the same content and quality scores allows an analysis of the strengths and weaknesses of all metrics available. Some of the databases are summarized in "Image and Video Quality Resources" [57].

1.6. Objective image quality assessment

The most well-known and widely used objective metric is MSE/PSNR. It is a FR metric that simply sums all pixel-by-pixel differences between a distorted image and its original version. The

metric is parameter free, and very inexpensive to implement, but it is widely criticized by the image quality community for its poor correlation with human perceived image quality [2].

MSE/PSNR, ironically, remains the most used quality metric in current signal and image processing systems, mainly because it is a convention. Researchers have taken different approaches to develop FR metrics with a better performance than MSE/PSNR, mainly by including aspects of the HVS. In order to be able to do so, functional aspects of the HVS needed to be modeled. Advances in human vision research increased our understanding of the structural and functional mechanisms of the HVS and allowed expressing these psychophysical findings into mathematical models [59, 60]. Although these models still remain limited in their sophistication and thus also in their reliability, they are already of great interest to explore their added value in image quality research.

One way to integrate HVS aspects in the design of an objective quality metric is defined rather ‘‘bottom-up’’ and simulates well-known functionalities of the early HVS [61]. The metrics of which numerous different implementations are discussed in literature are based on a so-called error-visibility framework [62]. This framework decomposes the image signal into channels of various frequencies and orientations in order to reflect human vision at the neural cell level.

Classical HVS models such as the contrast sensitivity function (CSF) per channel and interactions between these channels to simulate masking are then implemented. Pioneering work on this approach is described in [23] and more representative models that are consistent with the error- visibility framework are summarized in [62]. Although well studied there still are some limitations to these metrics. First, our knowledge of the HVS is far from complete and simulating precisely all related components of the HVS is impossible.

This intrinsically limits the accuracy of these metrics. Second, the HVS is a rather complex system that contains many nonlinear operations. But, most existing vision models are linear (or quasi-linear) and are developed using restricted and simplistic stimuli. Applying these vision models in objective metric design actually implies the acceptance of a number of strong assumptions. There are two recent and very successful alternatives to achieve a reliable FR metric. They both are based on a higher level ‘‘top-down’’ approach of the overall functionality of the HVS. It concerns the ‘‘structural similarity’’ (SSIM) [51] and the ‘‘visual information fidelity’’ (VIF) [52] metrics.

The principal idea behind SSIM is the observation that the HVS is highly adapted to extract structural information from visual scenes. Therefore, the metric intends to quantify image quality by measuring the structural similarity (or distortion) between a distorted image and its original version. SSIM defines nonstructural distortions as those that do not modify the structure of objects

in the visual scene, whereas all other distortions are defined as structural. The metric measures the similarity in three elements, i.e. luminance, contrast and structure within a local area of image content.

The design of VIF is based on an information communication and sharing point of view. It attempts to relate image quality to the amount of information that is shared between the distorted image and its original version. In other words, VIF exploits the relationship between statistical image information and Image Quality. It has been shown that both SSIM and VIF are much more consistent than MSE/ PSNR in predicting perceived Image Quality. A comprehensive evaluation of the performance of SSIM, VIF as well as other recent FR metrics is detailed in [63].

In a **full-Reference Method**, reference images are required to assess the quality of distorted images. Most of the quality assessment methods belong to this category. However, in most of the real-time applications reference images are unavailable so a FR metric is not useful for those applications.

In a **reduced-Reference Method** [39], partial information of the reference images is required to predict the visual quality of distorted images. This method is used as a substitute of FR method in case of missing an ideal FR image. For example, while developing new IQ feature, it is compared to the original image without the new IQ feature in order to see if there is an improvement in IQ or not.

No-Reference Method. The method that does not require any reference to assess the quality of a distorted image is known as a NR method [39]. In many real-time applications, the reference image is not available, and a *no-reference* or "blind" quality assessment approach is desirable. Moreover, a human observer can assess the quality of a distorted image without the use of any reference image. However, developing an objective NR quality measurement metric is difficult due to the limited knowledge of HVS. Therefore, a NR quality assessment is reliable only when the prior knowledge about the image distortion is available [31]. This research focus on this method.

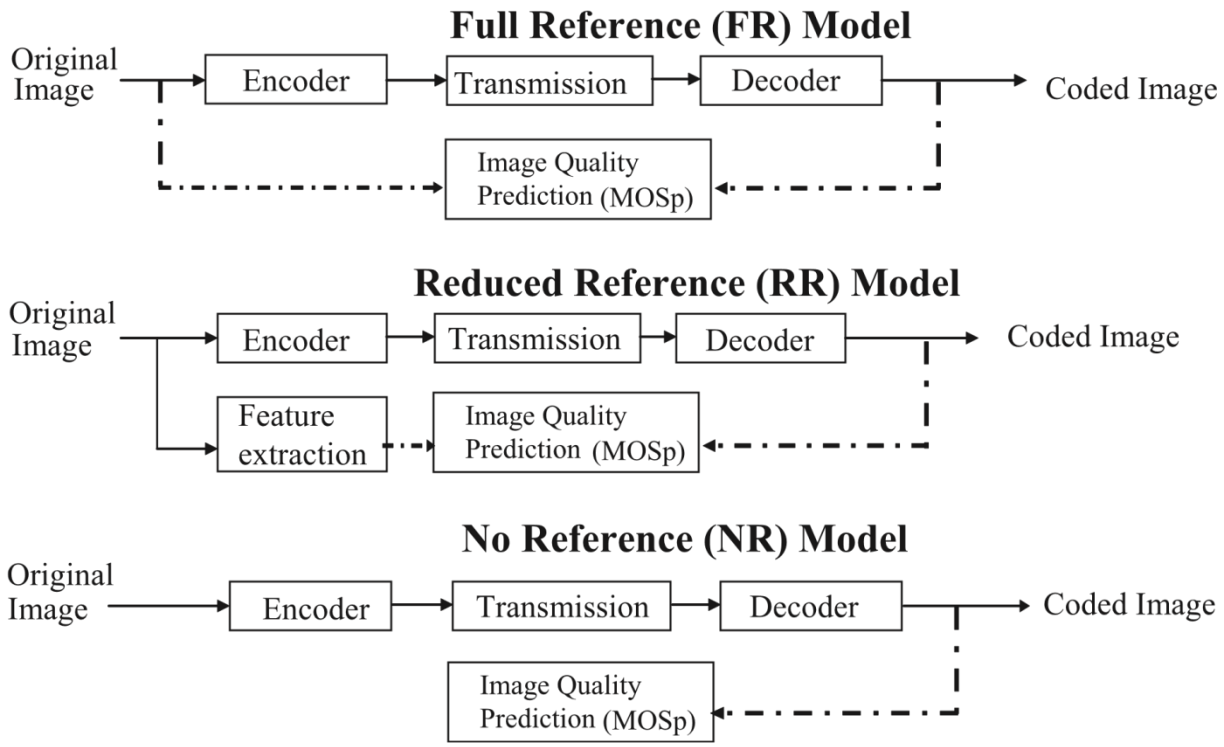


Fig.1.2. Image quality assessment models [2].

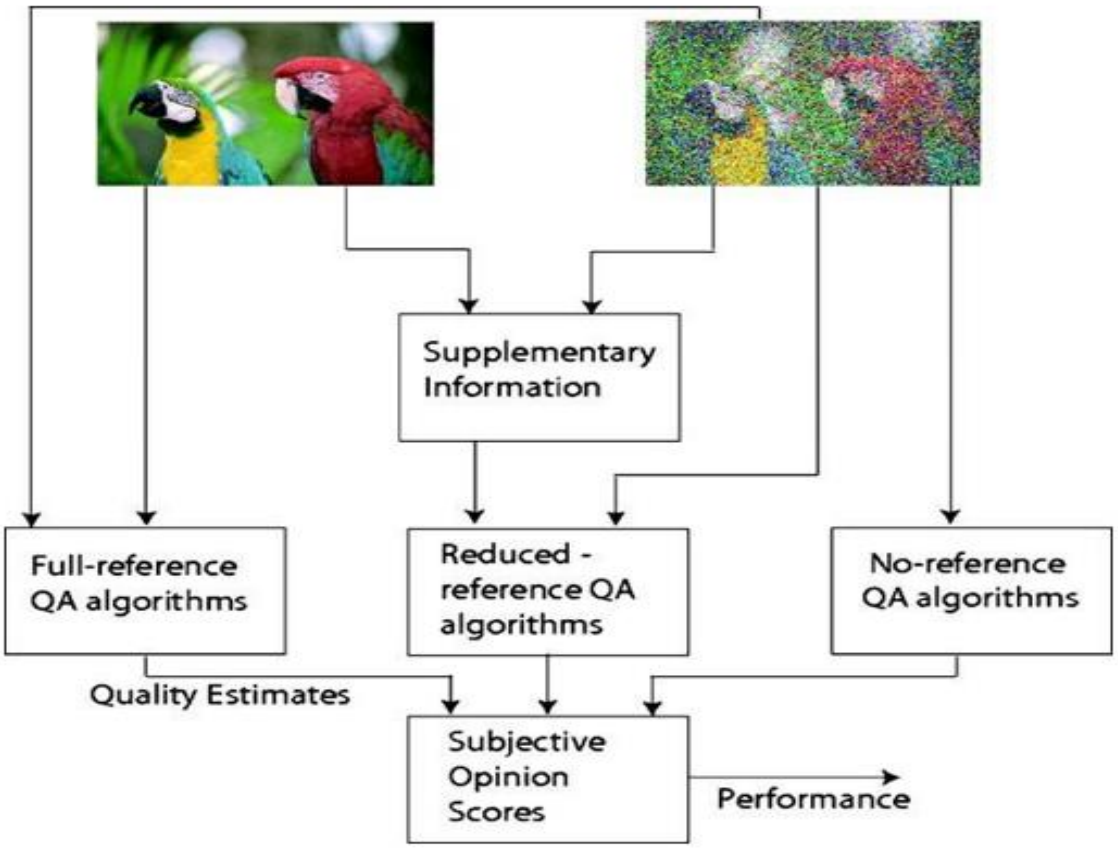


Fig.1.3. Flow chart of subjective image quality assessment [2].

1.7. The challenge of non-reference image quality assessment

Compared to the research on **FR metrics** that on **NR metrics** is still in a very preliminary stage. Nonetheless, research on NR metrics has recently received a lot of attention, because of their great practical potential in real-time applications. Assessing quality based on the distorted image only seems an easy task for human observers, yet it is the most difficult problem in objective image quality metric design [63].

Fortunately, in many practical applications, the processes involved in generating the distortions are known and fixed and so the design of a NR metric that handles a specific distortion type turns out to be much more realistic. Based on this idea, NR metrics can be categorized into ***general metrics and dedicated metrics.***

General NR metrics are intended to assess the overall perceived quality of an image degraded by a known distortion process which possibly contains various artifact types, e.g. a wavelet-based compressed image often exhibits blur and ringing artifacts simultaneously [64].

Dedicated NR metrics instead are based on directly measuring a specific artifact type created by a specific image distortion process such as blur caused by image acquisition or blocking artifacts resulting from block- based DCT coding [64]. In the design of a general NR metric the overall quality of specifically degraded images is often targeted using hypothesized assumptions about natural scenes or the HVS. The NR approach proposed [2] relies on the assumption that images of natural scenes exhibit strong statistical regularities and therefore, reside in a tiny area of the space containing all possible images. As a consequence, it quantifies the overall quality of images compressed by JPEG2000 based on detected variations in the statistics of image features calculated in the wavelet domain.

The performance of this approach however, largely depends on sophisticated modeling of natural scene statistics. As an alternative, NR image quality assessment is formulated as a machine learning problem in some research (such as e.g. in [67-69]). It avoids the explicit modeling of the HVS, but rather treats it as a black box, whose input-output relationship between image characteristics and a quality rating is to be learned by computational intelligent tools, such as neural networks. This type of NR metrics is generally defined as a regression or function approximation and therefore, usually requires extensive training on a large data set obtained from subjective quality rating experiments.

The two types of general NR metrics mentioned so far have been proved to be effective for the overall quality prediction of a specific combination of distortions, but they are unlikely to be able to handle other combinations of distortions. For example, a NR metric based on a neural

network that is trained to assess the quality of JPEG compressed images is not necessarily useful to predict the perceived quality of JPEG2000 compressed images. In the literature, a large number of NR metrics are designed to assess the quality degradation of a specific type of artifact, such as blockiness, ringing or noise. These dedicated NR metrics are highly beneficial for image/ video compression and transmission systems.

First, they usually provide a spatially varying quality degradation profile of a distorted image, indicating at each location in the image the visibility of the targeted type of artifact. Second, these metrics can each individually determine the quality degradation caused by a specific type of artifact, e.g. the annoyance of blockiness and ringing can be quantified simultaneously and separately for each JPEG compressed image.

Both aspects contribute to the optimization of signal enhancement at either local or global level in an imaging chain. For example, in the video chain of current television sets the artifact reduction scheme uses these metrics to quantify the occurrence of individual artifacts in the incoming video and automatically adjusts the algorithms and their parameter settings accordingly [70-73].

FR metrics cannot be applied since there is no access to the original video data at the receiving side. On the other hand, NR metrics have limited reliability since the type of distortions occurring in complex communication networks can be insufficiently predicted.

A RR approach provides a practical solution, it only sends partial information about the reference as additional data from the transmitter to the receiver. Obviously, the bandwidth needed for sending the additional information becomes a crucial aspect in the metric design. So far, advances in image quality assessment have shown the need and practical attainability of integrating relevant aspects of the HVS in objective metric design.

In the literature, lower level aspects of the HVS, such as contrast sensitivity, luminance masking and texture masking, are successfully modeled and integrated in various metrics. Studies evaluating whether also higher level aspects of the HVS, such as visual attention, are beneficial for objective quality prediction, and if so, how to apply them in metric design are still limited, but recently have emerged as an active research area [74-78].

1.8. Research on perceived image quality arena

As the smartphones with embedded camera became an essential part of our everyday life style. Image transportation is in extensive usage in the social networks (e.g. Facebook and Instagram).

The smartphones' camera and display frequently getting higher megapixels resolutions, the research on perceived image quality became a very "Hot" subject for image processing researchers and smartphones vendors.

In the recent 2-3 years, several researches on perceived image quality have been conducted. These researches were focused on mobile devices.

An overview on several of the most recent researches: ***Image quality evaluation for smart-phone displays at lighting levels of indoor and outdoor conditions*** [98]:

The image quality of two active matrix organic light emitting diode (AMOLED) smart-phone displays and two in-plane switching (IPS) ones was visually assessed at two levels of ambient lighting conditions corresponding to indoor and outdoor applications, respectively. Naturalness, colorfulness, brightness, contrast, sharpness, and overall image quality were evaluated via human visual experiment using test images selected from different scenes. The experimental results show that the AMOLED displays perform better on colorfulness because of their wide color gamut, while the high pixel resolution and high peak luminance of the IPS panels help the perception of brightness, contrast, and sharpness. Further statistical analysis indicates that ambient lighting levels have significant influences on the attributes of brightness and contrast

This study focused on revealing the critical factors that seriously affect IQ, as well as the correlations between the physical parameters and the perceptual IQ attributes. Four smart-phone displays from two types of prevailing display technologies were employed because their physical parameters. Five important perceptual attributes were chosen for the subjective IQ evaluation: naturalness, colorfulness, brightness, contrast and sharpness.

Meanwhile, smartphones are being used under indoor and outdoor environments, and the displayed images are perceived differently due to the different ambient lighting conditions. Therefore, the investigation was also aimed at finding out the impact of ambient lighting on the evaluated IQ attributes by adopting two lighting levels to simulate the actual applications of smartphones in indoor and outdoor environments, respectively.

Influence of ambient lighting levels on image quality **[**98]

Image quality evaluation for smartphone displays at lighting levels significantly influence the assessment results and that there are significant differences between the performances of IPS and AMOLED panels. As for the ambient lighting effect, only the brightness and contrast are influenced significantly, since the change of ambient lighting level has a direct impact on these two attributes. Consequently, maintaining an appreciable performance of brightness and contrast under very high lighting conditions is of significant importance for smartphones' outdoor

applications. The IQ performances of two AMOLED displays and two IPS have been visually evaluated and analyzed at two lighting levels of indoor and outdoor conditions.

The psychophysical experiment demonstrates that high pixel resolution and high peak luminance help the perception of brightness, contrast, and sharpness, especially for the IPS-2, while the AMOLED displays achieve better performance on colorfulness because of their wider color gamut.

Analysis of the results indicates that there are significant differences in IQ between IPS and AMOLED displays, and the ambient lighting levels have significant impact on the attributes of brightness and contrast.

Perceived image quality on mobile phones with different screen resolution [99]:

The diverse display screen imposes significant challenges for assessing the perceptual image quality across different mobile devices. In this study, the perceived image quality on different mobile phones is investigated. Firstly, subjective experiments for image quality evaluation are implemented on 9 popular mobile phones and a broadcast-quality monitor to evaluate the impact on perceived image quality regarding the screen resolution, screen size, image resolution, and image coding quality. Furthermore, the effect of mutual interaction between the image resolution and screen resolution is analyzed and an integrated assessment parameter is proposed to establish a device-dependent image quality assessment model.

In practice, the mobile device significantly influences the user's perceived quality in terms of the screen resolution, screen size, and devices type.

Subjective Experiment, two subjective experiments have been conducted to investigate whether the user's viewing experience will be significantly improved by enhancing the screen resolution and how the image resolution will affect the user's perceived image quality on different mobile devices. A total of 9 popular mobile devices were chosen as the test devices in the experiments, that is, P1 to P9, and a monitor (M1) was used as the benchmark. A uniform image browser, named Tidy, was used to display the images on each mobile phone. *Test Material* - A total of ten 4K (3840 × 2160) resolution color images were selected.

Conclusions

In this research, a wide range of popular mobile phones are selected in the subjective experiments to investigate the impact of image resolution, screen size, and screen resolution on user's perceived image quality. The quantitative and statistical analyses are conducted to check whether the increase of the screen size and resolution will 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 screen resolution. Furthermore, a device-dependent image

quality assessment model is proposed to evaluate the perceived image quality on different mobile phones. *The proposed quality assessment model is useful for image quality assessment on specific mobile phones.*

Perceived image quality assessment for color images on mobile displays [100]:

This theis presents a quantitative method to evaluate perceived image quality of color images on mobile displays. Three image quality attributes, colorfulness, contrast and brightness, are chosen to represent perceived image quality. Image quality assessment models are constructed based on results of human visual experiments. In this paper, three phase human visual experiments are designed to achieve credible outcomes while reducing time and resources needed for visual experiments. Values of parameters of image quality assessment models are estimated based on results from human visual experiments. Performances of different image quality assessment models are compared.

Relationship between single visual attribute and perceived image quality could be assumed to obey a reversed U-shaped curve. Figure 1.6. illustrates an example of such a relationship. Horizontal axis represents a level of contrast. Vertical axis denotes perceived image quality. When contrast of an image is enhanced, its perceived image quality would be also improved. But, when excessive enhancement is applied, perceived image quality would be worse.

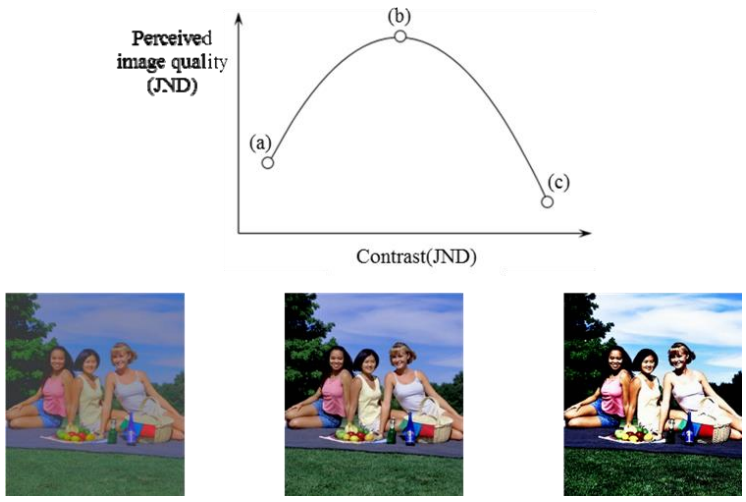


Fig.1.4. Different contrast levels [100].

The goal of this thesis is to develop a new no (or zero) reference (NR) subjective quality assessment model and framework that would enable the smartphones industry prediction of the perceived image quality by smartphone users.

This work includes the design of NR subjective and objective image quality metrics based on extensive HVTs and experiments in order to evaluate and predicted the perceived IQ with software tool for image analysis VIQET. This includes a new image quality assessment procedure based on VIQET tool that measure the perceived image quality by smartphones users.

The measurement of perceived image quality is based on the standard IQAs (brightness, contrast, colors and sharpness). These standard IQAs were extracted to the VIQET IQ criteria in order to be able to quantitatively predict the perceived image quality.

The research objectives:

- Situation analysis of the existing IQ assessment methods;
- Find the most effective IQAs that make the image more "pleasant" to smartphones users;
- Investigate the relationship between standard IQAs and VIQET IQ criteria;
- Develop new method to extract the standard IQAs to VIQET IQ criteria;
- Improve the new procedure and framework according to comparative analysis with HVTs results;
- Evaluate the performance of the new suggested process using the VQEG recommendations;
- Develop a reliable and efficient process for prediction of smartphones perceived IQ;

At baseline, the research examines how the smartphone users perceive the image quality of their devices via visual tests in NR subjective image quality assessment. Scores given by the observers have been analyzed and compared with VIQET scores. As result of the analysis of the HVT and VIQET scores, the VIQET image quality attributes coefficients were recalculated and modified in the tool as essential part of the new model development. This procedure has been ran until the highest correlation between HVT scores and VIQET scores have been achieved.

The research problem is elaborating a new diploid model for PIQ assessment that describes how to extract the standard IQAs, which used in large screens TVs into new IQAs of SW tool VIQET. This can be used by smartphone producers and vendors in order to shorten "time to market", or by any image quality experts, including in the academia in order to reduce time and cost of PIQ assessment process of smartphones with small high definition (HD) displays in comparison to the process that based on many human physical tests.

Research problem solving, firstly, existing objective and subjective image quality assessment metrics have been reviewed and discussed their limitations for not being consistent with human visual perception. Also, review of the methods that consider the human visual characteristics and discussed their limitations. The relationship of image quality features and human visual system was investigated through human visual experiments. In theses experiments the IQAs were changed (increased and decreased) in order to analyze the degradation of images for different levels of IQ attributes and various contents of images.

The outcome of these experiments is to develop a new metric for image quality assessment based on image quality attributes: brightness, contrast, color saturation and sharpness. We found that the brightness, contrast, color saturation and sharpness are the major attributes for perceived image quality. The finding of these IQ attributes present a good starting point to describe overall perceived IQ and considered as a step towards achieving a link between objective and subjective IQ.

This thesis describes how to extract the standard IQAs to new IQAs, which implemented in the VIQET in order to be able to estimate the expected MOS by HVT. The new approach explains the relationship between the standard IQAs values and the VIQET IQAs values. It is also demonstrates how to transform one to other in order to get the most reliable MOS. The calculated MOS predicts the perceived IQ by humans.

The outcome of this research problem solving is elaborating a new diploid procedure and framework for perceived IQ assessment that describes how to extract the standard IQAs, which used in large screens TVs into new IQAs of SW tool VIQET. This can be used by smartphone producers and vendors in order to shorten “time to market”, or by any image quality experts, including in the academia in order to reduce time and cost of PIQ assessment process of smartphones with small high definition (HD) displays in comparison to the process that based on many human physical tests.

This research provides a complete solution, which consists of framework and software tool. The new framework and software image quality analysis enable researches and IQ engineers to perform a real time assessment of perceived IQ of smartphones.

Conclusions

In this research, four different image quality assessment models are constructed for three natural color images on mobile displays. Three image quality attributes, colorfulness, contrast and brightness, are chosen to represent perceived image quality. Image quality assessment models are constructed based on results of human visual experiments. Values of parameters of image quality

assessment models are estimated based on results from human visual experiments. Performances of different image quality assessment models are evaluated with testing samples that are not utilized for model construction. Three images result in different image quality assessment models.

However, they can be utilized to evaluate effectiveness or performance of image enhancement techniques. Suppose that an image enhancement technique is developed and applied to an image having perceived image quality assessment model. From resulting image, quantitative measures of visual attributes that are directly proportional to the just noticed distortion (JND) values from human visual experiment can be calculated. Then, change in perceived image quality can be estimated by perceived image quality assessment model.

1.9. Summary

As image quality assessment plays an important role in various image-processing applications. This chapter provides a description and analysis of:

- Various methods for objective and subjective image quality assessment;
- Several objective and subjective image quality assessment metrics have been reviewed and discussed their limitations for not being consistent with human visual perception;
- Review of the methods that consider the human visual characteristics and discussed their limitations;
- Verify the relationship of image quality features and human visual system, several experiments have been ran in order to analyze the degradation of images for different levels of IQ attributes and various contents of images;
- By considering the limitations of these existing methods, proposing a new metric for image quality assessment based on image quality attributes: brightness, contrast, color saturation and sharpness;
- Found that the brightness, contrast, color saturation and sharpness are the major attributes for perceived image quality. This approach also lead to better image quality prediction accuracy;
- Existing IQ attributes were identified and categorized. The outcome of this is a proposal of a refined set of selection of the most meaningful IQ attributes for the evaluation of perceived IQ of mobile phones;
- The number of IQ attributes considered important in IQ evaluation has been reduced to a set of four IQ attributes: brightness, contrast, color and sharpness;

- IQ attributes present a good starting point to describe overall perceived IQ and considered as a step towards achieving a link between objective and subjective IQ;
- Reviewing the methods that consider the human visual characteristics and understanding their limitations in order to verify the relationship of image features and HVS;
- By considering the limitations of these existing methods, proposing a new metric for quality assessment based on IQ attributes within the image;

II. PERCEIVED IMAGE QUALITY ASSESSMENT

The second chapter provides a comprehensive overview on the perceived image quality measurements and VIQET parameters calibration process. The work was divided into two phases:

Phase I: Through HVTs with observers in each experiment (subjective image quality assessment), the most contributing image quality attributes to the perceived IQ and the optimal parameters level of each attribute were identified.

Phase II: Creating a set of processed images based on the selected image quality attributes to be used as test content for HVT. By performing HVTs and analyzing the same images with the VIQET. The scores of HVTs and VIQET were analyzed. Then VIQET calibrated due to the outcomes of the scores analysis. Once the VIQET has new image quality parameters a new HVT conducted and the whole process done again. This process was performed in cycles with observers. The objective was to achieve the highest correlation between the HVT scores and VIQET scores.

Once the VIQET scores were very close to the HVT scores, the final perceived image quality assessment conducted with forty images as test content and 35 observers scored the test content, then the same images were analyzed by the VIQET.

As a final VIQET new parameters validation, a quick IQ assessment performed in the counter wise process, analyzing new images first by VIQET than in HVT. The outcomes showed a very high correlation between the HVTs scores and VIQET scores.

This work went through the following process: as a first step, the IQ attributes must be identified, this was done by a survey of the existing literature in chapter 1, numerous of IQ attributes have been considered as important and evaluated by researchers to quantify IQ. These QAs include for example brightness, sharpness, contrast, noise/graininess, banding, details, naturalness, color , saturation, color rendition, process color gamut, artifacts, color reproduction , tone reproduction, color shift.

When reducing these ***IQ attributes***, which found in the literature, there are several important issues to consider, such as the intention of how IQ attributes should be used, and their origin. With this intention, the IQ attributes should be based on perception and account for technological IQ issues. The IQ attributes should be general enough to be evaluated by observers, and in order not to exclude novice observers the IQ attributes should be somewhat straightforward to evaluate. In addition, the IQ attributes should be suitable for IQ metrics. The existing sets of IQ attributes and models do not fulfill all of these requirements, and therefore a new set of IQ attributes is needed.

Many of the IQ attributes listed above are similar and have common denominators, which enables them to be grouped within more general IQ attributes in order to reduce the dimensionality and create a more manageable evaluation of IQ. There is usually a compromise between generality and accuracy when it comes to dimensionality. Linking most of the above IQ attributes to four different dimensions, considered as important for the evaluation of IQ. This results in a reasonable compromise between accuracy and complexity, as well as being close to the statement by Engeldrum [97] that observers will not perceive more than five IQ attributes simultaneously. IQ attributes found in the literature were reduced to the following four:

- **Brightness** is considered so perceptually important that it is beneficial to separate it from the color. Brightness will range from "light" to" dark".
- **Contrast** can be described as the perceived magnitude of visually meaningful differences, global and local, in lightness and chromaticity within the image.
- **Color** contains aspects related to color, such as hue, saturation, and color rendition, except lightness.
- **Sharpness** is related to the clarity of details and definition of edges.

2.1. Image quality attributes of VIQET

Through a number of experiments and evaluations and research [4] it is apparent that the following factors greatly influence the perception of good image quality:

Resolution: Refers to the total number of little squares (pixels) that make up an image. The higher the resolution, the higher the number of pixels.

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.

Color Warmth: Refers to the tint of the overall image. Images with a bluish tint are considered to have cool colors. Images with an orange - reddish tint are considered to have warm colors.

Dynamic Range: Is the range between the lightest and darkest regions in an image maintaining details of an image in both the lightest and darkest spots (represented in shades of grey).

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.

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.

Multi-scale texture acutance: Refers to the level of activity and detail in the scene. The higher the level of detail in the scene the higher the feature value.

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.

% over exposed: Refers to the percentage of the image area that is covered in white. This feature ranges from 0 to 100. A higher percentage value indicates a larger area of an image is over---exposed.

% Under---exposed: Refers to the percentage of the image area that is covered in black. This feature ranges from 0 to 100. A higher percentage value indicates a larger area of an image is under-exposed.

Lux: Refers to a measure of light intensity that hits or passes through a surface, as perceived by the human eye.

In the past decades, a number of objective image quality assessment and video quality assessment algorithms have been proposed to evaluate the image/video quality, as summarized in [86, 81, 87, 88]. However, these traditional works only focused on evaluating the quality of image and video sources, losing sight of the effect from the display or the specific mobile device. Although some literatures have been involved with subjective experiments on mobile devices for media quality evaluation [89–92, 83], they still addressed the characteristics of image or video sources or only focused on the impact of the assessment technologies. Moreover, the influence of usage location of mobile devices has been also investigated in [79, 80, 81–83]. In practice, the mobile device significantly influences the user's perceived quality in terms of the screen resolution, screen size, and devices type.

2.2. The image quality assessment method

A large number of subjective metrics have been developed that are capable of mimicking subjective IQ assessment to measure the visible differences between a pair of images [84 - 86]. Considering this wide range of applications, in this research separation of the objective research into two main categories: first, the methods that consider statistical or mathematical measurement (i.e., the image features extraction), and, second, methods that consider the HVTs characteristics.

In this approach [13], considering the VIQET measures with incorporation of HVTs, that is, image feature extraction using HVTs characteristics. A flow chart of the new proposed approach is shown in Figure 2.1.

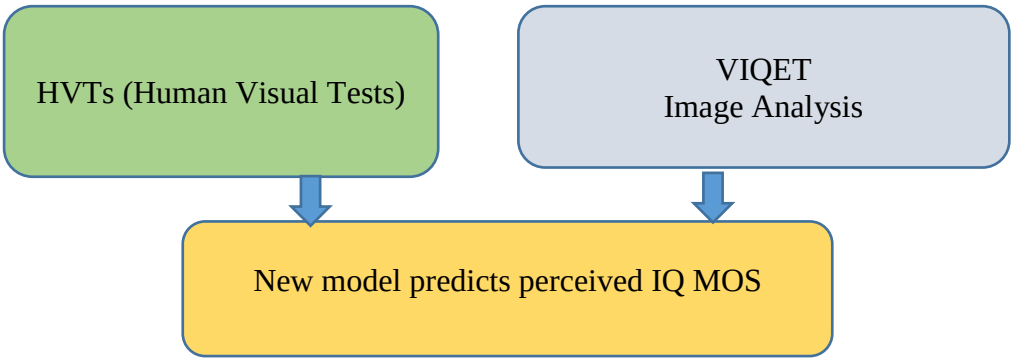


Fig.2.1. Image quality assessment flow [13].

2.3. Subjective image quality metrics

Most of the objective metrics consider the statistical or mathematical measurement for finding the image artifacts. The MSE [81] and the PSNR [87] are the most widely used pixel-based image quality metrics. These techniques are simple and fast, but widely criticized for not correlating well with human visual perception and require reference images. PSNR is a simple pixel-based comparison method whereas MSE is designed on statistical features for finding differences between reference and original images. They do not consider the relationship between pixels.

Although MSE or PSNR are considered as a quality metrics but these are not consistent with the HVS as they measure every pixel with equal priority. In addition, no information of structure, contrast, visibility, etc. are considered in these methods. These metrics consider the power of the error signal, but not how it affects the image. In reality pixels at different position create various effects on the HVS. Since image quality is strongly based on subjective observations, these metrics rarely work accurately on quality judgement.

MSE is the differences between corresponding scores of the reference and the distorted images and it can be defined as:

$$MSE = \frac{1}{V} \sum_{n=1}^{V} (S_n - MOS)^2 \quad (2.1)$$

where V is the number of viewers participated in the Visual tests . S is the corresponding score given by viewer per each individual image in the Visual test.

MOS represents the scores average of each image in visual tests.

PSNR maps the MSE in a logarithmic way which is defined as:

$$PSNR = 10\, log_{10} \frac{MAX}{MSE} \quad (2.2)$$

where *MAX* is the maximum value that an image can get according to the scoring table, which is: Poor =1 and Excellent = 5.

PSNR is a popular and widely used metric to evaluate and quantify performance of image processing algorithms. But it exhibits weak performance in perceived image quality assessment due to pixel-wise error computation.

Human vision is sensitive to contrast sensitivity of an image. Therefore, these mathematical models do not always correlate with human perception and fail to predict the perceived quality of an image.

Several experiments have been ran in order to analyze the prediction performance of MSE and PSNR over MOS. In order to show the performance of MSE over MOS for different level of image quality attributes of 10 different scenes with 5 images in each individual scene e.g. building, hall, bar and room.

2.4. HVS-based feature metrics

As image quality assessment should depend on the assessments made by humans, a better understanding of features of the **HVS** should lead to more effective comparisons, which in turn will assess more realistic and reliable perception. Recently, a great deal of effort has been made to the development of visual models that take advantage of the known characteristics of the HVS [90]. The aim of the HVS- based objective assessments are to evaluate how strong the distorted information is perceived by the metric, according to the characteristics of the HVS. However, most of these proposed approaches require the original image as a reference, therefore, these metrics cannot be used in real time applications.

Recently, extensive research into designing objective metric based on artificial neural network has done. Gastaldo et al. [89] proposed a neural network based objective metric using PSNR. They present their quality assessment on a circular back-propagation neural network model. Though the performance of their metric is good, it was not verified widely by other databases. Horita et al. [88] introduced a FR metric for stereoscopic color images by considering the HVS. They considered all the components such as luminance and color of the CIELab color space (i.e., colors are considered as combinations of red and yellow, red and blue, green and yellow, and green and blue).

Image impairments were quantified by the differences of coded and reference images. The MSE algorithms were used in their metrics for finding the pixel distortions. Gorley et al. [87] developed a point matching technique to identify the contrast and the luminance changes between the left and the right images. Most of these metrics were developed based on the structural similarity, segmentation and/or quality-awareness. Therefore, they fail when the reference image or at least the partial information of the reference image is unavailable.

Wang et al. [91] designed their metric by finding two major image artifacts, blockiness and blur within an image. The image impairments were identified from the shifting of the left and right image pair. The major drawback of their metric is high computational cost. A similar approach was proposed by Sazzad et al. where the metric was developed to find the texture areas of an image.

However, the metric was developed for image quality assessment. In the region-based matching, researchers investigated the sensitivity of the human eyes for occluded regions. Marziliano et al. presented a NR blur metric by measuring average edge transition widths, and this blur measure was used to predict the quality of JPEG2000 compressed images.

Many researchers proved the sensitivity of human eye on image features such as occluded, edge, or smooth regions of an image. Some of them focused on the brightness, contrast, or structural similarities of an image. A similar approach was developed by Sazzad et al. where three local features: edge, plain (i.e., uniformly colored areas), and textured areas have been considered to find the image impairments. Their metric has a large number of parameters, which poses a danger of over-fitting the metric. Moreover, the computational cost of their metric is very high.

The discussion above implies that for the past few decades, a lot of research has been done in the field of objective image quality assessment, but no comparable effort has been taken towards the quality assessment of perceived IQ. More specifically, very little research has been done on NR smartphones IQ assessment by considering the human visual system. Emerging smartphones with embedded camera and display technologies still require a larger number of IQ metrics and methodologies by taking into account the fundamental characteristics of the HVS and typical distortions of smartphones image content. Therefore, hopefully this research would be recognized as a great contribution in the field of smartphone IQ assessment.

2.5. Implementation of subjective image quality assessment

Creating a reliable test content of everyday scenes including large set of IQ attributes, a broad range of images should be used in order to reveal different quality issues. To achieve this, the test content was created according to the recommendations of VQEG, where the images were chosen based on the criteria for subjective IQ assessment. Pictures of natural image contents captured by smartphones camera in native resolution of 1920x1200 pixels. The original images will be used as a reference. Each original image has been processed by adding the IQ attributes (brightness, contrast, color and sharpness) in various levels of parameters.

As the test content was created according to the VQEG recommendations in "Recommendation P.913", ITU-T, 2014 [85]. Contents were carefully selected to represent a wide

range of different situations and demands for pictures. Also, recommendations of Photo-space standards set by I3A were considered when choosing the image contents.

Each original image was processed in order to enhance the image quality attributes in various levels: brightness, contrast, sharpness and color Saturation.

The overall test content for HVTs and VIQET analysis includes the following everyday scenes as demonstrated in Figure 2.2:

1. Outdoor day- landscape, people.
2. Indoor – without backlight.
3. Indoor – with backlight.
4. Outdoor night.

Images in Figure 2.2 were selected [13] as test content for the subjective IQA. Images (building, lake, man, taxi) are an example of outdoor day of some of the everyday city landscape of buildings.

Images (king, room) are an example of indoor without backlight of some of the everyday scene when pictures are taken indoor without window or door backlight. In this case the main objects are lighted. Image (hall) is an example of indoor with backlight of some of the everyday scene when pictures are taken indoor with window or door backlight. In this case the main objects are shaded. Images (sunset, bar, airplane) are an example of outdoor night with considerable amount of black in the photo and the space was very bright and well-lit. A camera needs to properly meter off the light so that the people don’t get blown out, and so the shadows are precisely dark

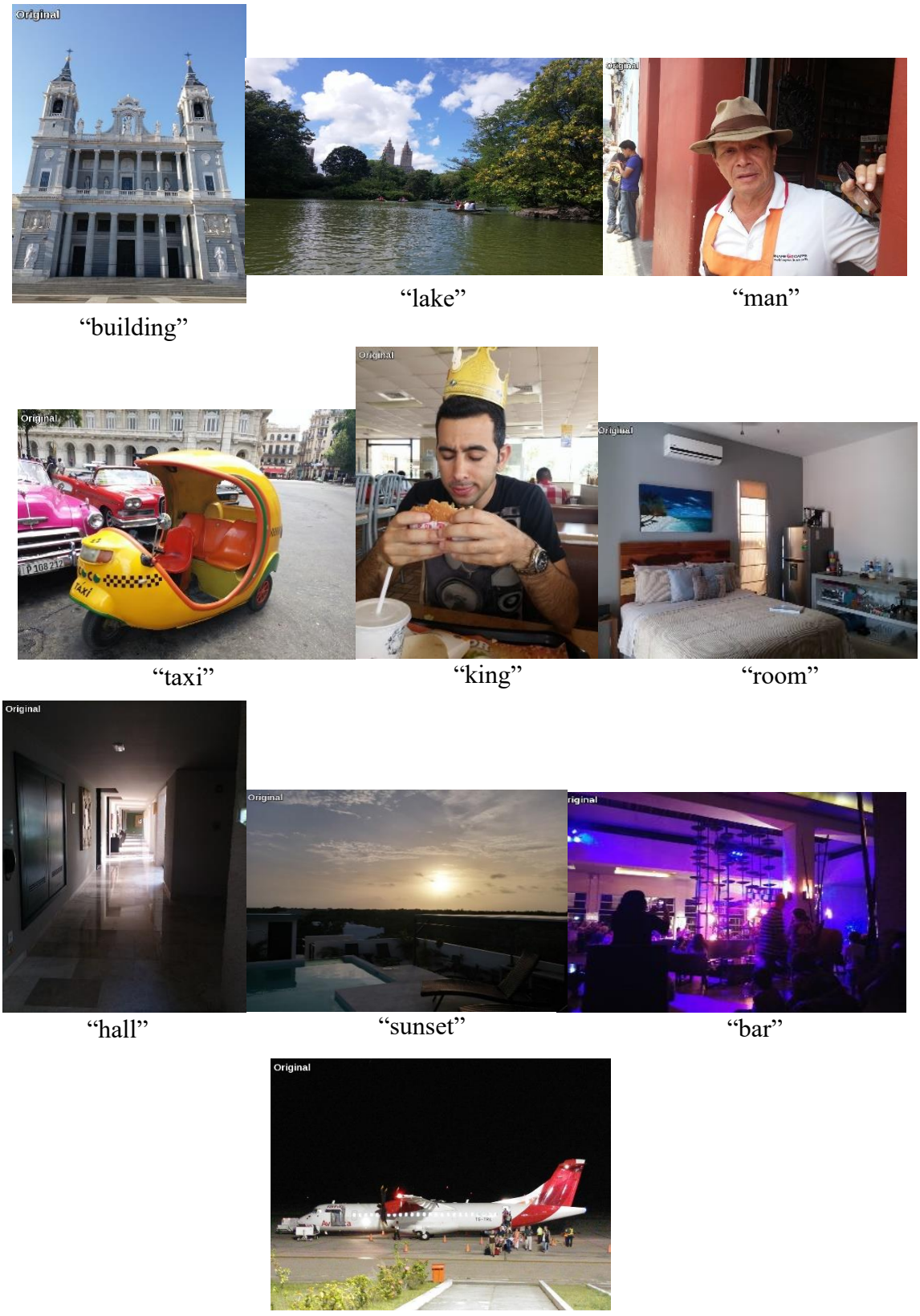


Fig.2.2. Subjective visual tests content of original images.

Test material with controlled image quality attributes

In order to measure the image quality attributes effect on perceived image quality, a set of natural images and four image quality attributes were added to each original image then added four different image quality attributes levels of each single original image[13, 65] (demonstrated in Figure 2.3).

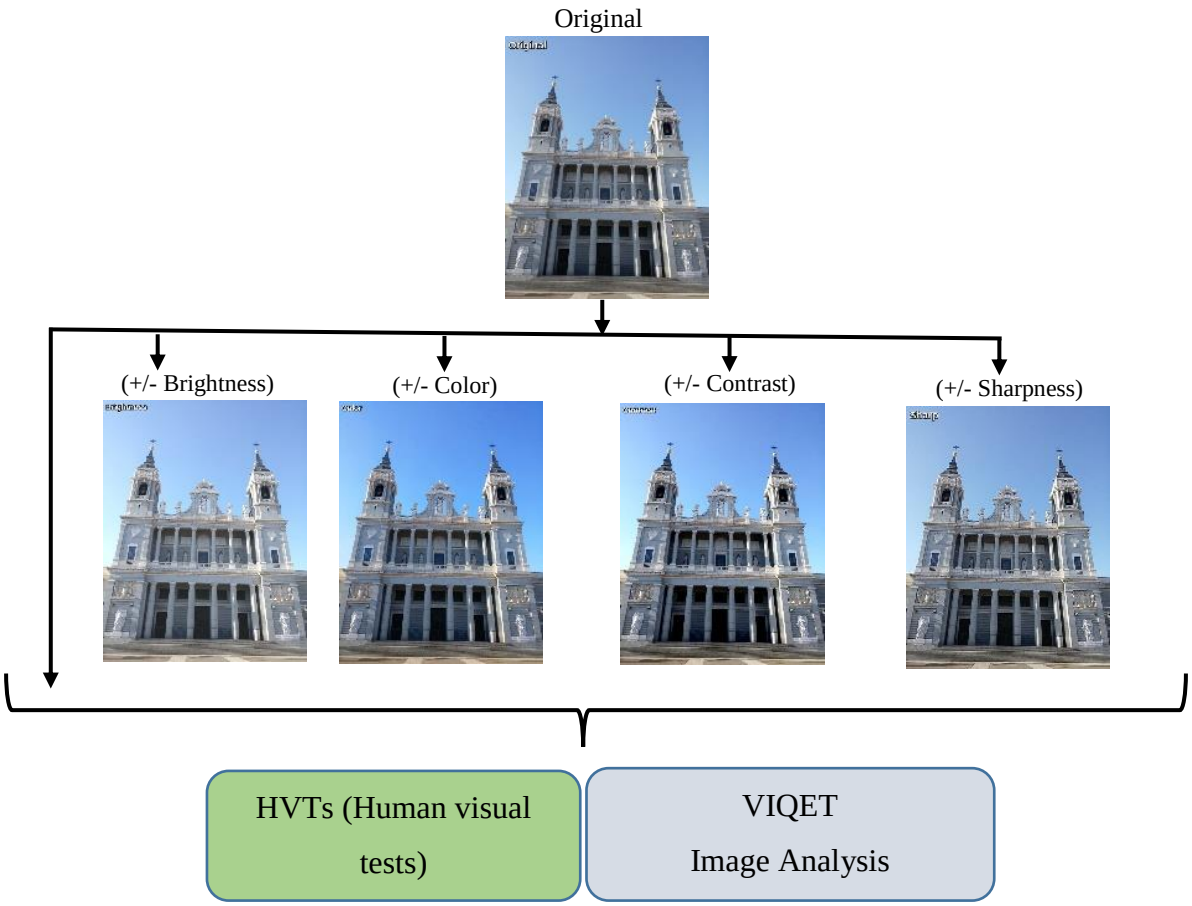


Fig.2.3. An example of test material processing “building” [13].

2.6. Implementation for perceived image quality prediction

The study was divided into two main phases (*phase I* and *phase II*) towards the new model development [13, 65]. The new model, which was developed as the result of the research work examined through a broad IQ assessment experiment. This model approval process performed in using a large number (fifty) of natural situation images as test content. Thirty-five “non-experts” observers gave the final visual experiment scores.

Human visual test process

The flow charts in Figure 2.5 and Figure 2.6 provide an overview of the methodology used in this research to test and perform subjective IQ assessment in phase I, phase II and final IQ

assessment for the new model approval. Figure 2.4 illustrates the visual tests setup for subjective IQ assessment experiment and VIQET image analysis. The IQ visual tests assess the subjective image quality of images material presented on a smartphone display (Samsung Galaxy S5) in a simulated viewing environment. The display resolution, however, was 1920 X 1080 in all tests. Each subjective experiment collected valid data from the participants. A statistical criterion was used to verify that the data from a viewer were correlated to the average of the other viewers' data. All viewers were screened prior to participation for normal (20/30) visual acuity with or without corrective glasses (per Snellen test or equivalent) and normal color vision (Appendix C, per Ishihara test).

The test material consisted of fifty images, which included the processed images with different IQ attributes. The duration of each image sample was one minute. For each experiment, the image samples were functionally divided into five subsets: original, brightness, contrast, color and sharpness.

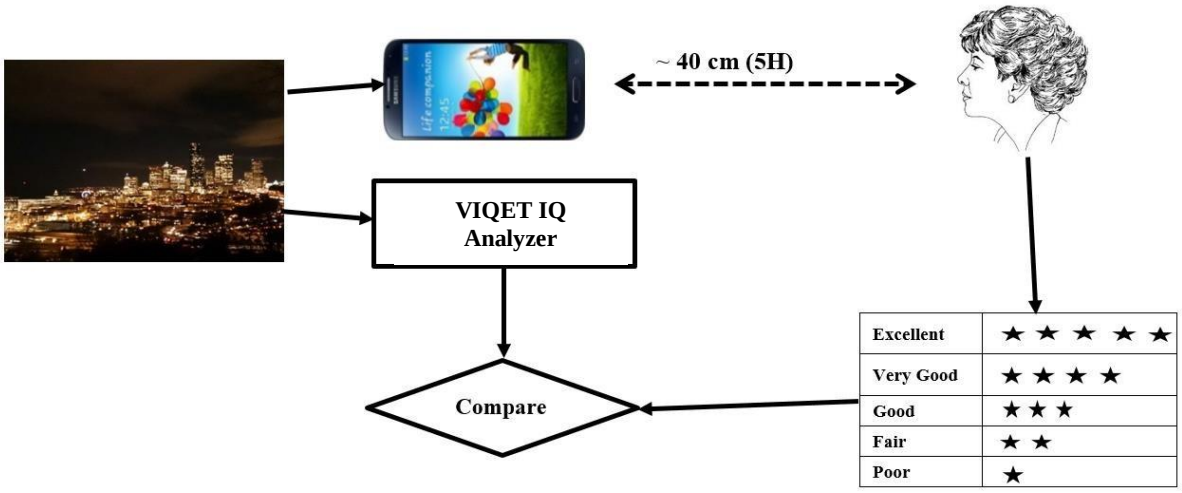


Fig.2.4. Experiments set-up for IQ visual tests and VIQET analysis [13, 65].

The subjective test methodology was the absolute category rating (ACR) method with hidden reference. The subjective picture quality of the image samples was assessed, in all subjective tests, using the ACR method [ITU-T Rec. P.910]. The ACR method is a single stimulus method in which the image samples are presented one at a time, and rated independently using the five-grade image quality scale shown in Figure 2.6.

During the data analysis the ACR scores given to the processed versions were subtracted from the ACR scores given to the corresponding reference (non-processed original image) in order to obtain a DMOS. This procedure is known as "hidden reference". This choice was made because ACR provides a reliable and standardized method that allows a large number of test conditions to be assessed in any single test session. For these, the viewers performed the experiment using

custom made software. The software controlled both the timing and order of presentation of the stimuli.

The order of image presentation of the samples was changed randomly for different groups of viewers. All viewers received instructions, which followed upon guidelines to ensure consistency across subjective experiments. To get the viewers familiar with the assessment tasks and with the levels of image quality used in the experiment, a small number of practice trials were performed at the beginning of the experimental session. To control the effects of fatigues, a short break was given after about half of the images sample had been assessed.

Accordingly, in this scenario, each experiment included the following steps [66]:

- Introduction and instructions to viewer.
- Practice images: these test images allow the viewer to be familiar with the assessment procedure and software. Ratings given to practice images were not used for data analysis.
- Short break.
- Assessment of second half of the image samples.

The test room conformed to Recommendation ITU-R BT.500-11. In general, a test session involved only one viewer per display assessing the test material. Viewers were seated directly in line with the center of the smartphone display (Figure 2.4) at a viewing distance equal to five times the height of the picture (i.e., 5H) in all experiments.

2.7. Human visual test and VIQET image analysis

This part of the study begins with an analysis of the images selected for test content for the HVT and with the VIQET.

Design requirements followed by a detailed description of design and development procedures of objective image quality assessment model. The algorithm consists of two parts: first, finding how image quality attributes effect observers' preferences through HVTs, and, second image analysis with the VIQET.

Taking brightness, contrast, color saturation and sharpness as major image quality attributes, because these are the most visible everyday images. Image quality attributes improve or degrade the perceived visual quality of an image.

However, the relationship between the image quality and the level of increasing/decreasing IQ attributes depends on the texture contents of an image. In order to verify this relationship, the IQAs parameters level was changed during *phase I* experiments. Therefore, the results indicate that visibility of image quality is strongly depended on the IQ attributes added to the image.

Phase I: Rating the image quality attributes

In this part of the research, ten visual experiments were performed in order to rate the contribution of each IQA to the **perceived IQ**. In this phase ten groups of five images of the same image with different IQAs levels were prepared (see Figure 2.3) as test content. Ten observers rated the IQ of each image in ten cycles. Scores were analyzed after each cycle and IQAs parameters were changed in the test content images before the next cycle. Phase I process flow chart is illustrated in Figure 2.5.

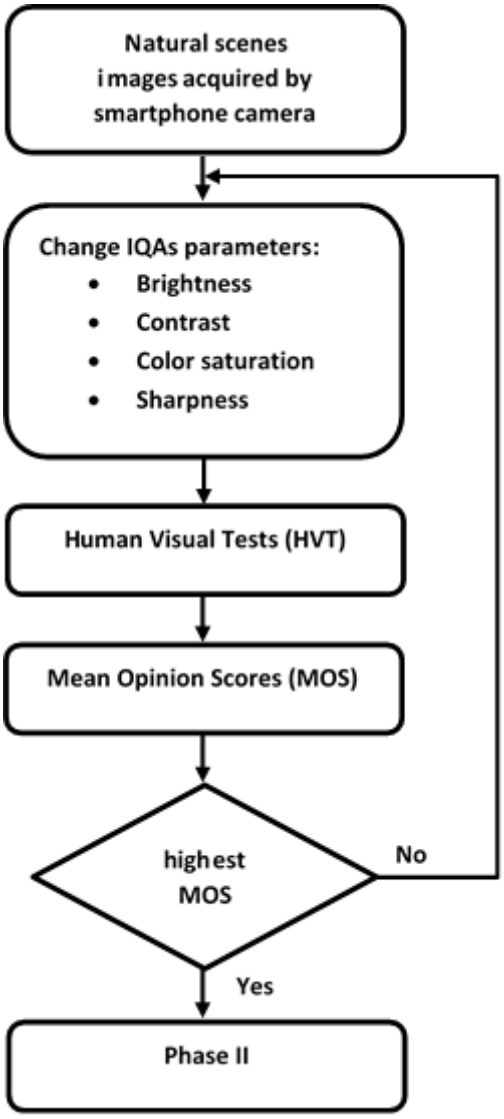


Fig.2.5. *Phase I* procedure flow chart for IQAs rating [13].

2.7.1. Subjective IQ assessment for IQAs rating

A total number of 98 non-expert (the term non-expert is used in the sense that the viewers' work does not involve television picture quality and they are not experienced assessors) subjects

participated in these experiments. All of the experiments viewers had normal or correct-to-normal sight. Each subject viewed the images in test content with a random order on each smartphone and viewed ten sets of five images in each set (original, + brightness, + contrast, + saturation, + sharpness).

He/she rated his/her perceived image quality in the **ACR** 5-point scale as shown in Figure *2.6* (corresponding to the perceived quality of **"excellent," "good," "fair," "poor," and "bad").**The environment to the experiments was set following the suggestion of ITU-R recommendation BT.500-13 [92]. Before the formal test, the subjects were asked to rate a few example images to get familiar with the scoring scale and the image browsers.

Excellent	★ ★ ★ ★ ★
Very Good	★ ★ ★ ★
Good	★ ★ ★
Fair	★ ★
Poor	★

Fig.2.6. Staring points ranking [65].

Subjective image quality experiments results

The subjective experiment was conducted in the Image Processing laboratory in the "Braude" engineering college. A double stimulus impairment scale (DSIS) method was used in the subjective experiment. The test format is shown in Figure 2.7. Images were displayed sequentially in the method. At the end of the presentation of each image, the subject was asked to assess the image quality.

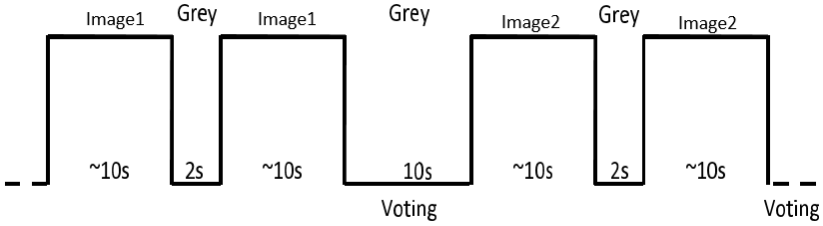


Fig.2.7. Double stimulus impairment scale test format [96].

Ninety-eight non-expert subjects were shown the test content (in 10 different sessions), most of them were college/high-school students. A 5.1-inch LCD (Samsung Galaxy S5) display (resolution: 1920 × 1080) was used in this experiment to display the images and the subjects were instructed about the limited horizontal viewing angle to perceive image correctly. MOS were then computed for each image after the screening of post-experiment results according to ITU-R Rec. 500-10 [96]. The subjective test conditions and parameters are summarized in Table 2.1.

Table 2.1. Subjective test conditions and parameters

Method	DSIS (Double stimulus impairment scale)
Evaluation scales	5 Grades (Impairment scales)
Subjects	98 (Non expert, students)
Display	5.1-inch, LCD Samsung Galaxy S5
Display resolution	1920×1080 pixels
Viewing distance	5H (H = picture height)
Room illumination	Dark

The MOS results given by visual test observers were calculated and analyzed. The IQAs of the images in test content were modified according to the MOS results and another visual experiment was ran (total of 10 cycles) until the highest MOS results achieved. These IQAs parameters will be used in *phase II* process.

Phase II: rating image quality with VIQET

The VIQET is an objective, non-reference photo quality evaluation tool. VIQET is an open source tool designed to evaluate quality of consumer photos. In order to perform photo image 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 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 2.8 demonstrates an example of VIQET RGB histogram and Figure 2.9 demonstrates VIQET Sharpness map.



Fig.2.8. An example of VIQET RGB histogram [13].



Fig.2.9. An example of VIQET Sharpness map [13].

Characteristic of the image quality attributes

Many of the IQAs were analyzed in many image quality researches are similar and have common denominators, which enables them to be grouped within more general IQAs In order to reduce the dimensionality and create a more manageable evaluation of IQ. There is usually a compromise between generality and accuracy when it comes to dimensionality. Linked most of the IQAs to four different dimensions, considered as important for the evaluation of IQ. This results in a reasonable compromise between accuracy and complexity. The IQAs found in the literature were reduced to: brightness, color saturation, contrast and sharpness.

Therefore, a grouping of these physical IQAs is needed. In order to create a simple and intuitive illustration of the IQAs and their influence on overall IQ we have turned to Venn diagrams. Venn diagrams may be used to show possible logical relations between a set of attributes. However, it is not possible to create a simple Venn diagram with a four fold symmetry. Therefore Figure. 2.10 illustrates the IQAs using only five folds, leaving the physical IQA out. This does not mean that the physical IQA is less important than the others.

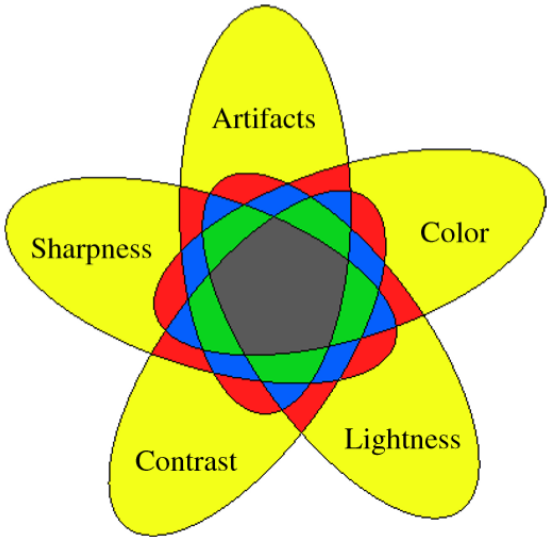


Fig.2.10. The Venn diagram illustrates how the overall IQ is influenced by lightness (brightness/contrast, lightness, color, sharpness and artifacts [97].

2.7.2. Extraction of IQAs to VIQET image quality categories

The calculation and analysis of the results of subjective IQ experiments which have been performed in *phase I* were diploid in the test content and have been extracted from general IQAs (brightness, contrast, color saturation and sharpness) to VIQET image quality attributes parameters, as performed in Table 2.3. The range of VIQET image quality categories is demonstrated in Table 2.2.

Table 2.2. VIQET image quality categories

Image quality categories	**Score**	**Range**
MOS	4.5	1 – 5
Multi- scale edge acutance	12.14	Higher is better
Noise signature index	99.39	0 - 589
Saturation	123.41	0 represents B&W image
Illumination	92.00	0 - 255
Dynamic range	106.72	Represents gray levels

Table 2.3. Image quality attributes extraction to VIQET image quality categories

Image quality attributes	**VIQET image quality categories**
Brightness	• Illumination • % Over-exposed • % Under-exposed • Lux
Contrast	• Dynamic range
Color	• Saturation • Color warmth
Sharpness	• Multi-scale edge acutance • Multi-scale texture acutance • Noise signature index

2.7.3. 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 image quality evaluation, VIQET requires a set of photos from the test device. It estimates an overall

MOS for tested images based on the individual image MOS scores in the set. The estimated MOS for each photo is based on a number of image quality categories 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 in ***phase I*** for rating IQ by **HVTs** were required for IQ rating by VIQET to analyze each individual image and get its IQ scores (IQ categories). This IQ assessment has been ran in eight cycles with ten viewers in each session.

The images were tested by VIQET and MOSp were calculated and analyzed, the VIQET parameters were modified according to the relationship between IQAs and VIQET categories model [13, 66]. Figure 2.11 illustrates the IQ assessment process flow chart with VIQET. ***Phase II*** ended once the highest correlation between ***phase I*** MOS (in subjective IQ assessment) and MOSp calculated by VIQET. The next step will be evaluating the new model in large group of viewers and VIQET simultaneously.

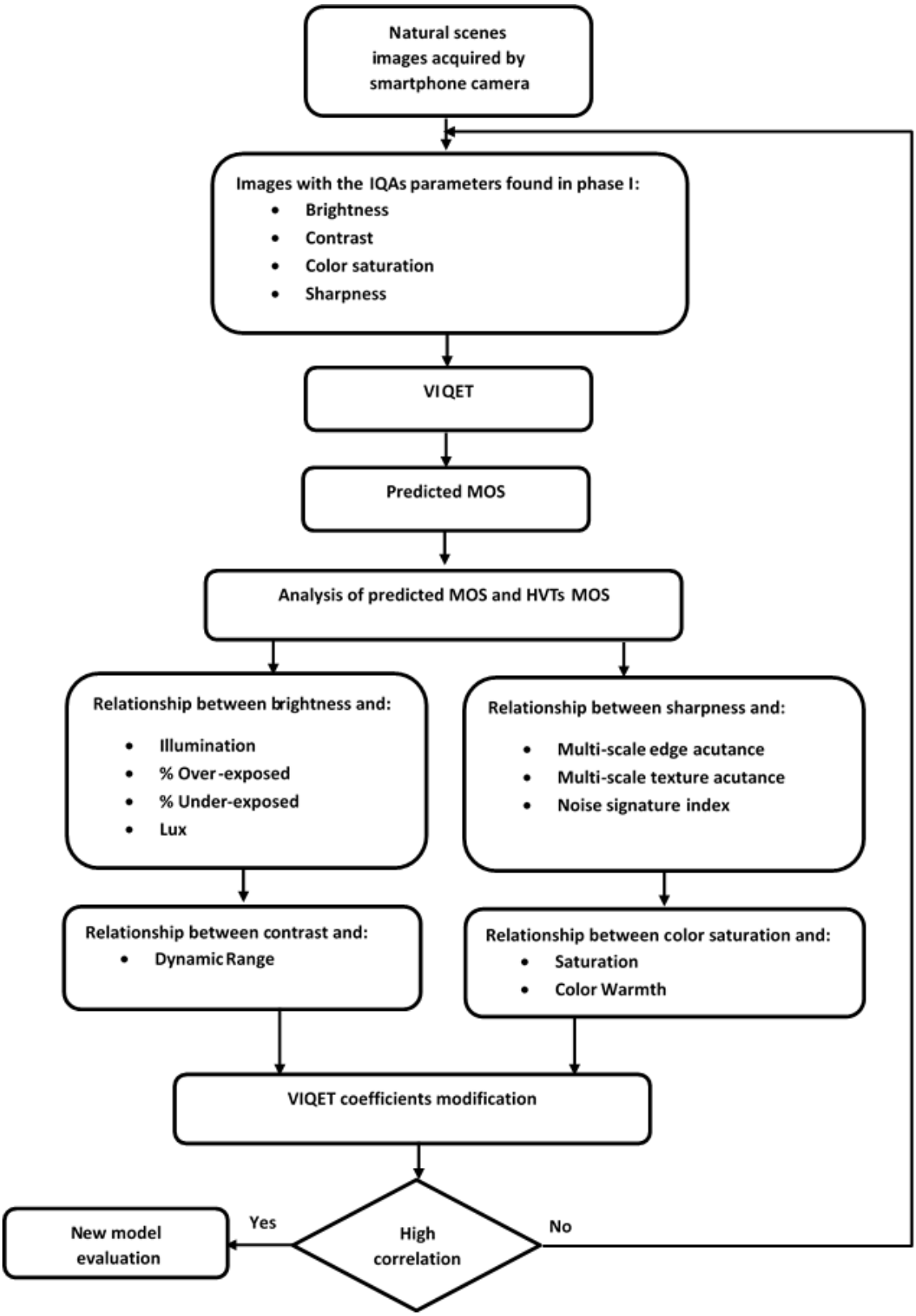


Fig.2.11. Phase II image quality assessment with VIQET [13].

2.8. Examination of the new model

The new model for prediction of perceived IQ was examined with large group of viewers in continues session of visual experiment [13]. The final visual experiments process is described in Figure 2.12. In this session, 35 observers scored the 50 images in the test content, the results were assumed as "Perceived smartphone image quality". The VIQET results were assumed as the "Prediction of perceived image quality". High correlation between the two groups of data will strongly support the reliability of the proposed new model in this research.

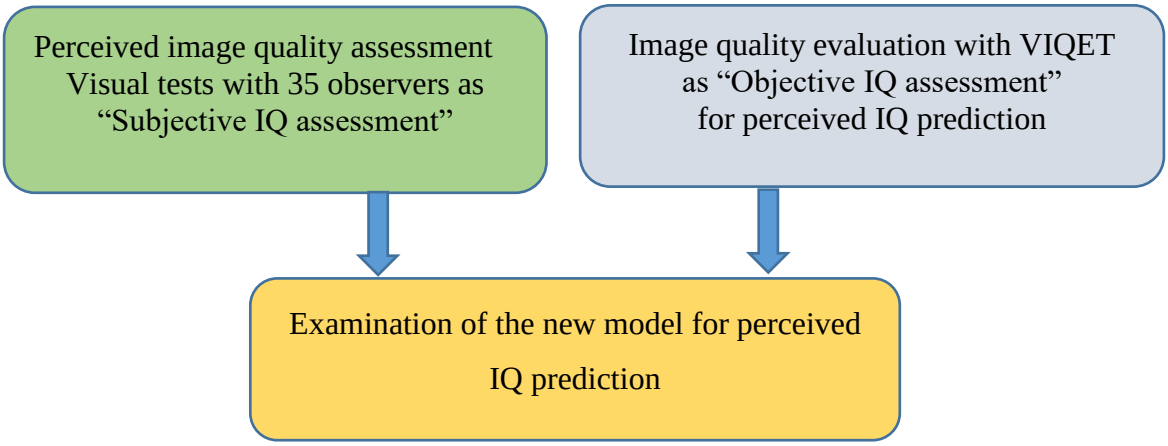


Fig.2.12. Process flow chart of the new model examination [13, 66].

Raw data processing

After the subjective tests, the credibility of assessment results was checked using the linear Pearson correlation coefficient suggested by ITU-T recommendation P.913 [93]. The Pearson correlation coefficient (see equation 2.3) measures the linear relationship between a model's performance and the subjective data. Its great virtue is that it is on a standard, comprehensible scale of -1 to 1 and it has been used frequently in similar testing.

The LPCC is calculated as follows:

$$CC=\frac{\sum_{i=1}^{N}(Xi-\overline{X})*(Yi-\overline{Y})}{\sqrt{\sum_{i=1}^{N}(Xi-\overline{X})^2}*\sqrt{\sum_{i=1}^{N}(Yi-\overline{Y})^2}} \quad (2.3)$$

Xi denotes the subjective score MOS(i) in HVT for processed image (IQ attribute added), X denotes the MOS ("objective") of processed image (IQ attribute added) and Yi denotes the subjective score MOSp(i) in HVT of original image (no IQ attribute added), Y denotes the MOS ("objective") of original image. N in equation (2.3) represents the total number of images considered in the analysis. Therefore, in the context of this test, the value of N in equation (2.3) is:

N=35. The sampling distribution of Pearson's CC is not normally distributed. "Fisher's z transformation" converts Pearson's CC to the normally distributed variable z. This transformation is given by the following equation:

$$z = 0.5 \cdot \ln\left(\frac{1+R}{1-R}\right) \tag{2.4}$$

The statistic of z is approximately normally distributed and its standard deviation is defined by:

$$\sigma_z = \sqrt{\frac{1}{N-3}} \tag{2.5}$$

Table 2.4. LPCC of each IQ attributes in HVT (*Phase I*)

IQ attribute	**Brightness**	**Contrast**	**Original**	**Color Saturation**	**Sharpness**
LPCC	0.22	0.85	0.28	0.72	0.25

The values of LPCC of each subject in HVT (***Phase I***) were calculated. As a result, the number of the valid subjects (i.e., 35) meets the requirement of the VQEG. Table 2.4 lists the LPCC of viewer's rating scores on each IQ attribute after the screening process. The perceived image quality of each image was measured in terms of the average score of all valid subjects, also known as the MOS [93].

The subjects in VIQET analysis were also screened according to the screening result in *Phase I*. The perceived image quality difference of each image pair was measured in terms of the average score of all valid subjects, also known as the DMOS [93]. Then, Cronbach's alpha value was computed to measure the internal consistency of the valid scores on each device. As per the results the value of alpha of each device is considerably large, which indicates that there is a strong internal consistency among the valid subjects.

Perceived image quality results of HVTs

In this section, the perceived image quality on diverse IQ attributes is firstly investigated based on the rated scores, that is, MOS, for the images categories: Outdoor Day, Indoor and Outdoor Night respectively. Considering the possible influence of the IQ attributes, these images have the same resolution (i.e., 1080P) but in different IQ attributes. Take the high and low quality images with ten randomly selected contents as an example. the relationship between the MOS, MSE, PSNR and the IQ attributes of outdoor day images is shown in Figure 2.13 - Figure 2.16,

the relationship between the MOS, MSE, PSNR and the IQ attributes of Indoor images is shown in Figure 2.18 – Figure 2.19 and the relationship between the MOS, MSE, PSNR and the IQ attributes of outdoor night images is shown in Figure 2.11 – Figure 2.20.

It can be seen that there is no significant increase or decrease in the perceived quality, when the Brightness is increased. The viewer's perceived quality is not significantly influenced by the change of Brightness during the viewing process.

Figure 2.13 presents the MOS, MSE and PSNR values and their relationship for the "building" image with different IQ attributes. These values were received from the human observers in the HVT intervals. These values represent the perceived IQ by viewers. The data will be studied and analyzed during the standard IQ attributes extraction process to VIQET IQ criteria.

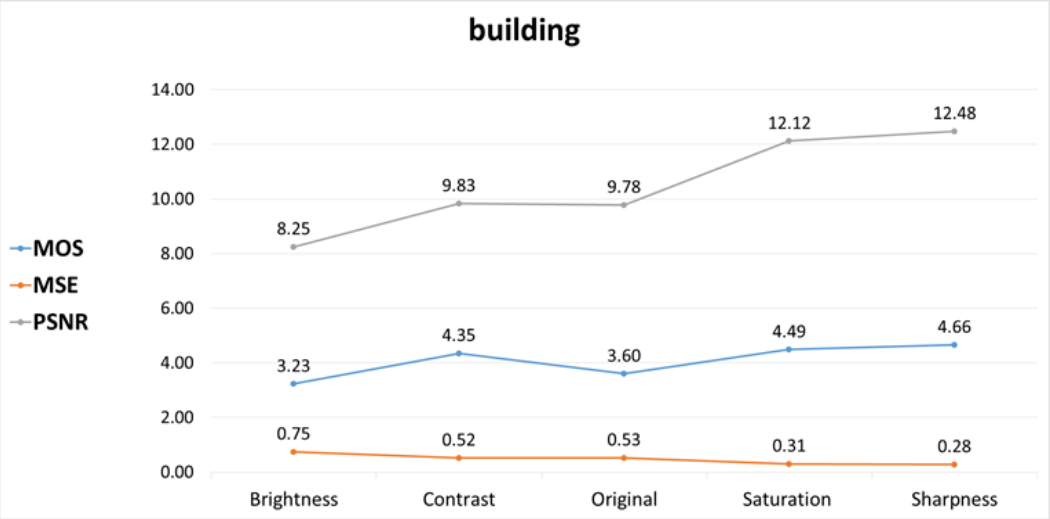


Fig.2.13. Relationship between the perceived IQ MOS, MSE and PSNR of "building".

The cures show that the sharpness and color saturation have the majority in contribution to the perceived IQ. Figure 2.14 presents the MOS, MSE and PSNR values and their relationship for the "lake" image. The curves show that the color saturation, contrast and sharpness have the majority in contribution to the perceived IQ.

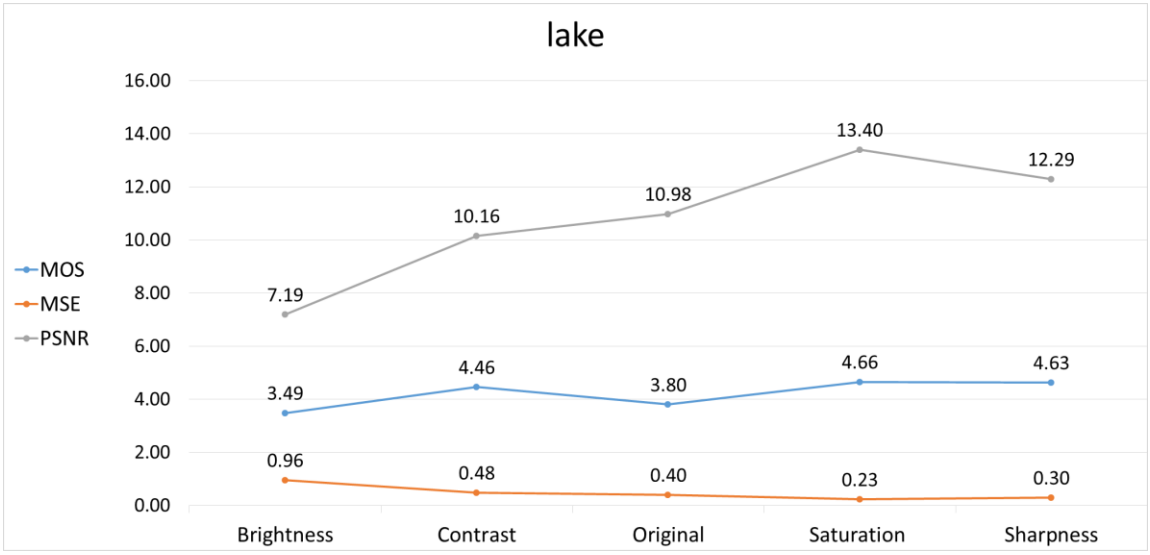


Fig.2.14. Relationship between the perceived IQ MOS, MSE and PSNR of “lake”.

Figure 2.15 presents the MOS, MSE and PSNR values and their relationship for the “taxi” image showing again that the color saturation, contrast and sharpness have the majority in contribution to the perceived IQ.

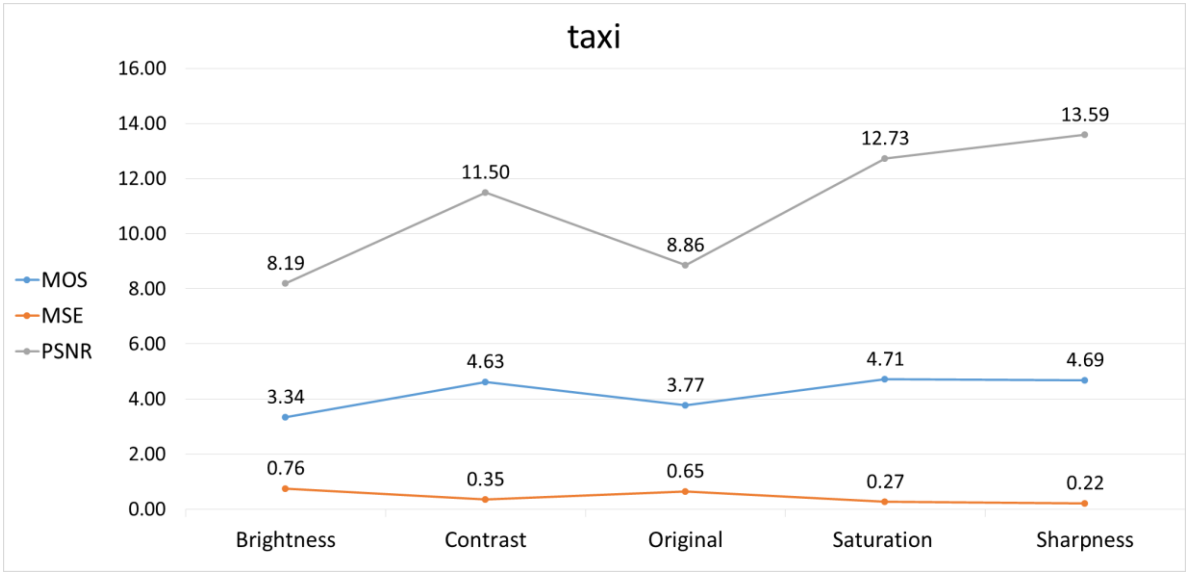


Fig. 2.15. Relationship between the perceived IQ MOS, MSE and PSNR of “taxi”.

Figure 2.16 presents the MOS, MSE and PSNR values and their relationship for the “man” image. The cures show that the color saturation, contrast and sharpness have the majority in contribution to the perceived IQ.

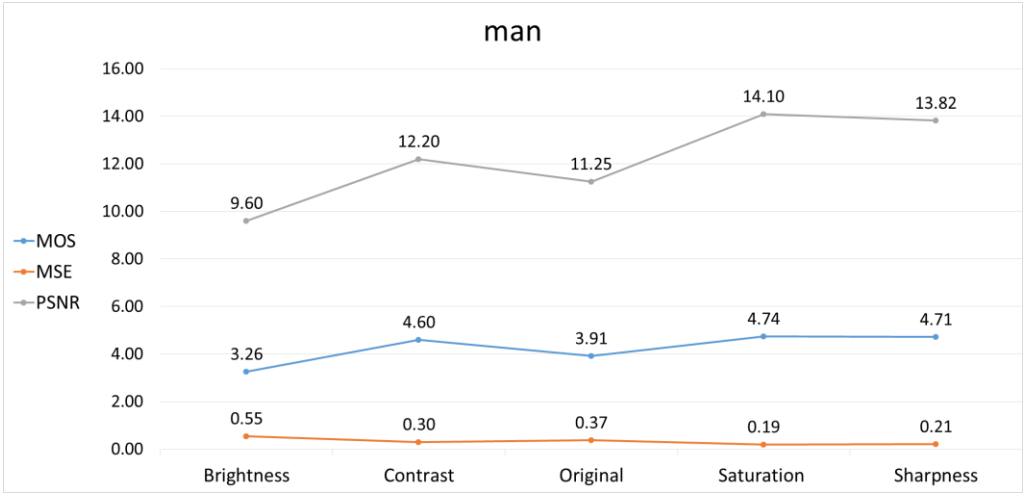


Fig. 2.16. Relationship between the perceived IQ MOS, MSE and PSNR of “man”.

Figure 2.17 presents the MOS, MSE and PSNR values and their relationship for the “room” image. The cures show that the color saturation and sharpness have the majority in contribution to the perceived IQ.

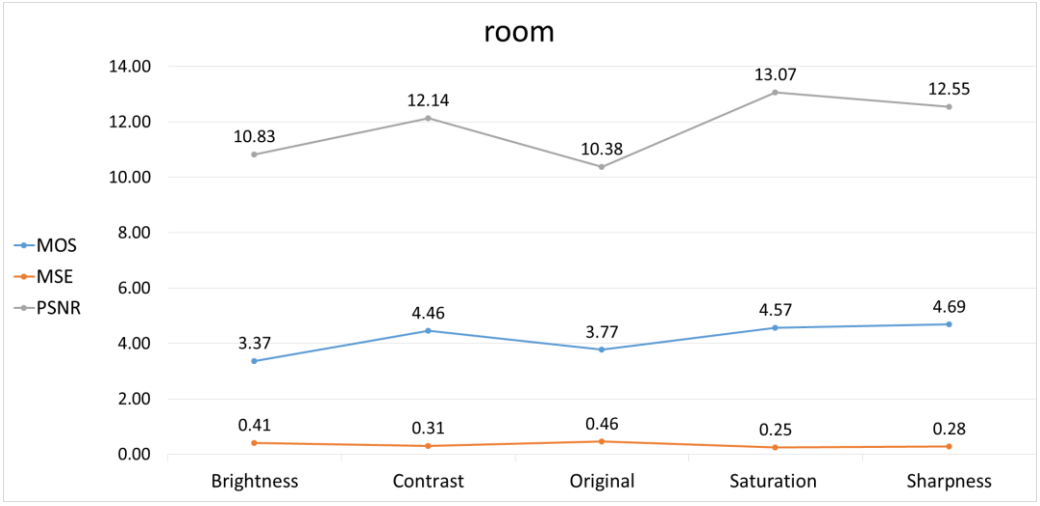


Fig.2.17. Relationship between the perceived IQ MOS, MSE and PSNR of “room”.

Figure 2.18 presents the MOS, MSE and PSNR values and their relationship for the “king” image. The color saturation, contrast and sharpness has the majority in contribution to the perceived IQ.

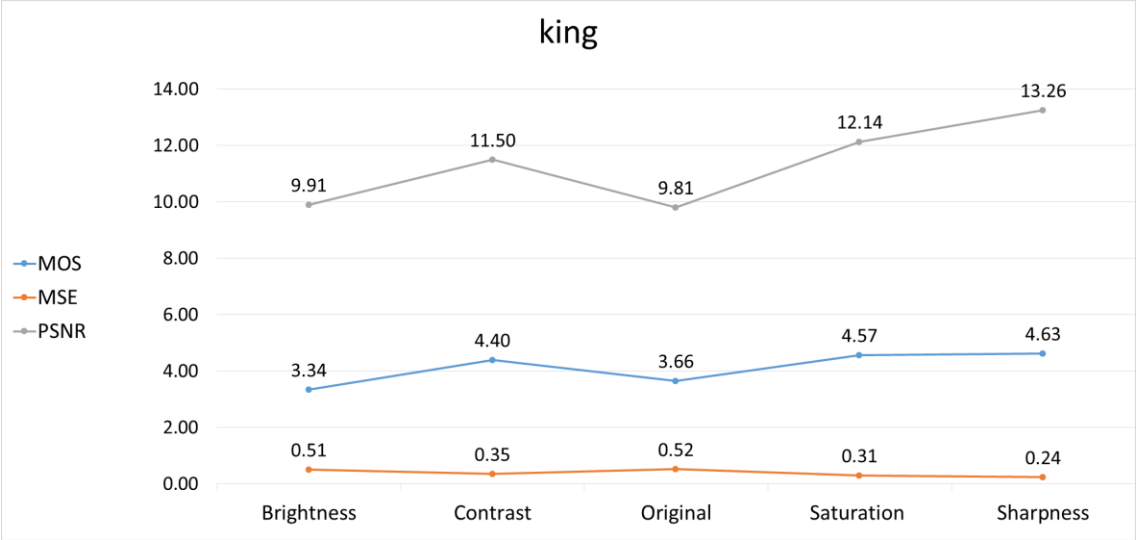


Fig.2.18. Relationship between the perceived IQ MOS, MSE and PSNR.

Figure 2.19 presents the MOS, MSE and PSNR values and their relationship for the “hall” image. The color saturation, contrast and sharpness has the majority in contribution to the perceived IQ.

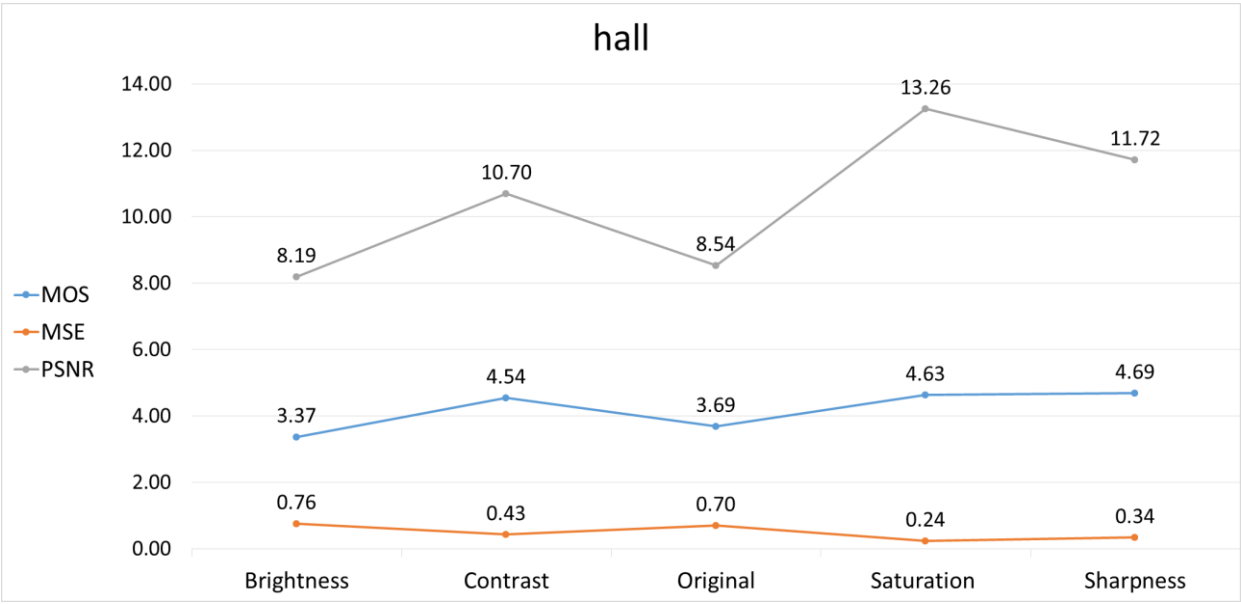


Fig.2.19. Relationship between the perceived IQ MOS, MSE and PSNR of “hall”.

Figure 2.20 presents the MOS, MSE and PSNR values and their relationship for the “sunset” image. The color saturation, contrast and sharpness has the majority in contribution to the perceived IQ.

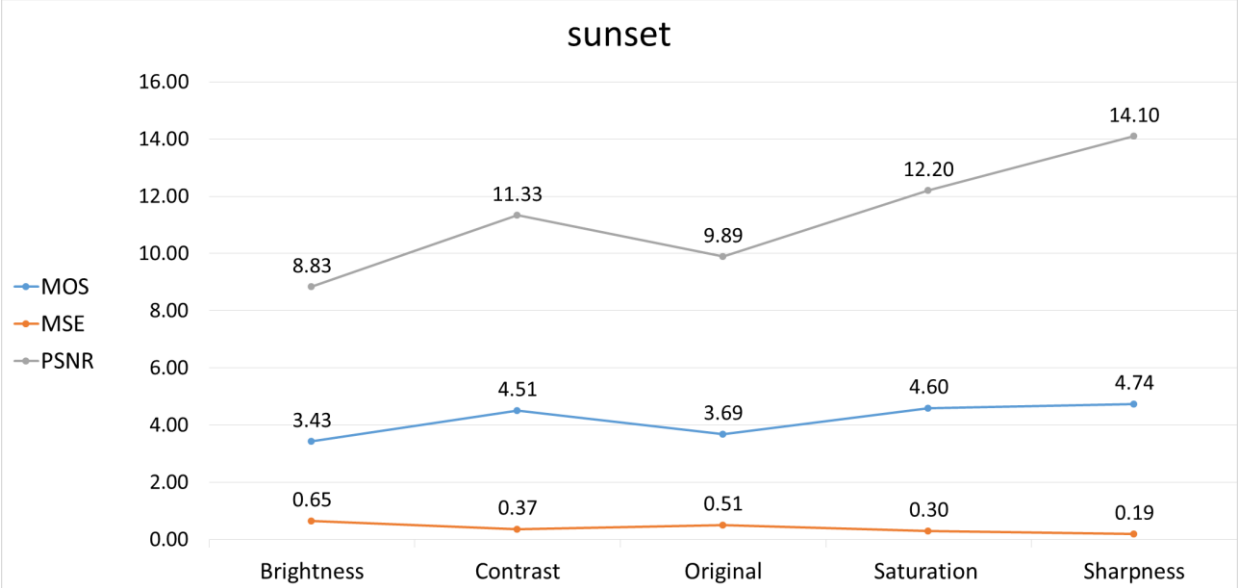


Fig.2.20. Relationship between the perceived IQ MOS, MSE and PSNR of “sunset”.

Figure 2.21 presents the MOS, MSE and PSNR values and their relationship for the “bar” image. The color saturation, contrast and sharpness has the majority in contribution to the perceived IQ.

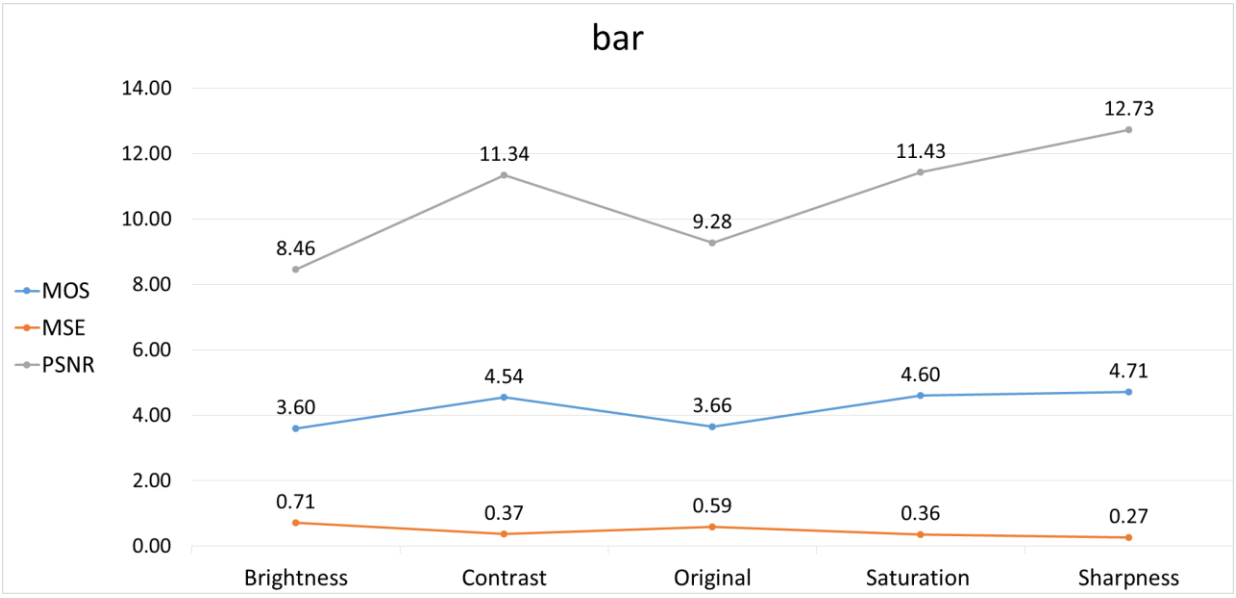


Fig.2.21. Relationship between the perceived IQ MOS, MSE and PSNR of “bar”.

Figure 2.22 presents the MOS, MSE and PSNR values and their relationship for the "airplane" image. Also, the color saturation, contrast and sharpness has the majority in contribution to the perceived IQ.

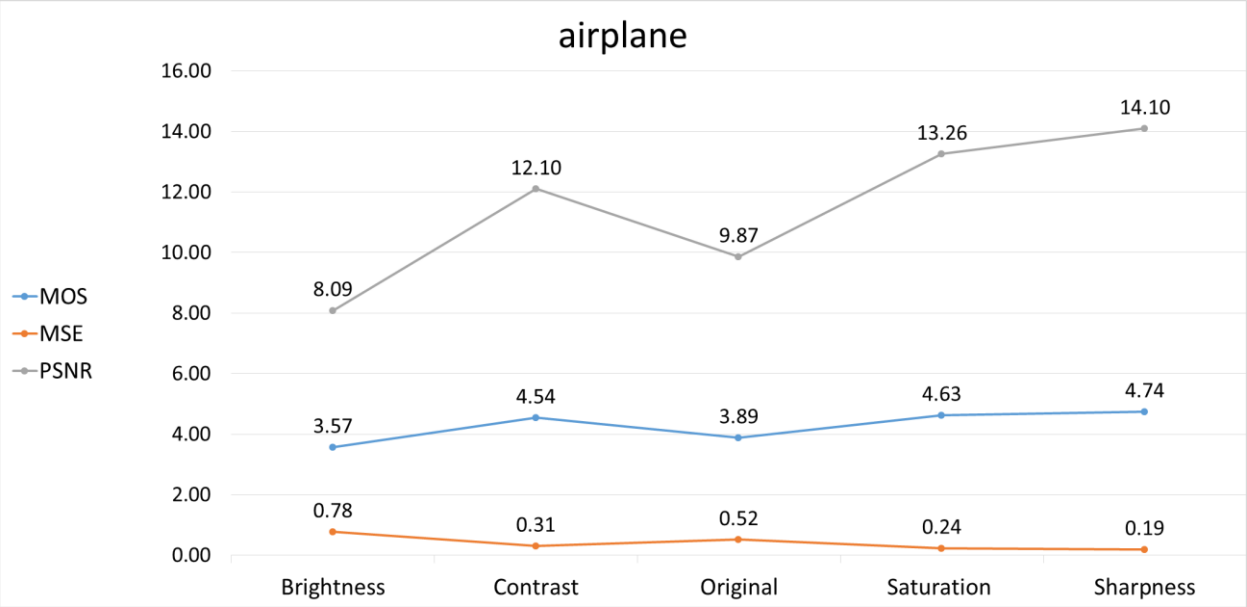


Fig.2.22. Relationship between the perceived IQ MOS, MSE and PSNR of "airplane".

Image quality valuation with VIQET

In a general sense, the MOS of the images displayed on all smartphones are used to illustrate the difference of **perceived image quality** across four IQ attributes (brightness, contrast, color saturation and sharpness) [13, 65, 66]. Figures 2.23 – Figure 2.32 illustrate the rates of different IQ attributes defined by VQEG that measured and calculated by VIQET. Furthermore, a statistical analysis, that is, the one-way analysis of variance (ANOVA), is further performed to check the significance of influence of the IQ attributes on the perceived image quality.

The test is firstly implemented on HVT while observers gave scores to each image displayed on mobile phone display. The analysis is conducted under different IQ attributes. Secondly, all images in test content were analyzed by VIQET and received scores. The significant of scores in test was performed using RMSE.

Figure 2.23 presents the VIQET IQ criteria values and their relationship for the "building" image. It is clear that image with higher sharpness and color saturation received the highest values.

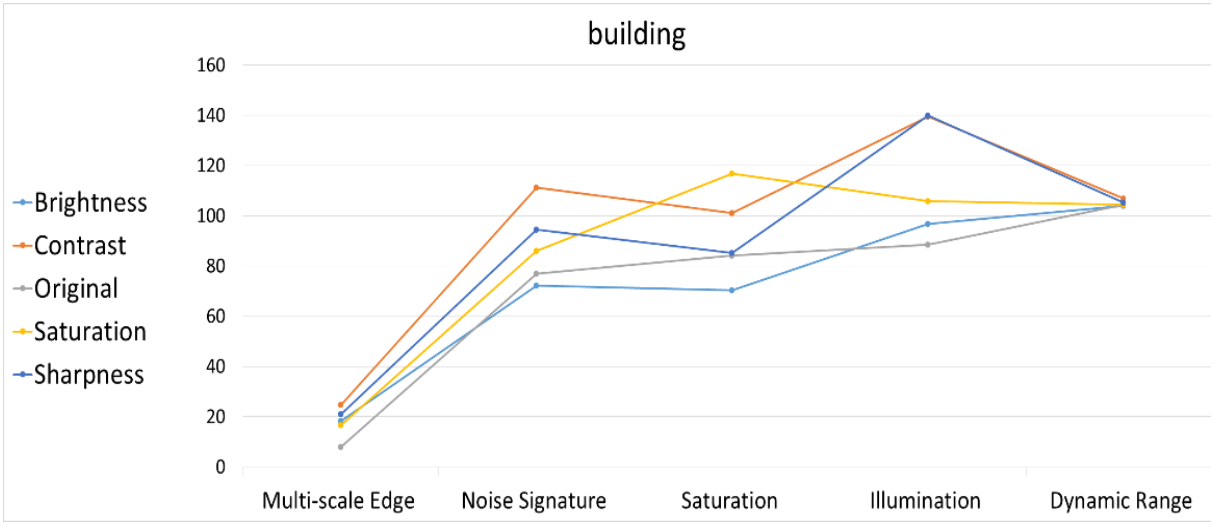


Fig.2.23. Relationship between VIQET IQ scores of “building”.

The results in Figure 2.24 “lake” presents the high contribution of sharpness to the IQ criteria measured values.

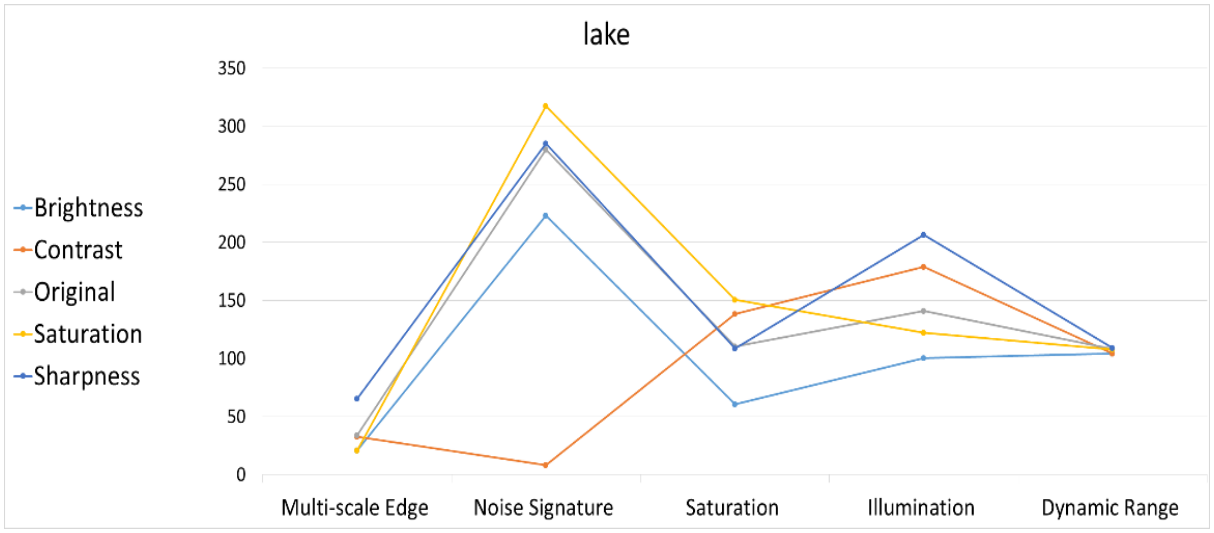


Fig.2.24. Relationship between VIQET IQ scores of “lake”.

In the “man” image, the contrast increases the IQ criteria values, the color saturation also has big contribution on all IQ criteria values.

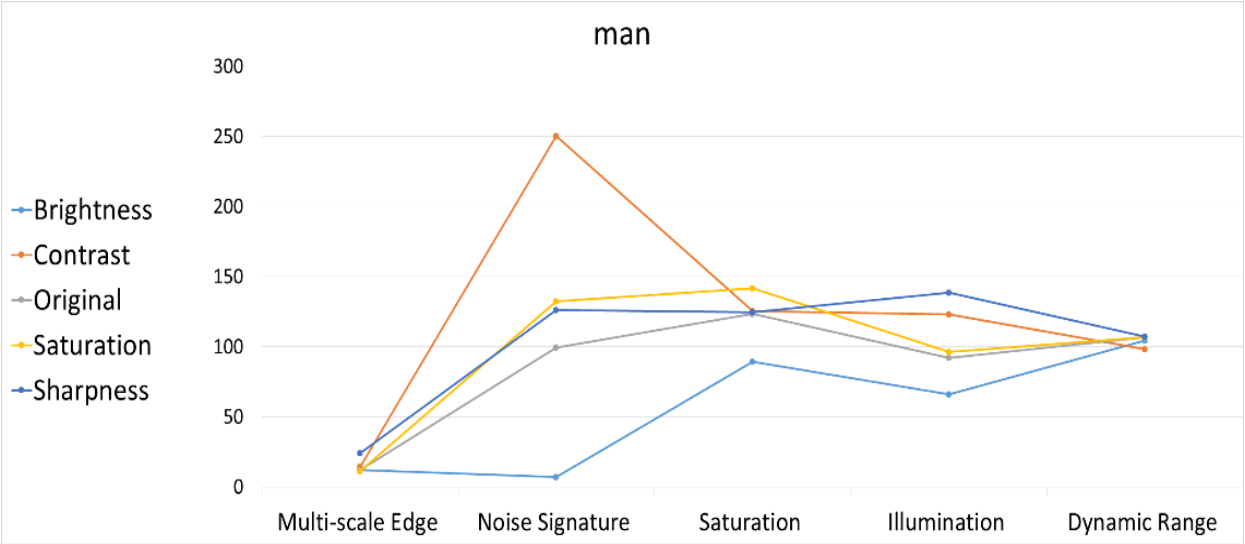


Fig.2.25. Relationship between VIQET IQ scores of “man”.

Figure 2.26 presents the high contribution of contrast and sharpness to the IQ criteria values.

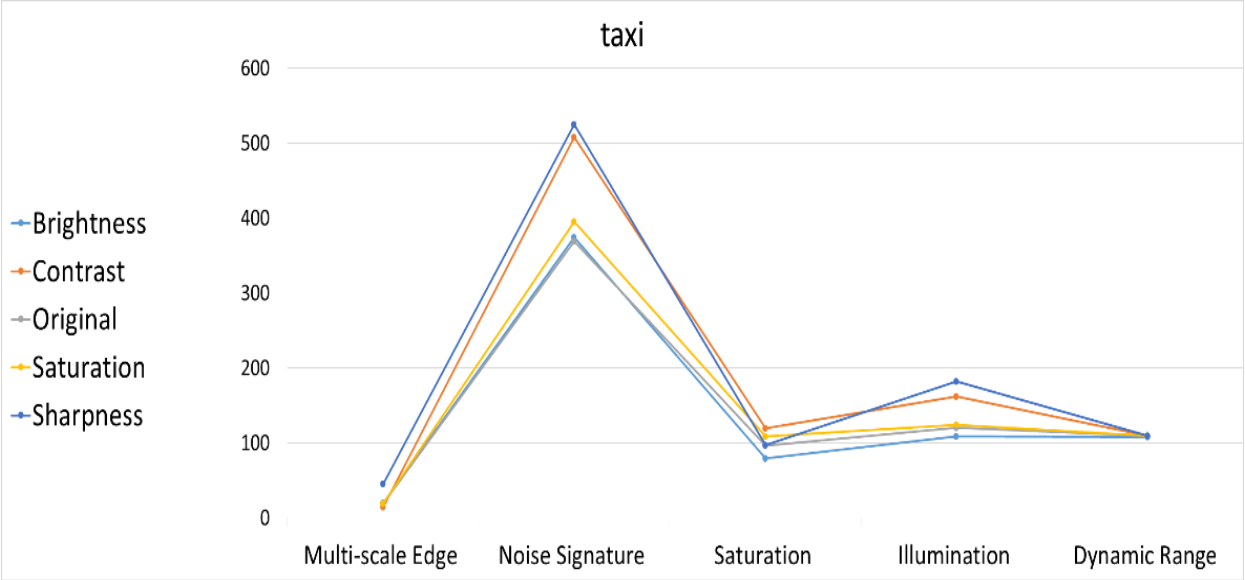


Fig.2.26. Relationship between the VIQET IQ scores of “taxi”.

Figure 2.27 presents the MOS, MSE and PSNR values and their relationship for the “room” image. The contrast has the majority in contribution to the perceived IQ.

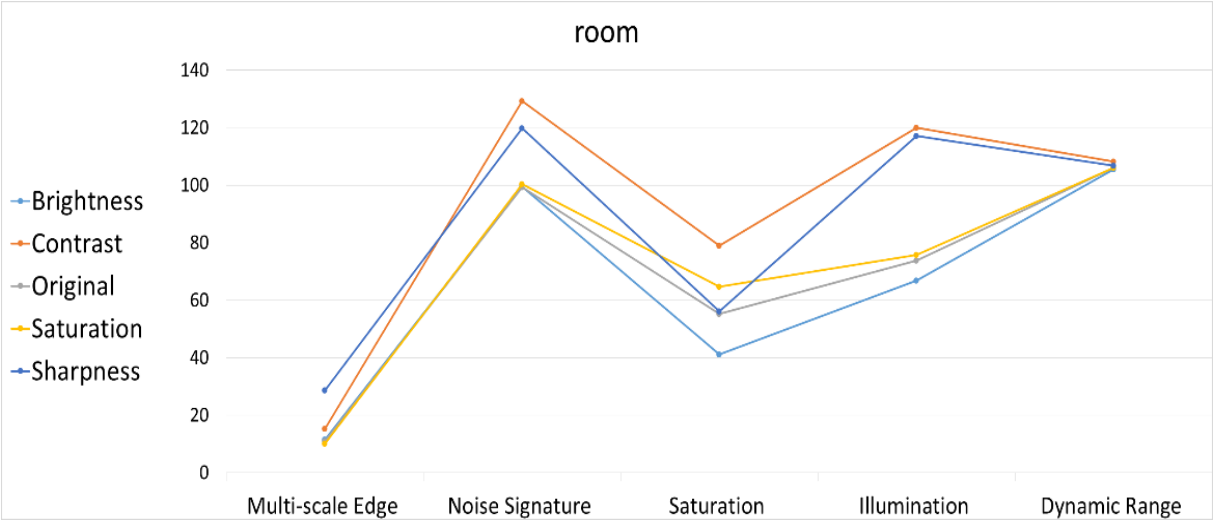


Fig.2.27. Relationship between the VIQET IQ scores of “room”.

Figure 2.28 presents the MOS, MSE and PSNR values and their relationship for the “king” image. In this indoor image the contrast and sharpness dominate the image and perform the highest value of noise signature.

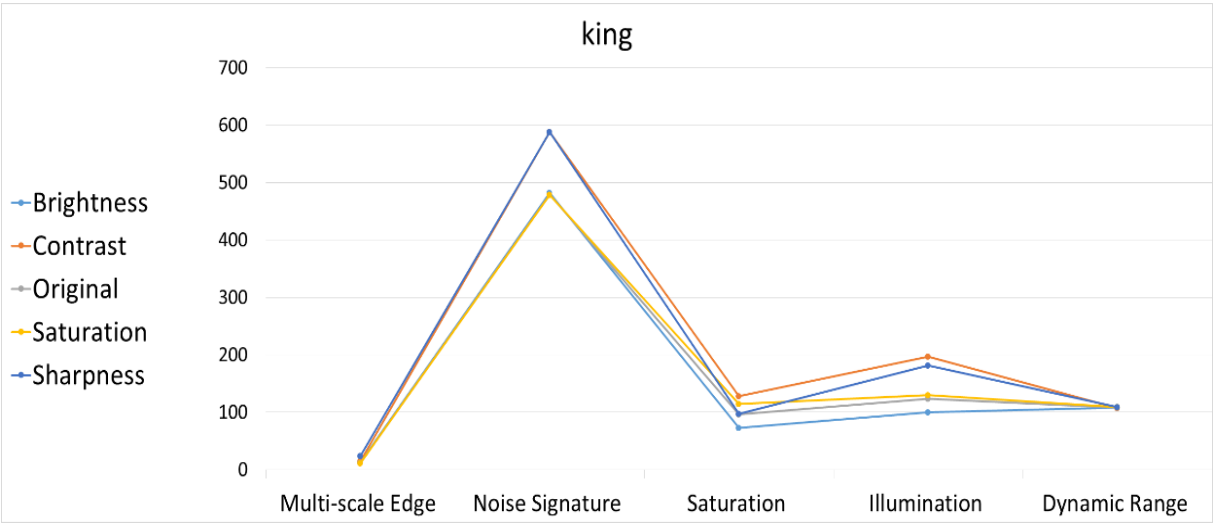


Fig.2.28. Relationship between the VIQET IQ scores of “king”.

Figure 2.29 presents the MOS, MSE and PSNR values and their relationship for the “hall” image. This image is indoor view where the contrast and sharpness are dominating the image.

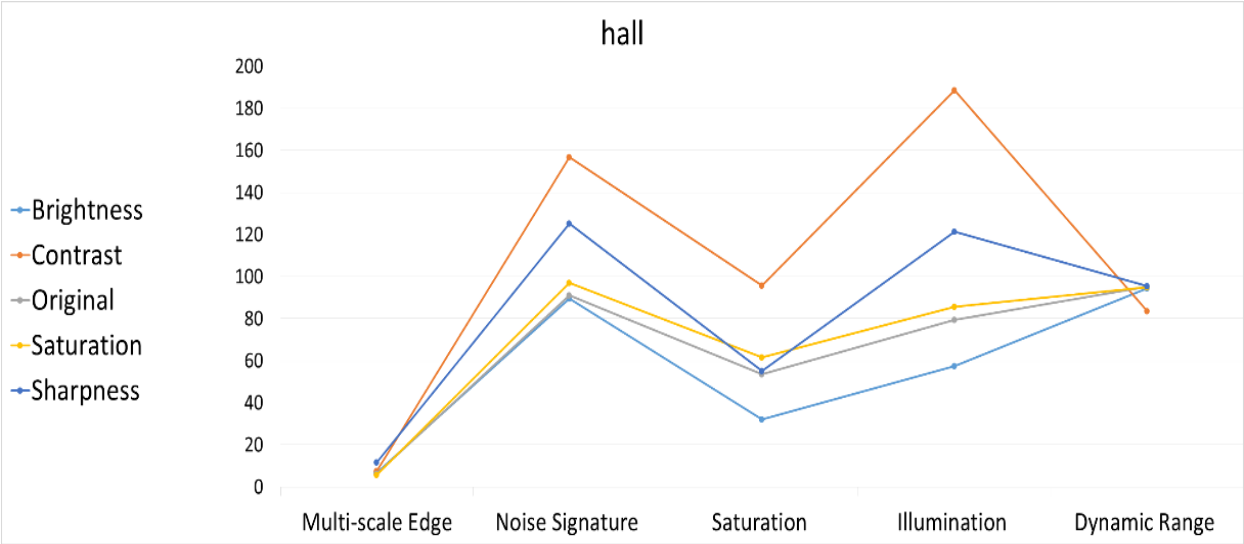


Fig.2.29. Relationship between the VIQET IQ scores of “hall”.

Figure 2.30 presents the MOS, MSE and PSNR values and their relationship for the “sunset” image. Similar IQ attributes value, no significant results.

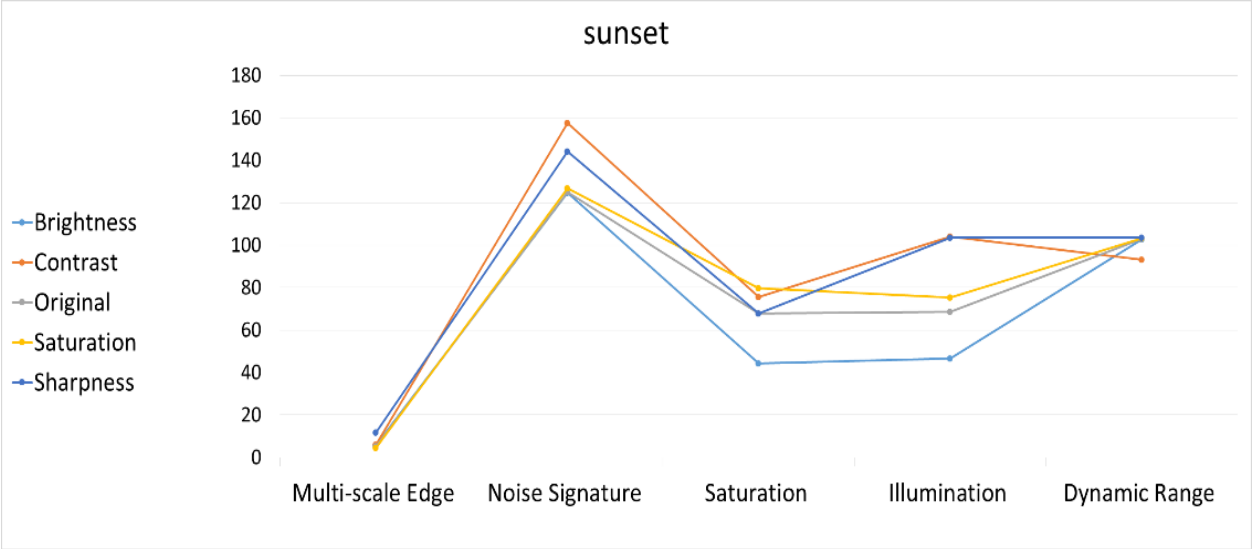


Fig.2.30. Relationship between the VIQET IQ scores of “sunset”.

Figure 2.31 presents the MOS, MSE and PSNR values and their relationship for the “bar” image. This image is “flat”, therefore all the IQ attributes have low levels except of the color saturation.

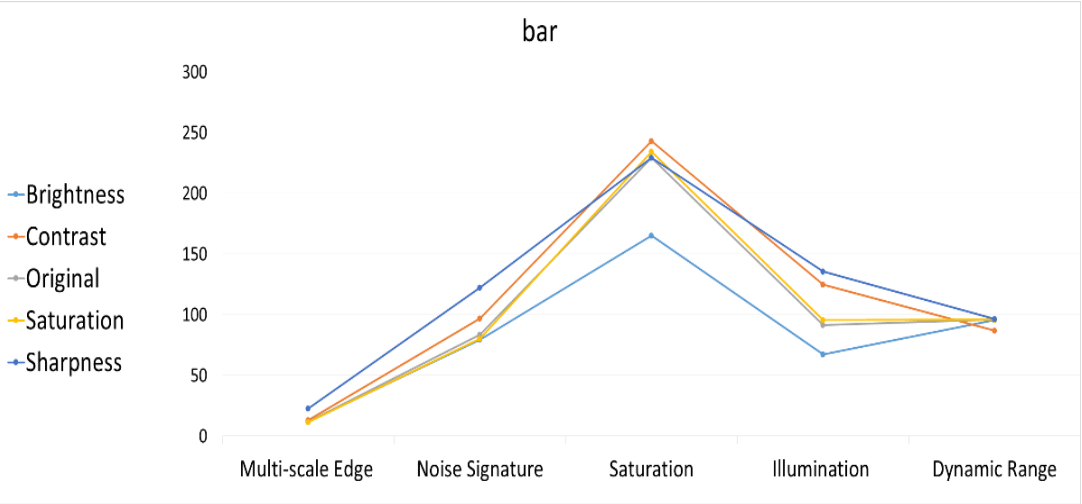


Fig.2.31. Relationship between the VIQET IQ scores of “bar”.

Figure 2.32 presents the MOS, MSE and PSNR values and their relationship for the “airplane” image. In this outdoor night image the contrast and sharpness dominating the image and increase the noise signature and illumination.

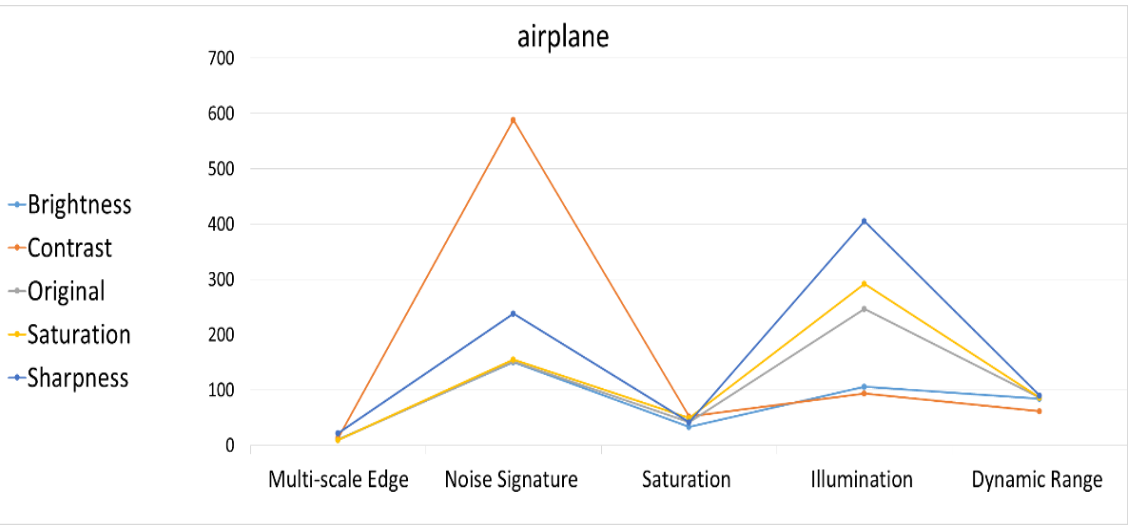


Fig.2.32. Relationship between the VIQET IQ scores of “airplane”.

Root Mean Square Error

The accuracy of the objective metric is evaluated using the RMSE evaluation metric. The difference between measured and predicted DMOS is defined as the absolute prediction error Perror:

$$Perror(i) = Score(i) - MOS \tag{2.6}$$

Where the index i denotes the image sample. While score (i) is the score gave by observer in HVT

and MOSp is the predicted MOS (which is the average of all observers' scores). The root-mean-square error of the absolute prediction error Perror is calculated with the formula:

$$RMSE = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{i=N} perror(i)^2\right)} \quad (2.7)$$

Where N denotes the total number of images considered in the analysis. The accuracy and signification of the MOS results of the subjective IQ assessment during the visual tests sessions were examined using RMSE formula. The RMSE values are demonstrated in Table 2.5.

Table 2.5. Results of the accuracy and signification calculation using RMSE formula

IQ attribute	Brightness	Contrast	Original	Color saturation	Sharpness
Building	0.86	0.72	0.73	0.55	0.53
Lake	0.98	0.69	0.63	0.48	0.54
Man	0.74	0.55	0.61	0.44	0.46
Taxi	0.87	0.59	0.81	0.52	0.47
Room	0.64	0.55	0.68	0.50	0.53
King	0.71	0.60	0.72	0.55	0.49
Hall	0.87	0.65	0.84	0.49	0.58
Bar	0.84	0.61	0.77	0.60	0.52
Sunset	0.81	0.61	0.72	0.55	0.44
Airplane	0.88	0.56	0.72	0.49	0.44

Table 2.6 presents the results of the VIQET IQ evaluation according to its IQ criteria. The accuracy and signification of the MOSp results of the VIQET IQ evaluation were examined using RMSE formula. The RMSE values are demonstrated in Table 2.7.

Table 2.6. Results of the VIQET image quality analysis

IQ attribute	Multi-scale Edge Acutance	Noise Signature Index	Saturation	Illumination	Dynamic Range
Brightness	13.09	170.33	68.81	81.56	101.19
Contrast	15.71	259.47	115.78	143.20	95.95
Original	13.11	185.46	95.86	112.56	102.32
Saturation	12.23	196.91	112.25	120.26	102.43
Sharpness	27.74	236.84	96.40	173.21	103.28

Table 0.7. Results of the accuracy and signification calculation using RMSE formula

IQ attribute	Multi-scale Edge Acutance	Noise Signature Index	Saturation	Illumination	Dynamic Range
Brightness	15.28	447.74	117.80	68.73	22.11
Contrast	23.38	653.78	156.83	108.77	45.81
Original	25.17	423.56	161.52	158.08	22.69
Saturation	16.04	434.79	164.69	190.45	22.67
Sharpness	48.79	537.87	160.80	263.72	20.39

2.8.2. The impact of different IQ attributes

The impact of different IQ attributes on the perceived image quality is evaluated by analyzing the MOS rated on brightness, contrast, colors and sharpness in *phase I*. The analysis is performed individually for the same image resolution, since it is needed to avoid the impact of image resolution when investigating the impact of IQ attributes. It is noted that the image with a high quality level has been paid more attention during the analysis, while the image with medium and low quality levels are studied on the way.

The images, which can be captured by the camera on the smartphones conveniently nowadays, are firstly selected to observe the difference of the perceived image quality on the screen with different IQ attributes.

Figure 2.32 and Figure 2.33 shows the perceived image quality of ten images with different IQ attributes displayed on the screen of mobile phone. The images with IQ attributes compared to original images. It can be seen that the perceived image quality of most images can obtain a slight improvement when the IQ attributes are increased. However, this increasing trend seems to be not obvious when the brightness increased. In another word, the brightness does not guarantee significantly better users' experience. This phenomenon can also be observed in Figure 2.33 and Figure 2.36 with VIQET analysis results.

It can be found that the perceived image quality of some images even decreases when the brightness is increasing. Consequently, when the subjects view the high quality images with the IQ attributes it seems that the subjects may not perceive a higher image quality.

It can be seen that there is no decrease of the perceived quality of any image when increasing the Contrast and Sharpness. This phenomenon indicates that the Contrast and Sharpness can

provide a meaningful gain on user's perceived image quality of the images. This improvement brought by the IQ attributes can be distinguished by the viewers.

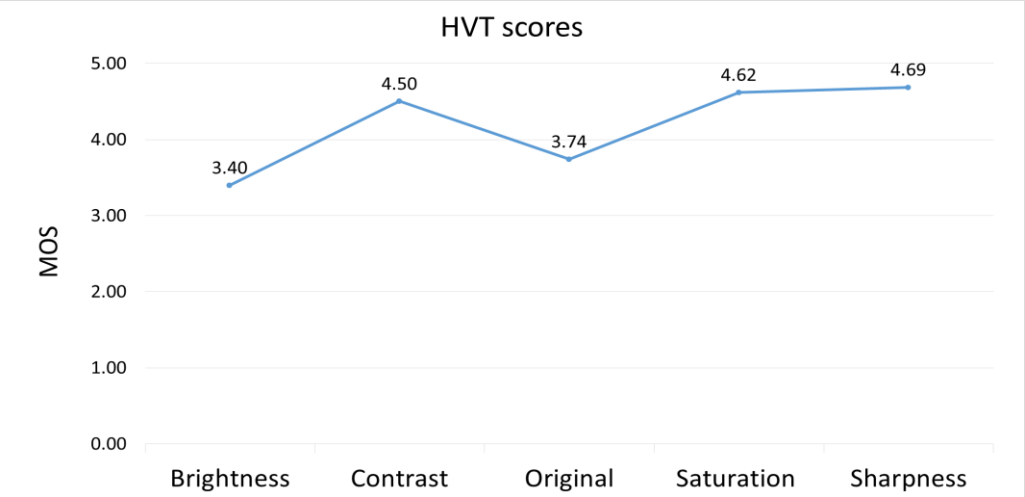


Fig.2.33. MOS of ten high quality images with different IQ attributes.

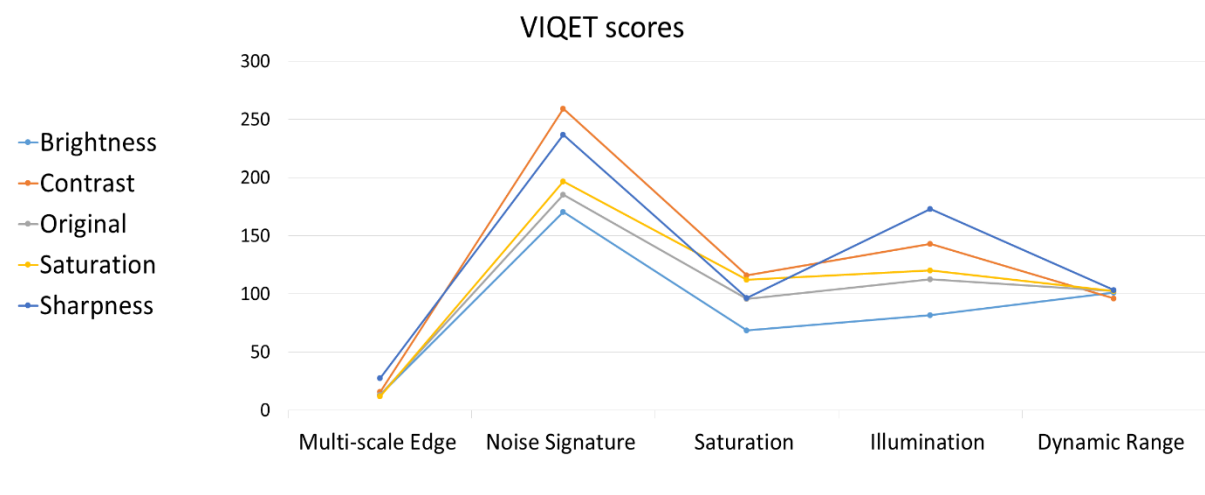


Fig.2.34. VIQET scores of ten high quality images with different IQ attributes.

Not all images can obtain a quality increment with the IQ attributes, some have a much more decrease. For the sharpness and saturation the variation of perceived image quality of these image is similar to the contrast.

Table 2.8 lists the details of the variation of perceived quality of different image quality levels, from which it can be seen that the user's perceived image quality with brightness has no advantage compared to the original image, with different quality levels considered. The results in Table 2.9 also show this phenomenon.

Table 2.8. Comparison between perceived image quality of IQ Attributes

IQ attribute	Brightness	Contrast	Original	Color Saturation	Sharpness
MOS	3.46	4.50	3.74	4.62	4.69

Calculating DMOS Values

The data analysis was performed using the DMOS. DMOS values were calculated for each IQ attribute. DMOS values were calculated using the following formula:

$$DMOS = MOS_{iq} - MOS \quad (2.8)$$

While MOSiq is the average of MOS of IQ attribute and MOSo is the average of MOS of the original image. In using this formula, higher DMOS values indicate better quality.

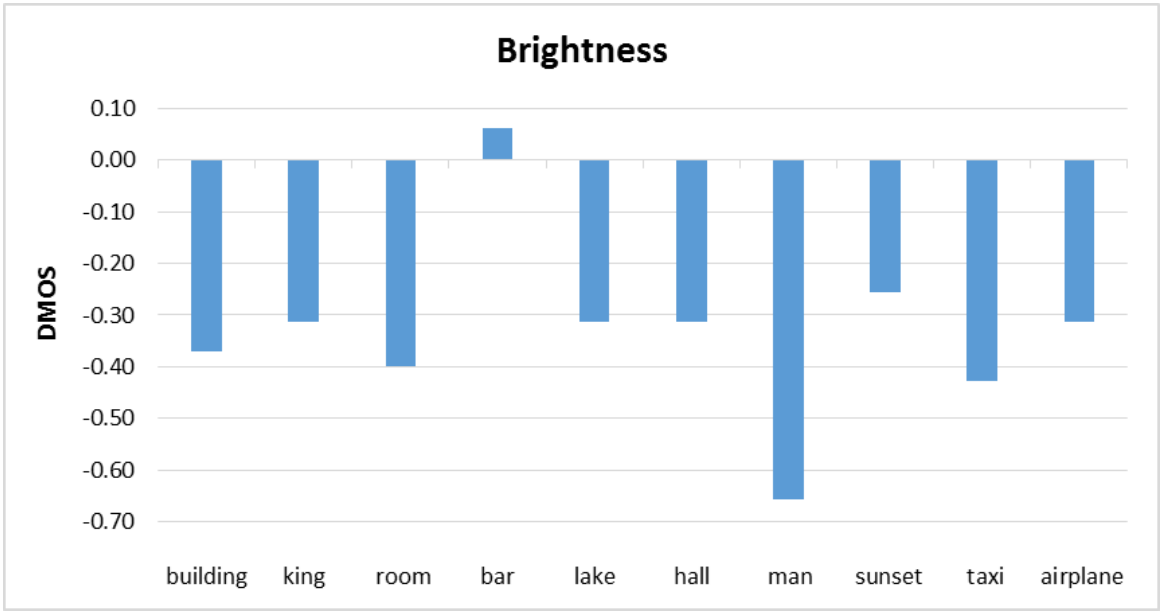


Fig.2.35. Illustration of the DMOS of brightness.

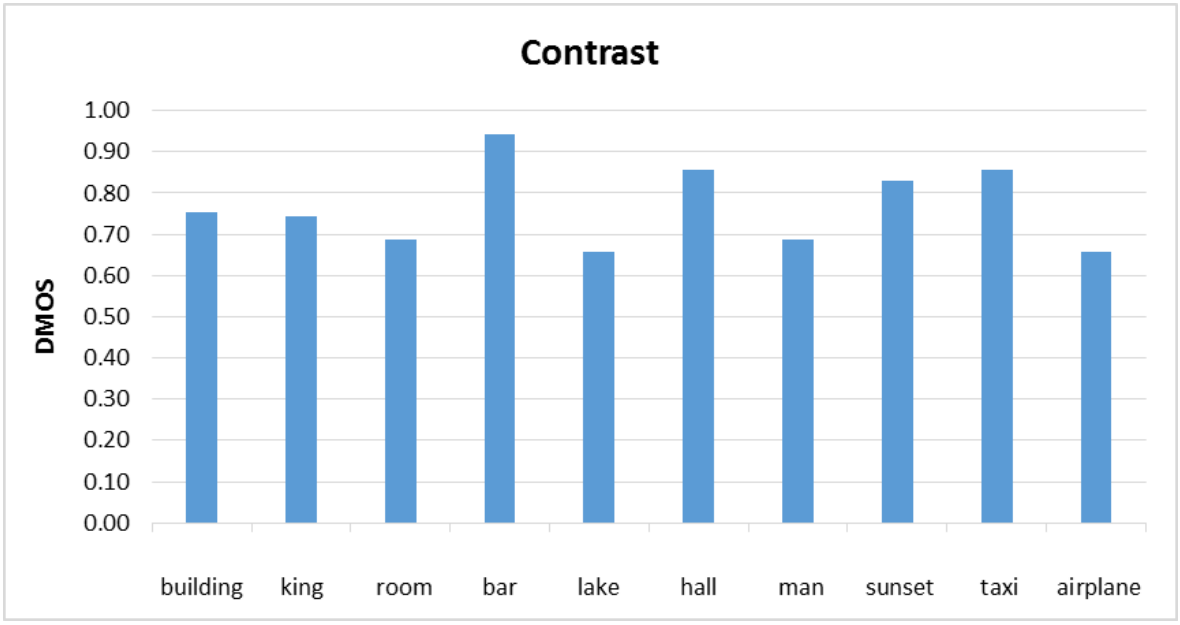


Fig.2.36. Illustration of the DMOS of contrast.

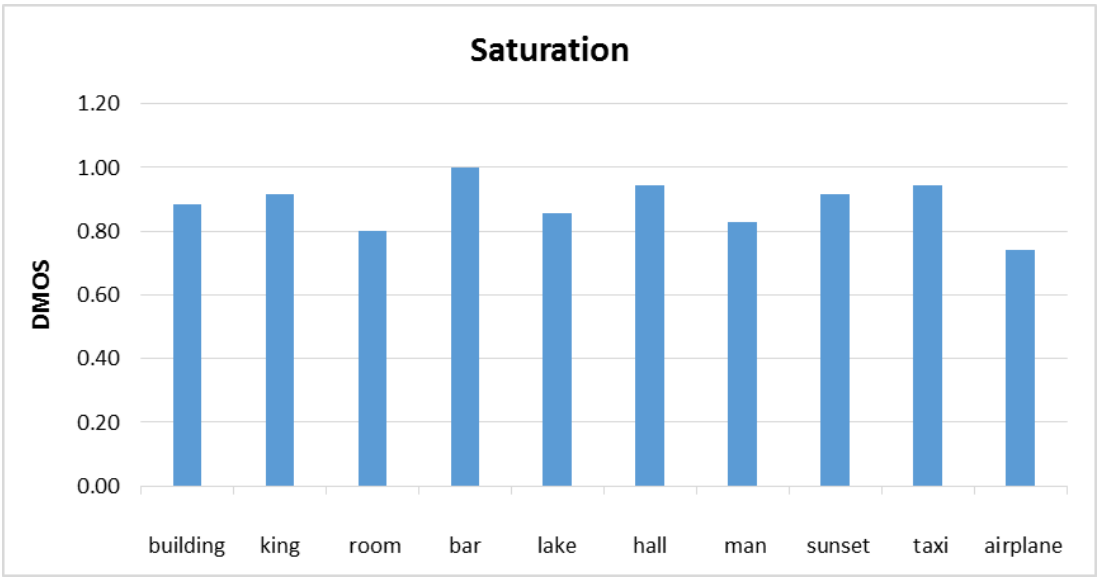


Fig. 2.37. Illustration of the DMOS of saturation.

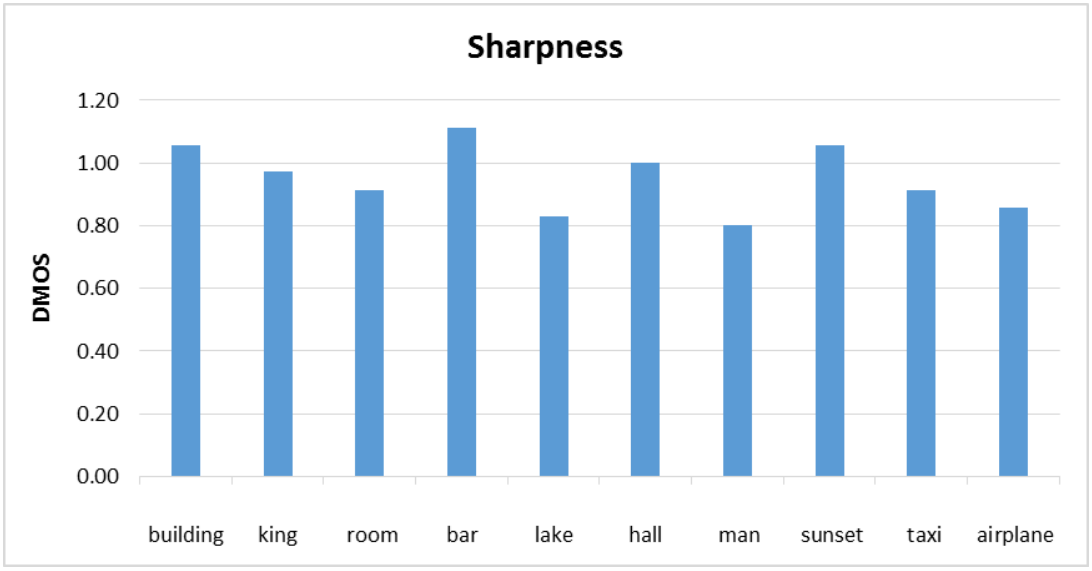


Fig.2.38. Illustration of the DMOS of sharpness.

Compared with the perceived image quality with the contrast, a slight improvement on perceived image quality is obtained when the images are with higher sharpness and saturation as shown in Figure 2.37 and Figure 2.38.

Table 2.9. DMOS calculations of IQ attributes

IQ attribute	Brightness	Contrast	Color saturation	Sharpness
Building	-0.37	0.75	0.89	1.06
Lake	-0.31	0.66	0.86	0.83
Man	-0.66	0.69	0.83	0.80
Taxi	-0.43	0.86	0.94	0.91
Room	-0.40	0.69	0.80	0.91
King	-0.31	0.74	0.91	0.97
Hall	-0.31	0.86	0.94	1.00
Bar	0.06	0.94	1.00	1.11
Sunset	-0.26	0.83	0.91	1.06
Airplane	-0.31	0.66	0.74	0.86

Table 2.9 presents the DMOS values of ten images with different IQ attributes. Higher values mean better image quality. Sharpness, color saturation and contrast received the highest values respectively.

Statistical significance analysis

The performance of each objective quality model was characterized by three prediction attributes: accuracy, monotonicity and consistency. The statistical metrics RMSE, Pearson correlation, and outlier ratio together characterize the accuracy, monotonicity and consistency of a model's performance. These statistical metrics are named evaluation metrics in the following.

The calculation of each evaluation metric is performed along with its 95% confidence intervals. To test for statistically significant differences among the performance of various models, a test based on the RMSE, tests based on approximations to the Gaussian distribution were constructed for the Pearson correlation coefficient and the Outlier Ratio.

The evaluation metrics were calculated using the objective model outputs and the results from viewer subjective rating of the test video clips. The objective model provides a single number (figure of merit) for every tested images. The same tested video clips get also a single subjective figure of merit. The subjective figure of merit for an image represents the average value of the scores provided by all subjects viewing the image.

The evaluation analysis is based on DMOS scores for the RR models, and on MOS scores for the NR model. Discussion below regarding the DMOS scores was applied identically to MOS scores. For simplicity, only DMOS scores are mentioned for the rest of the chapter. The objective quality model evaluation was performed in three steps. The first step is a mapping of the objective data to the subjective scale. The second calculates the evaluation metrics for the models and their confidence intervals. The third tests for statistical differences between the evaluation metrics value of different models.

2.8.3. Evaluation of smartphone perceived image quality

In this section, four impact factors, that is, brightness, contrast, saturation and sharpness are investigated to establish an objective quality assessment model. Here, the influence of each IQ attribute itself will be checked at first. Then, the mutual interaction of these four impact factor is evaluated.

In order for the observers to use a sufficiently large set of IQ attributes, a broad range of images should be used in order to reveal different quality issues. To achieve this we followed the recommendations of VQEG, where the images were modified based on the following criteria:

- Higher levels of brightness (lightness).
- Higher levels of contrast.
- Higher levels of color saturation.
- Higher levels of sharpness.

Most of the images were pictorial with a wide range of scenes like landscapes, portraits, and personal items. This helps to characterize the impacts for IQ attributes and ensures that the observers examine a wide variety of IQ attributes. In addition to the pictorial images a set of test charts were included, since these are content-free and have a selection of ”interest area” colors suitable for evaluation of different aspects of IQ. A total of 10 images, as seen in Figure 2.2 were used in this experiment.

- Outdoor day- landscape, people.
- Indoor – without backlight.
- Indoor – with backlight.
- Outdoor night.

The images captured by the author. These ten images were in original format which were captured by the smartphone camera. Images IQ attributes were changed by Photoshop. The images were captured at a resolution of 8 Mega pixels.

2.8.4. Perceived image quality based on IQ attributes

As the basic quality of the perceived image quality, the original image used as a reference in order to investigate the influence of IQ attributes on the perceived image quality. The MOS of five IQ attributes that are selected to check the relationship between the IQ attributes and the MOS. The interactive impact between IQ attributes can be investigated on the way.

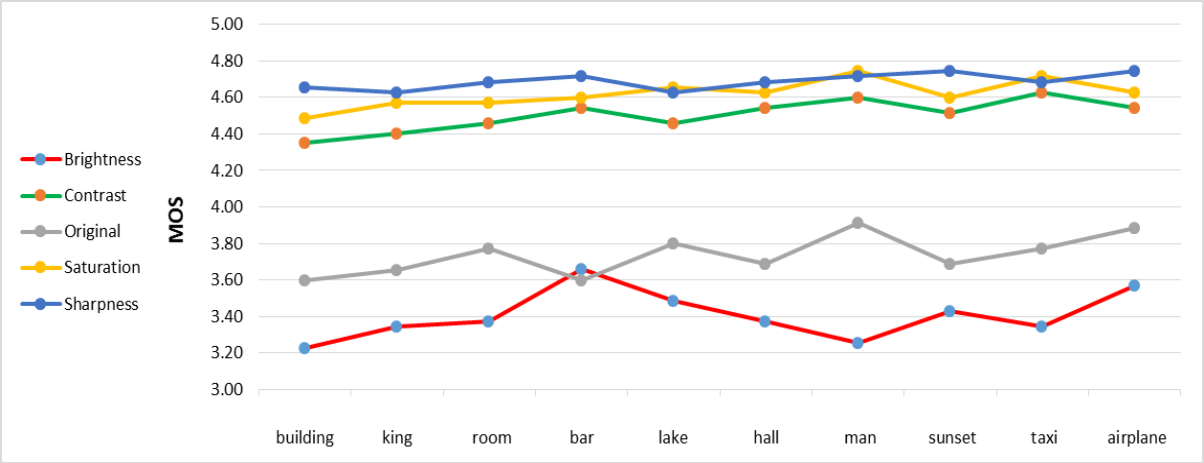


Fig.2.39. MOS of ten images with IQ attributes that scored on the HVT

Figure 2.39 takes this relationship of the ten images as an example. For the IQ attributes the values of MOS are increasing with the increment of the contrast (green curve), saturation (yellow curve) and sharpness (blue curve).However, when the brightness (red curve) is higher than of the original image the value of MOS will decrease. Likewise, for the images with higher contrast and saturation the values of MOS also keep nearly constant.

Consequently, a preferable perceived image quality can be provided by improving the image contrast, saturation and sharpness, but this improvement of perceived image quality will be limited by the IQ attributes level.

To reflect the limitation caused by the IQ attributes on the perceived image quality, an integrated assessment parameter, that is, the level of IQ attributes added to the original images is proposed.

The IQ attributes levels were increased and decreased in several levels but didn’t reach the saturation levels in order to keep image quality in normal level.

VIQET image analysis

The VIQET was used for images analysis and IQ features measurement for four IQ attributes that are selected to check the relationship between the IQ attributes and the VIQET image quality criteria.

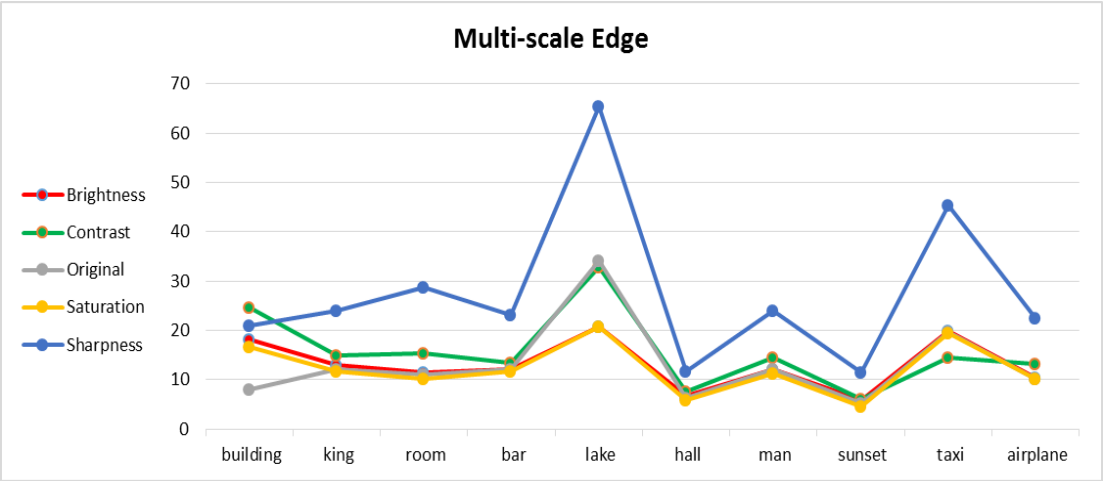


Fig.2.40. Multi-scale edge levels of ten images with IQ attributes by VIQET.

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. Figure 2.40 takes this relationship of the ten images as an example. For the IQ attributes the values of multi-scale edge are increasing with the increment of the sharpness (blue curve) and contrast (green curve).

However, higher brightness (red curve) and saturation (yellow curve) have the same level as the original image (gray curve).

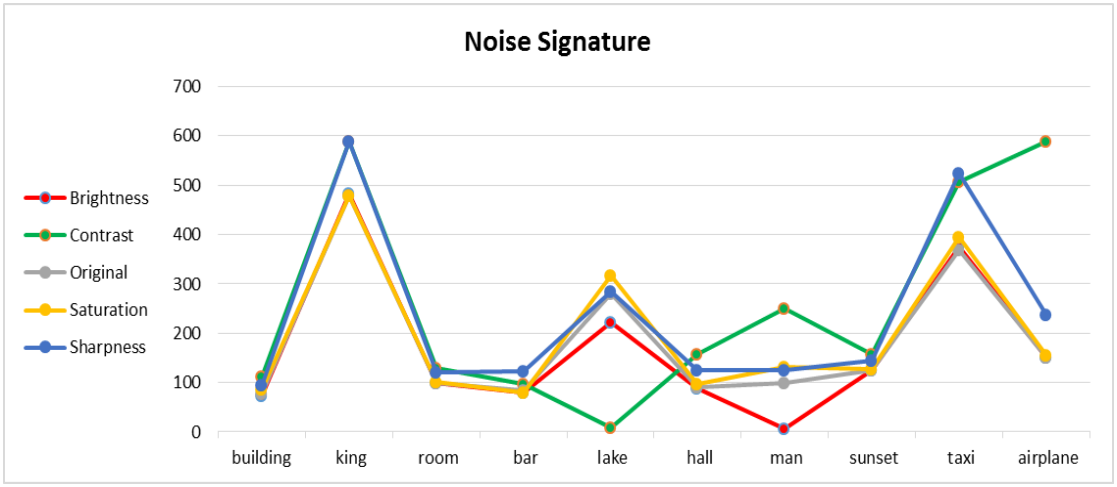


Fig.2.41. Noise Signature levels of ten images with IQ attributes by VIQET.

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.

Figure 2.41 illustrates the relationship between IQ attributes and noise signature of the ten images. For the IQ attributes the values of noise signature index are increasing or decreasing randomly depends on the image content.

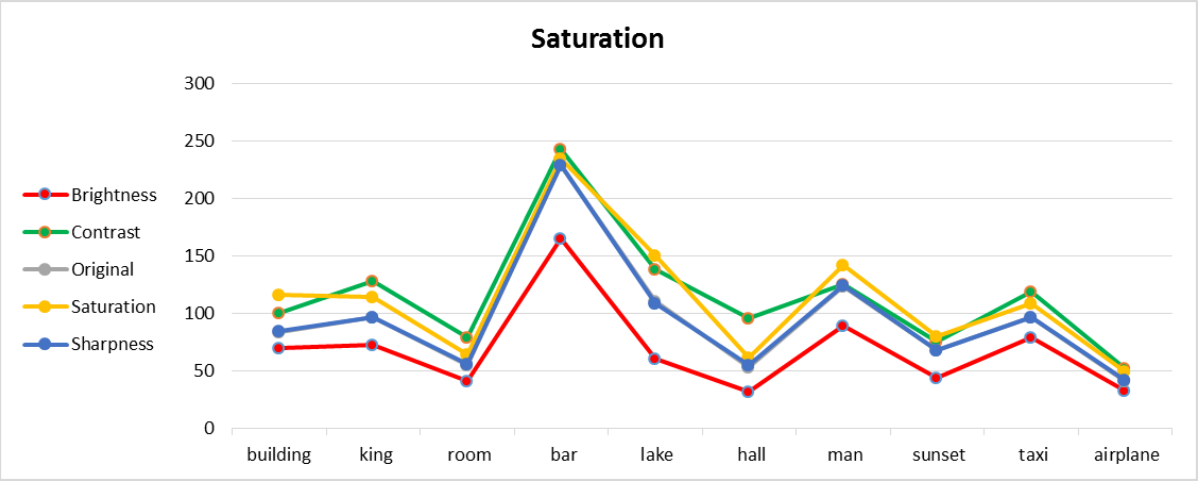


Fig.2.42. Saturation levels of ten images with IQ attributes by VIQET.

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.

Figure 2.42 illustrates the relationship between IQ attributes and Saturation of the ten images. For the IQ attributes, the values of Saturation are increasing with the increment of the color saturation (yellow curve), sharpness (blue curve) and contrast (green curve).

However, higher brightness (red curve) does not effect the saturation level.

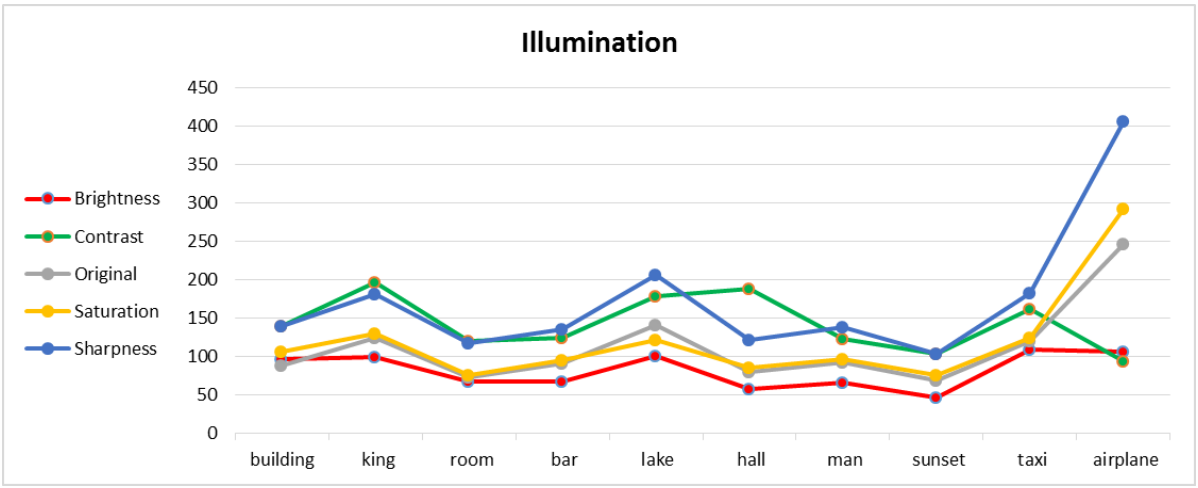


Fig.2.43. Illumination levels of ten images with IQ attributes by VIQET.

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.

Figure 2.43 illustrates the relationship between IQ attributes and Illumination of the ten images as an example. For the IQ attributes the values of illumination are increasing with the increment of the sharpness (blue curve) and Contrast (green curve).

However, higher brightness (red curve) and saturation (yellow curve) doesn't effect the illumination level.

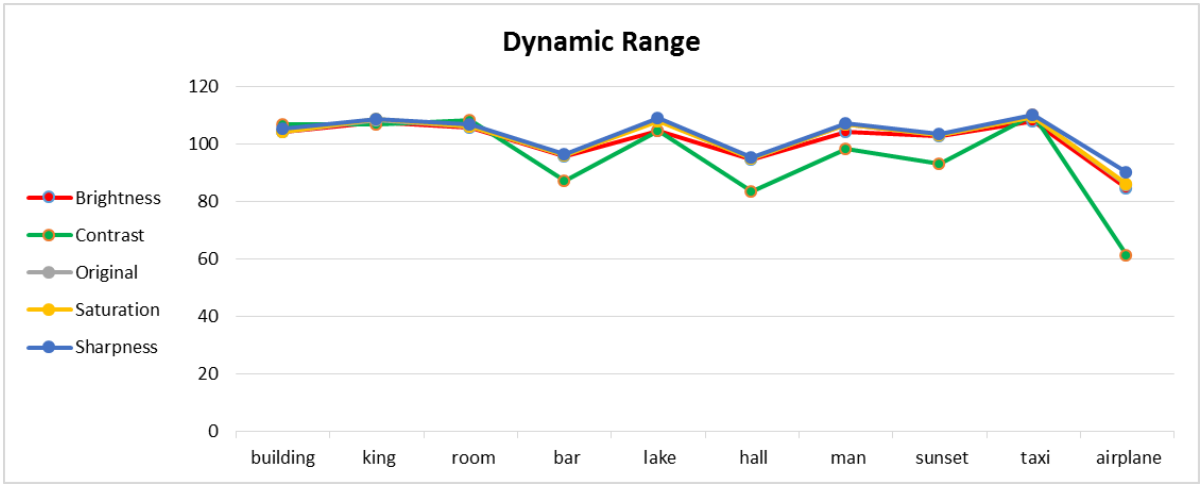


Fig.2.44. Dynamic range levels of ten images with IQ attributes by VIQET.

Dynamic Range is the range between the lightest and darkest regions in an image maintaining details of an image in both the lightest and darkest spots (represented in shades of grey). Figure 2.44 illustrates the relationship between IQ attributes and dynamic range of the ten images. For the IQ attributes the values of dynamic range are almost the same with the increment of the sharpness (blue curve), brightness (red curve) and saturation (yellow curve). Dynamic range was decreased when increasing the contrast (green curve) of the mage.

2.9. Summary

Chapter 2 describes the main work of this research towards modeling perceived IQ of smartphones. The IQ evaluation also shows that method is very compatible with human assessment scores for all different types of images and various IQ attributes.

A wide range of images captured by mobile phones were selected in the subjective experiments to investigate the impact of IQ attributes on user's perceived image quality.

The outcome of this work are:

- New model which has been developed through many experiments of image quality visual tests;
- The new model able to predict the expected perceived IQ of smartphones using only VIQET instead of complicated and expensive visual experiments;
- IQ experiment has been carried out to identify the most effective IQ attributes. Results obtained from the experiment support the proposed set of IQ attributes;
- New process and framework for perceived IQ prediction;

The work was divided into two main phases. In *phase I* the perceived IQ of subjective assessment was performed through visual experiments. The objective of *phase I* was to identify the most important IQAs that make the image pleasant for humans. In *phase II* the images were evaluated by VIQET while the IQ criteria were calibrated according to the new model. The calibration based on relationship between IQAs and VIQET IQ criteria.

As the high correlation between the HVTs results and VIQET results achieved, the new model approval is required, then, the new model was examined with large group of viewers.

In *phase I,* performing an experimental study of perceived image quality of smartphones with controlled IQ attributes: brightness, contrast, color saturation and sharpness. Experiments were based on HVT on smartphone display.

In *phase II,* results of *phase I* and *phase II* were analyzed , and the VIQET IQ attributes coefficients were calibrated. In the end of the experimental process the VIQET got calibrated and the tool image quality scores are a good reference for perceived image quality prediction. This well described and examined in the next chapter.

III. NEW MODEL PERFORMANCE EVALUATION

"It would be possible to describe everything scientifically, but it would make no sense; it would be without meaning, as if you described a Beethoven symphony as a variation of wave pressure."

(Albert Einstein)

Chapter 3 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 3.1, showing scatter plots of the perceived image quality versus the predicted image quality.

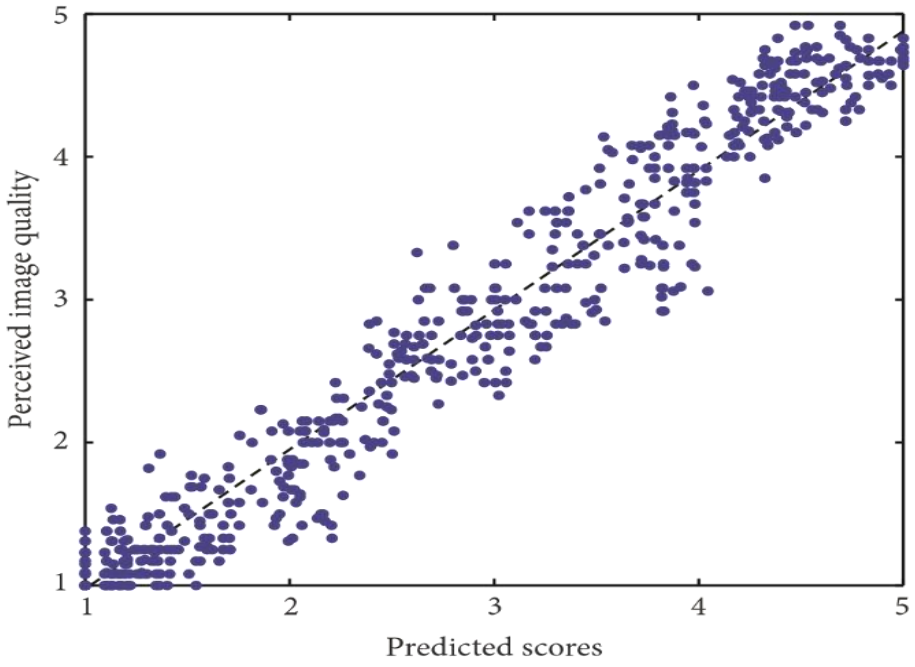


Fig.3.1. An example of perceived image quality and predicted scores [2].

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.

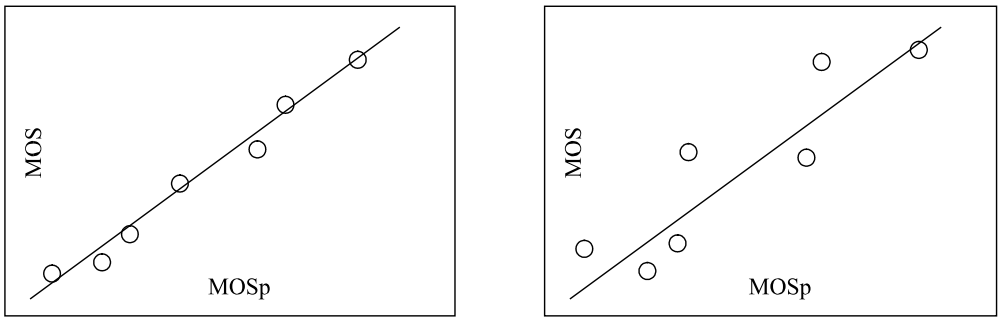
The VIQET correlation coefficients will be used to prediction the perceived smartphone image quality. To evaluate the performance of the proposed quality assessment model, this study followed the standard performance evaluation procedures of VQEG [6]. The standard was developed for calculating the prediction error between a mathematical model and subjective scores (i.e., human viewers' opinion).

3.1. VQEG's standard performance evaluation procedures

According to the VQEG [6], the performance of an objective quality model is characterized by three prediction attributes:

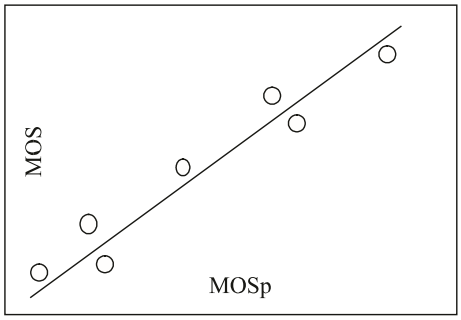
Accuracy — is the ability to predict the distortions between MOS and MOSp. In an ideal case, the relationship between the MOS and MOSp is expected to be linear. Figure 3.2 illustrates the hypothetical relationships between the MOS and the MOSp of two models. Model-I is more accurate than the Model-II because most of the images evaluations are reasonably closer to the straight line.

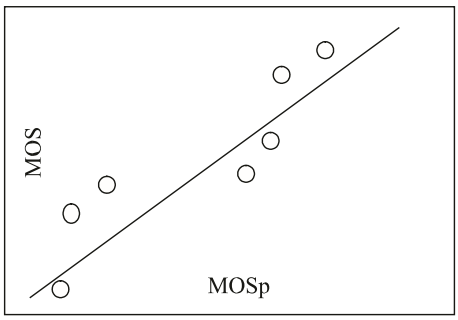
Monotonicity — is the degree to which the model's predictions agree with the relative magnitudes of subjective quality ratings. The prediction monotonicity is the extent of agreement between the subjective test and the objective model of variations in picture quality. As an example, viewers rank image 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.3.2. Two hypothetical models with different prediction accuracy [95].

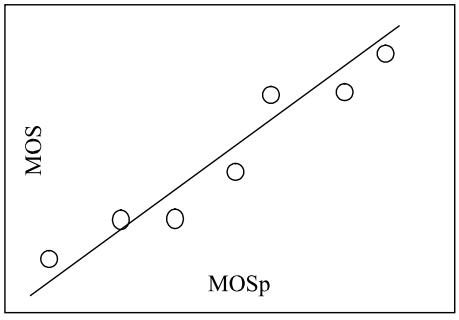


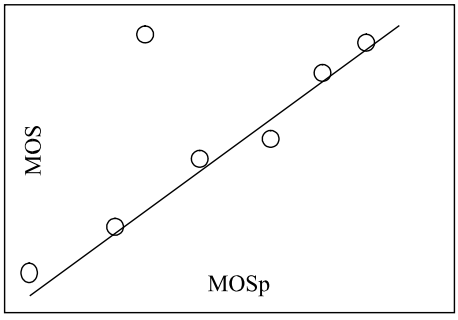


Model I (monotonic) Model II (non-monotonic)

Fig. 3.3. Two hypothetical models with different prediction monotonicity [95].

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 3.5 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.3.4. Two hypothetical models with different prediction consistency [95].

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

Metric 1*:* Pearson correlation coefficient (CC) between objective MOSp and subjective MOS scores. MOSp is the mean opinion score prediction that is the output of objective (i.e., mathematical) model and whereas MOS is the mean opinion score of human assessments. This metric provides an evaluation of *prediction accuracy*, *which* can be defined as:

$$CC = \frac{\sum_{i=1}^{N}\left(MOS(i)-\overline{MOS}\right)\left(MOS_p(i)-\overline{MOS_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(MOS(i)-\overline{MOS}\right)^2}\sqrt{\sum_{i=1}^{N}\left(MOS_p(i)-\overline{MOS_p}\right)^2}} \tag{3.1}$$

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

The Pearson correlation coefficient is used to measure the strength of a linear association between two variables, where the value *r* = 1 means a perfect positive correlation and the value *r* = -1 means a perfect negative correlation. So, for example, you could use this test to find out whether people's height and weight are correlated (they will be - the taller people are, the heavier they are likely to be).

Requirements for Pearson's correlation coefficient:

- Scale of measurement should be interval or ratio
- Variables should be approximately normally distributed
- The association should be linear
- There should be no outliers in the data

Metric 2: Spearman rank order correlation coefficient (SROCC) between objective MOSp and subjective MOS scores. It is considered as a measure of *prediction monotonicity* and it is defined by:

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

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

The prediction monotonicity is the extent of agreement between the subject test and the objective model in terms of the sign of change in picture quality.

Spearman's correlation coefficient is a statistical measure of the strength of a monotonic relationship between paired data. In a sample, it is denoted by and is by design constrained as follows and its interpretation is similar to that of Pearson's, e.g. the closer is to the stronger the monotonic relationship. Correlation is an effect size and so we can verbally describe the strength of the correlation using the following guide for the absolute value of:

- .00-.19 "very weak"
- .20-.39 "weak"
- .40-.59 "moderate"
- .60-.79 "strong"
- .80-1.0 "very strong"

Metric 3: Outlier ratio (OR) represents number of "outlier-points" to the total points N. It is considered as a measure of *prediction consistency*, which can be defined by the following equation:

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

where an outlier is a point for: $|MOS(i) - MOS_p(i)| > 2\times\sigma(\text{MOS(i)})$, where σ(MOS(i)) represents the standard deviation of the individual scores associated with the image sample *i*. The individual scores are approximately normally distributed and therefore twice the σ value represents the 95% confidence interval.

Thus, $2\times\ \sigma$(MOS(i)) value represents a good threshold for defining an outlier point. It is desirable that an objective model performs well over a wide range of test conditions. In other words, besides having the prediction error as small as possible, consistency in the magnitude of the prediction error is also preferable.

Metric 4: Average/Mean Absolute Error (MAE)

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:

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

As the name suggests, the mean absolute error is an average of the absolute errors where is the prediction and the true value. Note that alternative formulations may include relative frequencies as weight factors.

The mean absolute error is on same scale of data being measured. This is known as a scale-dependent accuracy measure and therefore cannot be used to make comparisons between series on different scales.

The mean absolute error is a common measure of forecast error in time series analysis, where the terms "mean absolute deviation" is sometimes used in confusion with the more standard definition of mean absolute deviation. The same confusion exists more generally.

The mean absolute error is one of a number of ways of comparing forecasts with their eventual outcomes. Well-established MASE and the MSE.

These all summarize performance in ways that disregard the direction of over- or under-prediction. a measure that does place emphasis on this is the mean signed difference.

Where a prediction model is to be fitted using a selected performance measure, in the sense that the least squares approach is related to the mean squared error, the equivalent for mean absolute error is least absolute deviations.

Metric 5: RMSE between objective MOSp and subjective MOS scores is defined by:

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

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

An excellent objective model should 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.

3.2. 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 [6]. 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 forty 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. Figure 3.5 and Figure 3.6 demonstrates the VIQET user interface window with the loaded images and their scores calculated by the tool according to the mathematical model.

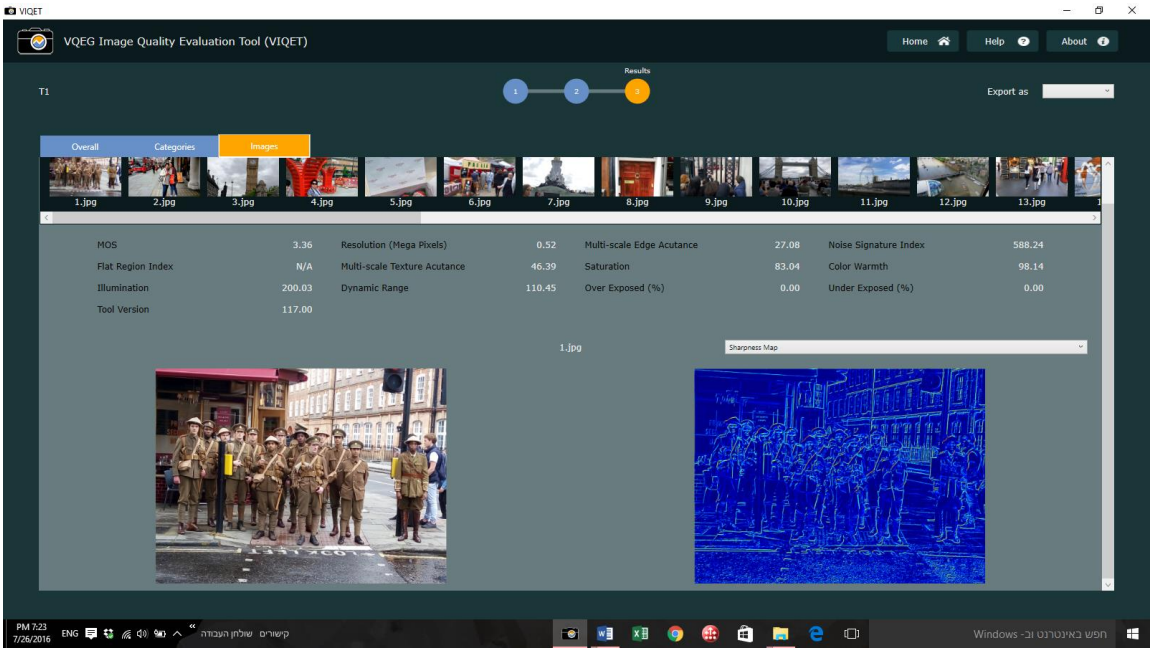


Fig.3.5. VIQET image quality analysis scores and sharpness map.

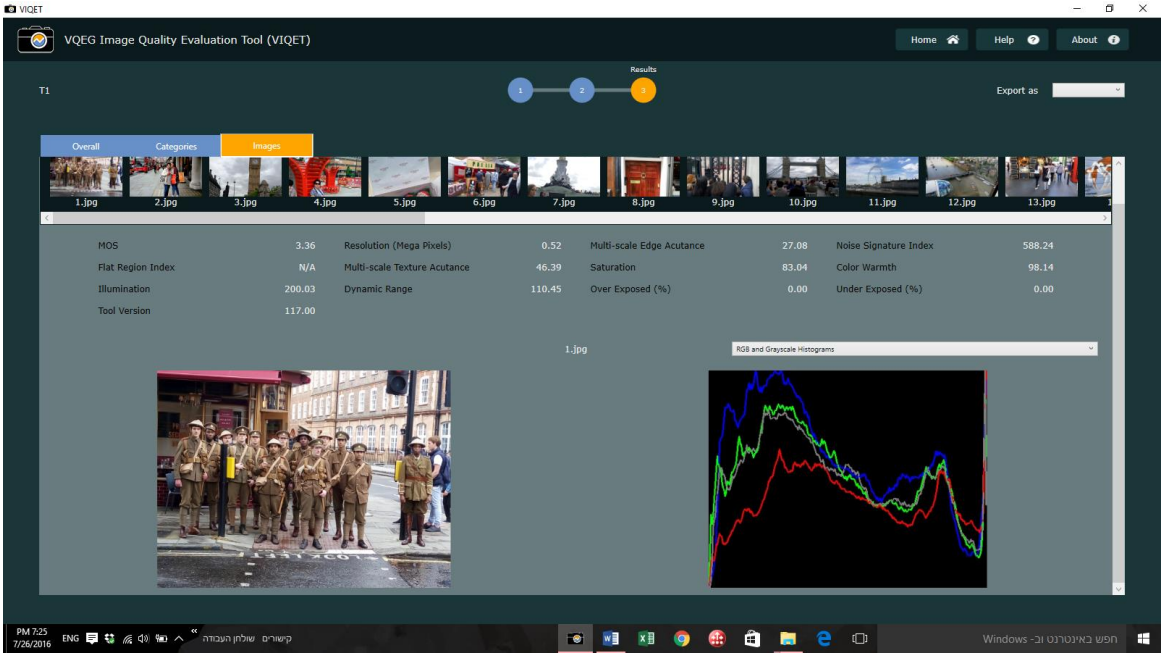


Fig.3.6. VIQET image quality analysis scores and RGB histogram.

3.2.1. 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. The MOSp was calculated according to the mathematical model, which is an outcome of this thesis based on HVT.

Table 3.1. VIQET scores and MOSp of the outdoor day images

Filename	MOSp	Category
1.jpg	3.91	Outdoor day
2.jpg	4.03	Outdoor day
3.jpg	3.83	Outdoor day
4.jpg	4.3	Outdoor day
5.jpg	4.23	Outdoor day
6.jpg	4.5	Outdoor day
7.jpg	4.14	Outdoor day
8.jpg	4.5	Outdoor day
9.jpg	4.15	Outdoor day
10.jpg	4.46	Outdoor day
11.jpg	4.4	Outdoor day
12.jpg	3.61	Outdoor day
13.jpg	3.84	Outdoor day
14.jpg	4.39	Outdoor day
15.jpg	3.44	Outdoor day
16.jpg	4.05	Outdoor day
17.jpg	4.38	Outdoor day

The VIQET results of outdoor day images are presented in Table 3.1. values of image quality attributes were calculated according to the image analysis configuration in Software. The predicted MOSp of each individual image was calculated by VIQET by using the weight of each image quality attributes according to the mathematical model.

Table 3.2. VIQET scores and MOSp of the indoor images

Filename	MOSp	Category
18.jpg	3.97	Indoor
19.jpg	3.85	Indoor
20.jpg	2.97	Indoor
21.jpg	3.46	Indoor
22.jpg	3.2	Indoor
23.jpg	4.41	Indoor
24.jpg	3.02	Indoor
25.jpg	3.93	Indoor

26.jpg	3.19	Indoor
27.jpg	4.25	Indoor
28.jpg	3.8	Indoor
29.jpg	3.89	Indoor
30.jpg	3.8	Indoor

The VIQET results of Indoor images are presented in Table 3.2. Values of image quality attributes were calculated according to the image analysis configuration in Software. The predicted MOSp of each individual image was calculated by VIQET by using the weight of each image quality attributes according to the mathematical model.

Table 3.3. VIQET scores and MOSp of the outdoor night image

Filename	MOSp	Category
31.jpg	2.01	Outdoor Night
32.jpg	3.45	Outdoor Night
33.jpg	3.44	Outdoor Night
34.jpg	3.04	Outdoor Night
35.jpg	2.79	Outdoor Night
36.jpg	2.83	Outdoor Night
37.jpg	3.02	Outdoor Night
38.jpg	2.1	Outdoor Night
39.jpg	3.84	Outdoor Night
40.jpg	1.46	Outdoor Night

The VIQET results of outdoor night images are presented in Table 3.3. Values of image quality attributes were calculated according to the image analysis configuration in Software. The predicted MOSp of each individual image was calculated by VIQET by using the weight of each image quality attributes according to the mathematical model.

3.3. The new image quality assessment model evaluation

The evaluation results of five mentioned methods in above are summarized in Tables 3.4 – Table 3.8. Each table shows that the proposed model's performances for every one of the evaluation metrics are sufficient for both the training and the testing datasets.

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 3.4 and the MOSp relation to MOS is illustrated in Figure 3.7.

Table 3.4. MOS and MOSp values of the outdoor day images

Image	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
MOS(i)	4.21	3.75	3.64	4.04	4.00	4.04	3.93	4.14	3.89	4.00	3.82	3.71	4.11	4.61	3.82	3.79	3.96
MOSp(i)	3.91	4.03	3.83	4.30	4.23	4.50	4.14	4.50	4.15	4.46	4.40	3.61	3.84	4.39	3.44	4.05	4.38

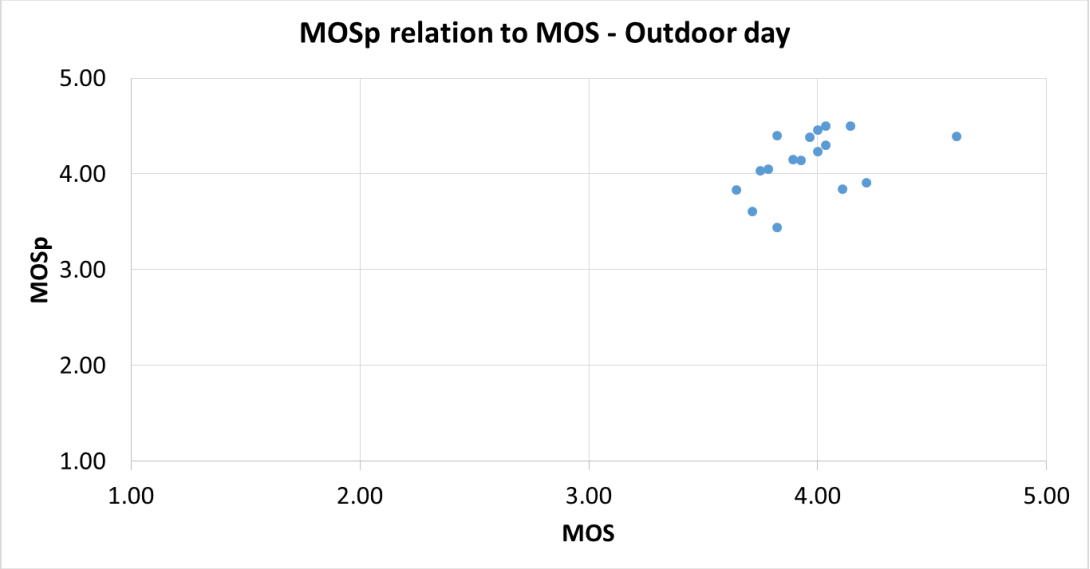


Fig.3.7. MOSp relation to MOS of outdoor day images [13].

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

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

Table 3.5. Pearson correlation coefficients of the outdoor day images

MOS	MOSp	Pearson correlation coefficient
3.97	4.13	0.44

It has also been observed from the Table 3.5 that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of outdoor day images

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. The plot of the ranked data in Figure 3.8, then we see that they are linearly related.

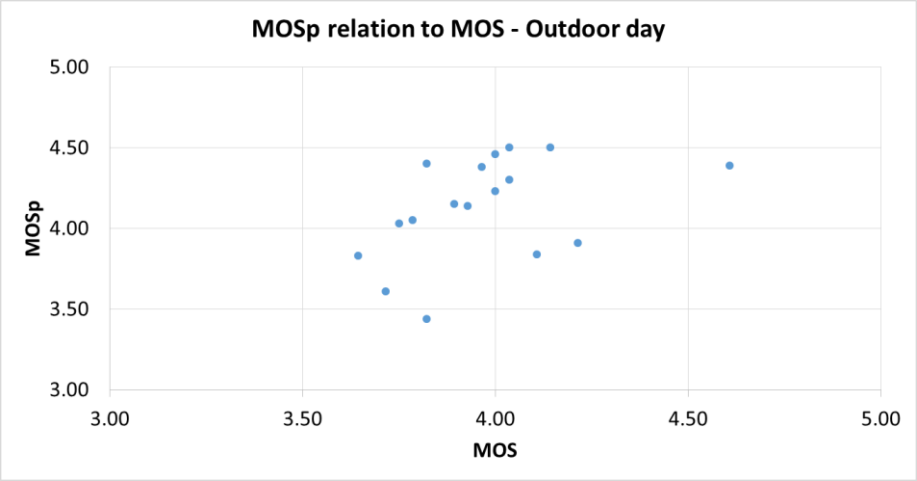


Fig.3.8. Ranked data of MOSp relation to MOS of outdoor day images [13].

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

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

Table 3.6. SROCC of the outdoor day images

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 images

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 3.9 presents the Outlier Ration values, a smaller outlier fraction means the model's predictions are more consistent.

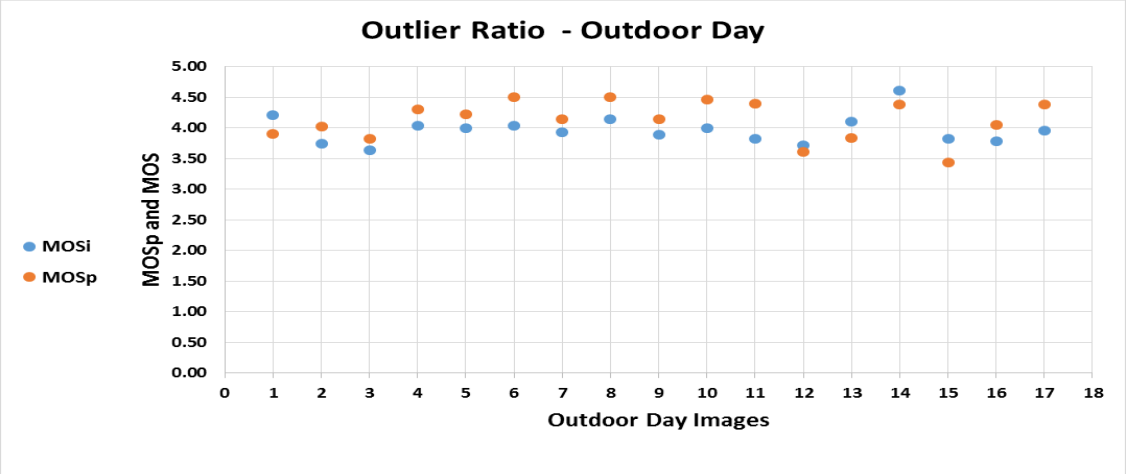


Fig.3.9. Outlier Ration values of outdoor day images.

The objective test plan specifies this metric as follows:

Outlier Ratio = outliers / N

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

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

Table 3.7. Outlier ratio of the outdoor day images

MOS	MOSp	Outlier ratio
3.97	4.13	0.036

The smallest Outlier Ratio is better. Table 3.7 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 images

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.

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

Table 3.8. AAE of the outdoor day images

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.

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

Table 3.9. RMSE of the Outdoor Day images

MOS	MOSp	Root mean square error
3.97	4.13	0.16

Lower value of RMSE is better, Table 3.9 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 3.10 and the MOSp relation to MOS is illustrated in Figure 3.10.

Table 3.10. MOS and MOSp values of the indoor images

Image	18	19	20	21	22	23	24	25	26	27	28	29	30
MOSi	3.71	4.14	3.32	3.64	3.75	4.14	3.50	3.71	3.71	3.93	4.00	3.79	3.57
MOSp	3.97	3.85	2.97	3.46	3.20	4.41	3.02	3.93	3.19	4.25	3.80	3.89	3.80

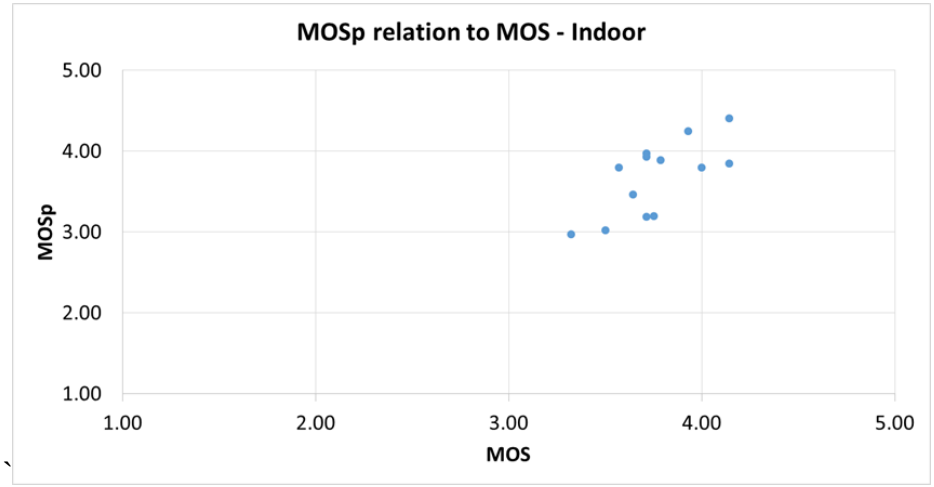


Fig. 3.10. MOSp relation to MOS of indoor images

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

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

Table 3.11. Pearson correlation coefficients of the indoor images

MOS	MOSp	Pearson correlation coefficient
3.97	4.13	0.72

Table 3.11 shows that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of indoor images

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 3.11, then we see that they are linearly related.

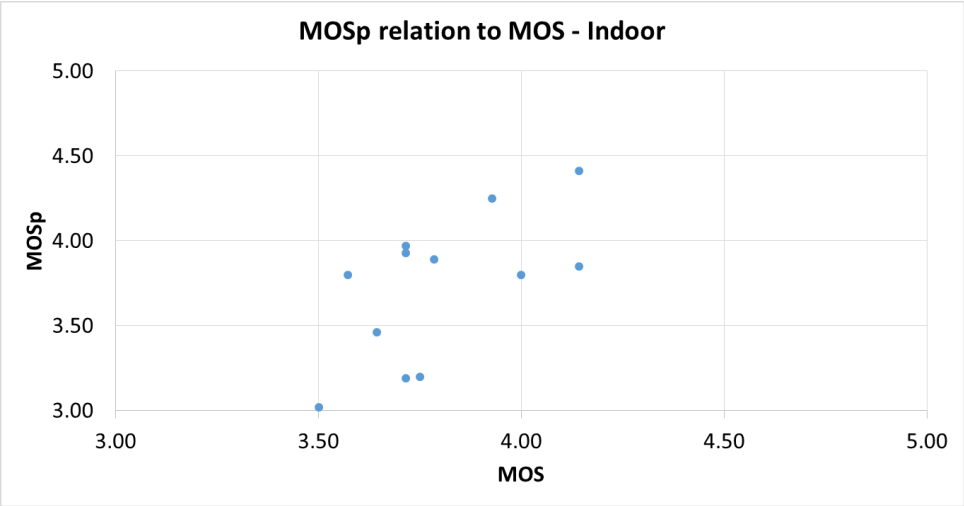


Fig. 3.11. Ranked data of MOSp relation to MOS of indoor images [13]

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

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

Table 3.12. SROCC of the indoor images

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 images

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 3.12 presents the Outlier Ration values, a smaller outlier fraction means the model's predictions are more consistent.

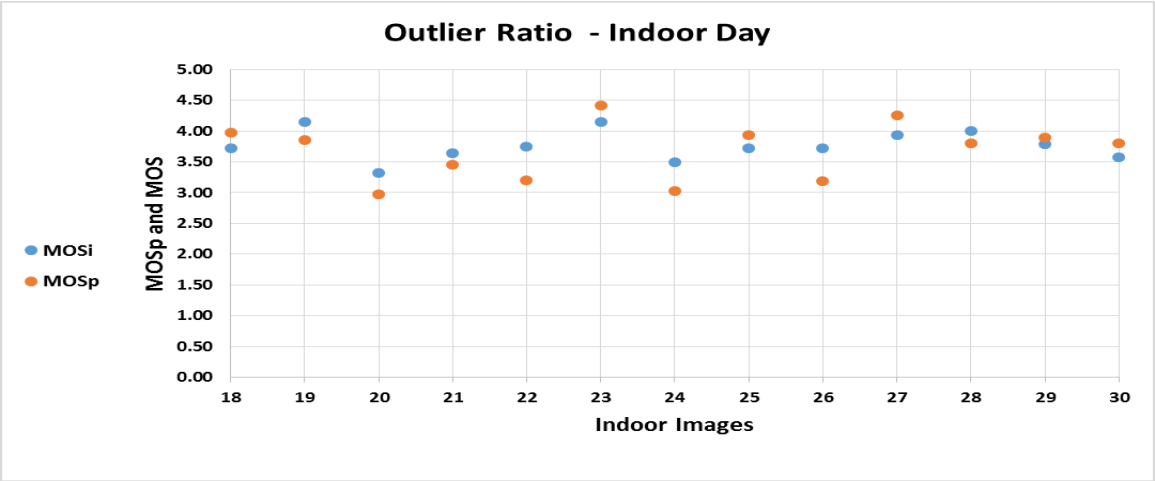


Fig.3.12. Outlier ratio values of indoor images.

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

The smallest Outlier Ratio is better. Table 3.12 shows the outlier ratio (OR = 0.11) for the new model calculated over the main partitions of the subjective data.

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

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

Table 3.13. Outlier ratio of the indoor images

MOS	MOSp	Outlier ratio
3.97	4.13	0.11

Metric 4: Average/Mean absolute prediction error of Indoor images

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.

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

Table 3.14. AAE of the indoor images

MOS	MOSp	Average absolute prediction error
3.97	4.13	0.31

Metric 5: Root mean square error of indoor

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

Table 3.15. RMSE of the indoor images

MOS	MOSp	Root mean square error
3.97	4.13	0.09

Lower value of RMSE is better, Table 3.15 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 3.16 and the MOSp relation to MOS is illustrated in Figure 3.13.

Table 3.16. MOSi and MOSp values of the outdoor night images

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
MOSp	2.01	3.45	3.44	3.04	2.79	2.83	3.02	2.10	3.84	1.46

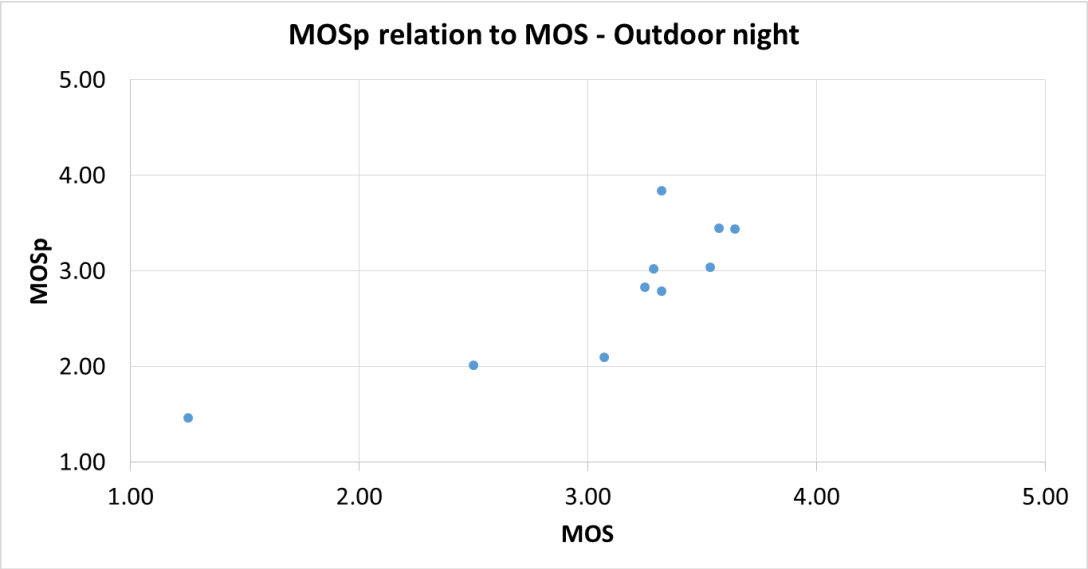


Fig.3.13. MOSp relation to MOS of outdoor night images.

$$CC = \frac{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)\left(MOS_p(i) - \overline{MOS_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(MOS(i) - \overline{MOS}\right)^2}\sqrt{\sum_{i=1}^{N}\left(MOS_p(i) - \overline{MOS_p}\right)^2}} \tag{3.16}$$

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

Table 3.17. Pearson correlation coefficients of the outdoor night images

MOS	MOSp	Pearson correlation coefficient
3.97	4.13	0.84

The high Pearson Correlation Coefficient value (0.84) which observed in Table 3.17 is strongly support that the proposed method provides sufficient prediction accuracy (higher CC).

Metric 2: SROCC of outdoor night images

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 3.14, then we see that they are linearly related.

$$SROCC = 1 - \frac{6\sum_{i=1}^{N}\left(MOS(i) - MOS_p(i)\right)^2}{N(N^2-1)} \tag{3.17}$$

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

Table 3.18. SROCC of the outdoor night images

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 images

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.

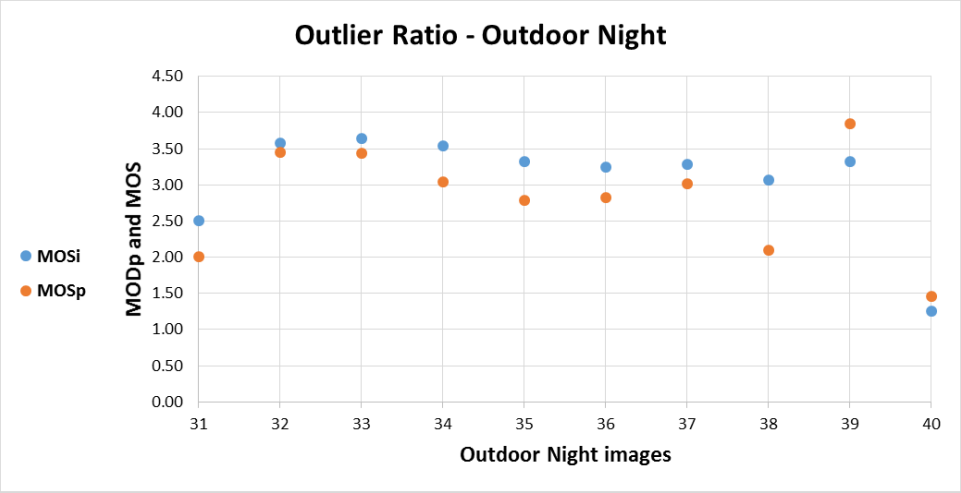


Fig.3.14. Outlier Ration values of outdoor night images.

The objective test plan specifies this metric as follows:

Outlier Ratio = outliers / N

$$OR = \frac{(total\ number\ of\ outliers)}{N} \tag{3.18}$$

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

Table 3.19. Outlier ratio of the outdoor night images

MOS	MOSp	Outlier ratio
3.97	4.13	0

The smallest OR is better. Table 3.19 shows the outlier ratio (OR = 0) for the new model calculated over the main partitions of the subjective data.

Metric 4: MAE of outdoor night images

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:

$$AAE = \frac{1}{N}\sum_{i=1}^{N}|MOS(i) - MOS_p(i)| \tag{3.19}$$

Table 3.20. AAE of the outdoor night images

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.***

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(MOS(i) - MOS_p(i))^2} \tag{3.20}$$

Table 3.21. RMSE of the Outdoor Night images

MOS	MOSp	Root mean square error
3.97	4.13	0.28

Lower value of RMSE is better, Table 3.21 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.

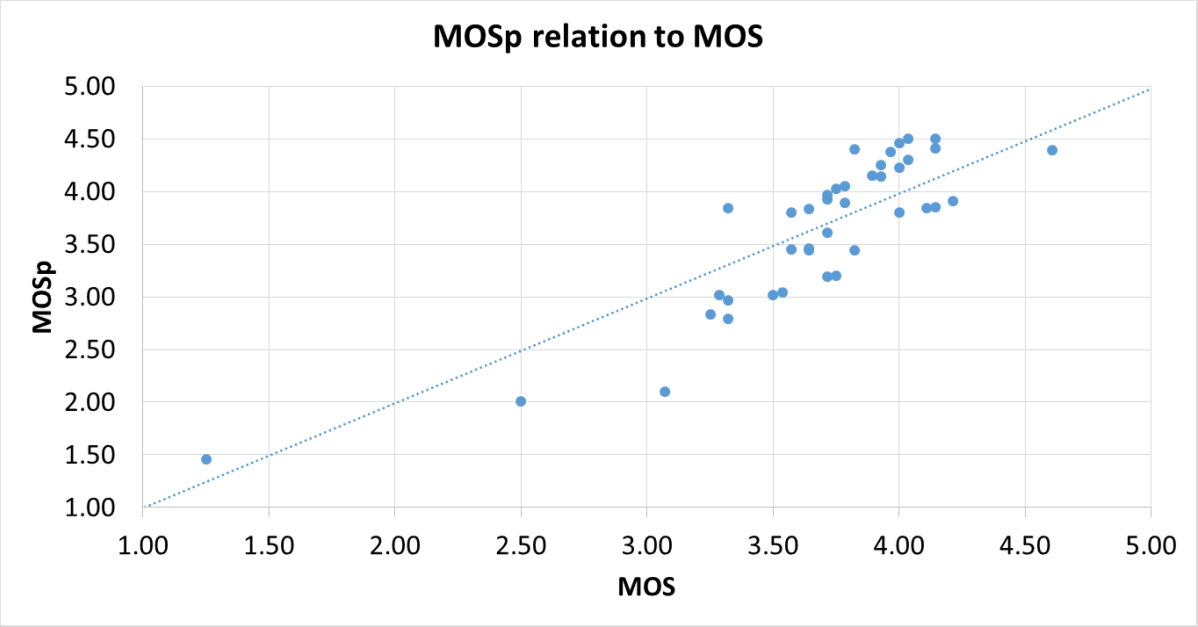


Fig. 3.15. MOSp relation to MOS of all images.

3.4. Model for perceived image quality evaluation

The model development work is presented in details in Chapter 2 and the model evaluation is presented in Chapter 3 of this thesis.

Figure 2.11 illustrates the workflow of the new model and framework of perceived IQ prediction. The deliverables of this study are: formulas for image quality attributes calculation, framework and open source code of VIQET.

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 (M) – rang: 0 -255;
- Noise Signature Index (N) – range : 0 – 590 ;
- Color saturation (C) – range: 0 -255;
- Illumination (I) – range: 0 – 255;
- Dynamic Range (D) – 0 – 255;

The image quality evaluation model represented in Formulas 3.16 – 3.22. The weighted coefficients (A_m, A_n, A_c, A_i and A_d) were constructed by applying linear regression of the image quality assessment resulted of HVTs and VIQET. The weighted coefficients are proposed to formulate perceived image quality. In order to obtain an image quality evaluation model for a single given image, Formula 3.16 should be performed and for overall perceived image quality of a smartphone, Formula 3.22 should be performed.

Model for perceived IQ evaluation of a single image

In order to evaluate the perceived IQ of an individual image, Formula 3.21 should be performed. This formula represents the five IQ criteria measured by VIQET. Each IQ criteria multiplied by the respectively coefficient, which was obtained by linear regression of the obtained results of HVTs and VIQET.

$$MOS_p = \frac{M_i}{M_{max}} A_m + \frac{N_i}{N_{max}} A_n + \frac{C_i}{C_{max}} A_c + \frac{I_i}{I_{max}} A_i + \frac{D_i}{D_{max}} A_d \quad (3.21)$$

Where M_{max}, N_{max}, C_{max}, I_{max} and D_{max} are the maximum values of IQ criteria measured by VIQET

A_m, A_n, A_c, A_i and A_d present the IQ coefficients for each IQ criteria.

Calculation of the *multi-scale edge acutance coefficient* - A_m

The A_m coefficient value was obtained by linear regression calculation of the results of HVTs and VIQET measurements. These results reflect the relationship between MOSi values and the values of multi-scale edge acutance, which measured by the VIQET.

Calculation of the *noise signature index* – A_n

The A_n coefficient value was obtained by linear regression calculation of the results of HVTs and VIQET measurements. These results reflect the relationship between MOSi values and the values of the noise signature index, which measured by the VIQET.

Calculation of the *color saturation* – A_c

The A_c coefficient value was obtained by linear regression calculation of the results of HVTs and VIQET measurements. These results reflect the relationship between MOSi values and the values of the color saturation, which measured by the VIQET.

Calculation of the *Illumination* – A_i

The A_i coefficient value was obtained by linear regression calculation of the results of HVTs and VIQET measurements. These results reflect the relationship between MOSi values and the values of the illumination, which measured by the VIQET.

Calculation of the *dynamic range* – A_d

The A_d coefficient value was obtained by linear regression calculation of the results of HVTs and VIQET measurements. These results reflect the relationship between MOSi values and the values of the dynamic range, which measured by the VIQET.

Coefficients index and their values:

- A_m - Multi-scale Edge Acutance, value = 3.04;
- A_n - Noise Signature Index, value = 1.01;
- A_c - Color saturation, value = 1.22;
- A_i - Illumination, value = 1.12;
- A_d - Dynamic range, value = 1.84;

Model for perceived IQ evaluation of smartphone or any mobile device

In order to evaluate the overall perceived IQ of a smartphone or any other mobile device, a number of images is required. The accuracy of the overall perceived IQ is depended on the number of images (bigger number = higher accuracy). For overall perceived IQ evaluation, Formula 3.27 should be performed. This formula represents the five IQ criteria measured by VIQET. Each IQ criteria is divided by the RMS (root mean square) value, to be calculated by Formulas 3.22 – 3.26 and multiplied by the respectively coefficient.

RMS calculations of IQ attributes values of VIQET

The RMS (root mean square) value of each VIQET image quality criteria should be performed by the following formulas respectively:

$$M_{rms} = \sqrt{\frac{1}{N} \sum_{i=0}^{N} M_i^{\,2}} \tag{3.22}$$

Where M_{rms} is the RMS (root mean square) of the Multi-scale Edge Acutance (Mi) values measured by VIQET.

$$D_{rms} = \sqrt{\frac{1}{N} \sum_{i=0}^{N} D_i^{\,2}} \tag{3.23}$$

Where D_{rms} is the RMS (root mean square) of the Dynamic Range (D_i) values measured by VIQET.

$$C_{rms} = \sqrt{\frac{1}{N} \sum_{i=0}^{N} C_i^{\,2}} \tag{3.24}$$

Where C_{rms} is the RMS (root mean square) of the Color saturation (C_i) values measured by VIQET.

$$N_{rms} = \sqrt{\frac{1}{N} \sum_{i=0}^{N} N_i^{\,2}} \tag{3.25}$$

Where N_{rms} is the RMS (root mean square) of the Noise Signature Index (N_i) values measured by VIQET

$$I_{rms} = \sqrt{\frac{1}{N} \sum_{i=0}^{N} I_i^{\,2}} \tag{3.26}$$

Where I_{rms} is the RMS (root mean square) of the Illumination (I_i) values measured by VIQET.

$$MOS_p = \frac{M_{rms}}{M_{max}} A_m + \frac{N_{rms}}{N_{max}} A_n + \frac{C_{rms}}{C_{max}} A_c + \frac{I_{rms}}{I_{max}} A_i + \frac{D_{irms}}{D_{max}} A_d \tag{3.27}$$

Where M_{max}, N_{max}, C_{max}, I_{max} and D_{max} are the maximum values of the IQ criteria measured by VIQET.

Where M_{rms}, N_{rms}, C_{rms}, I_{rms} and D_{rms} are the RMS (root mean square) of the IQ criteria values measured by VIQET.

A_m, A_n, A_c, A_i and A_d present the IQ coefficients for each IQ criteria.

Coefficients index and their values:

- A_m - Multi-scale Edge Acutance, value = 3.04.
- A_n - Noise Signature Index, value = 1.01.
- A_c - Color saturation, value = 1.22.
- A_i - Illumination, value = 1.12.
- A_d - Dynamic range, value = 1.84.

3.5. The study novelty and practical implementation

This study proposes a new approach with full solution for perceived IQ evaluation in smartphones. It is also beneficial for other devices with larger displays since the IQ attributes are the same. Apple introduced the first iPhone in year 2007, since then the smartphones industry is in rapid growth. Display and camera resolutions (number of megapixels) increasing in every day, vendors racing each other in increasing display and camera resolutions. Image transportation in the social media networks increasing as well. Research on IQ getting more importance. Many researches on perceived IQ evaluation methods have been performed in recent 5 years. These researches investigated different aspects of the perceived image quality in mobile devices. The combination of high definition (HD) displays with small size display brings a new challenge for IQ researches.

The recent researches propose new image quality assessment metrics, new algorithms and new models. These researches are targeting image processing and image quality experts (researches and engineers). In this study, we implemented a new subjective and objective IQ assessment metric. The new metric combines special IQ visual tests (subjective) and SW tool for image analysis (VIQET) simultaneously. The visual tests and VIQET IQ evaluation enabled the extraction of the four classic IQ attributes into five new IQ criteria. This makes the IQ assessment more accurate in perceived image quality evaluation.

The idea was “training” the VIQET (SW application) to the preferred image quality attributes by human in order to predict the perceived IQ. The “training” performed through many HVTs and experiments. The calculations of the score of perceived image quality (MOS) are the results of this work.

What is new in this study?

- Using four classic IQ attributes for human visual experiments;
- Find the preferred IQ attributes level by human;
- Extraction of the four classic IQ attributes to five IQ criteria, which implemented in VIQET;
- Using VIQET open source SW application for IQ measurement;
- Simple process, just install the VIQET and use it;
- Everyone can evaluate the perceived IQ, not only IQ experts;

Get perceived IQ in four steps

Figure 3.16. illustrates the four steps for full process in order to get the mean opinion score (MOS) of perceived IQ.

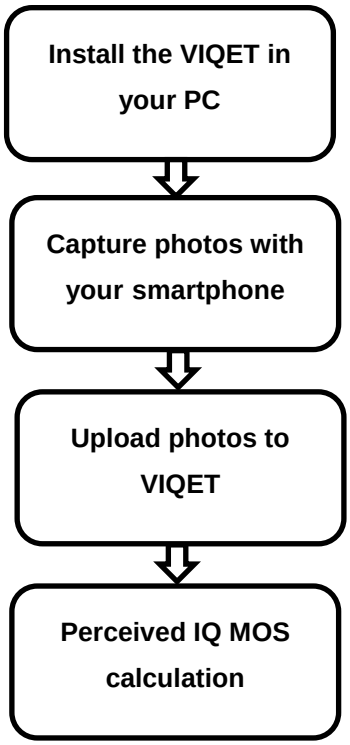


Fig.3.16. Get perceived IQ MOS in four steps.

Practical implementations of the perceived image quality assessment model

The proposed new model and framework for perceived image quality evaluation can be used in various image quality evaluation applications in the research arena and in industry.

It is implemented already in the "ORT Braude engineering college" in Israel and presented to several high-tech organizations for evaluation. It will be proposed to social networks with extensive media transportation in order to perform online image quality evaluation and image quality improvement.

Here are few examples for the new model implementation:

In academia – researches in image processing can use this model for IQ evaluation of new image processing features in order to get immediate feedback on their work. Students in "ORT Braude" [58] use it in their final engineering projects in image processing course. When they develop a new algorithm, they measure the IQ of the new image and compare with the previous version.

In Industry – The new model has been proposed (free of charge) to several high-tech companies in Israel. These companies are developing new technologies and devices based on image processing for general consumer market and for medical machines [101]. The image quality is an essential part of these products. The evaluation process might take few months.

In public image database (e.g. museums, art galleries etc.) – Evaluating the images IQ, try to improve and decide to save or delete. This will improve the database memory capacity.

In space research – The new model can be implemented in the satellites image processing systems. While satellites capture images in the space, the image quality will be evaluated and improved in real time before transmitting to the stations on the earth.

In social media networks [101] **–** Social networks use extensive transportation of images and videos. Since the image capture done and uploaded in real-time regardless to the image quality, implementation of the new model in the social networks servers will improve the image quality of the uploaded and downloaded images and videos (in future). Figure 3.17. demonstrates the implementation of the new model in several present social networks.

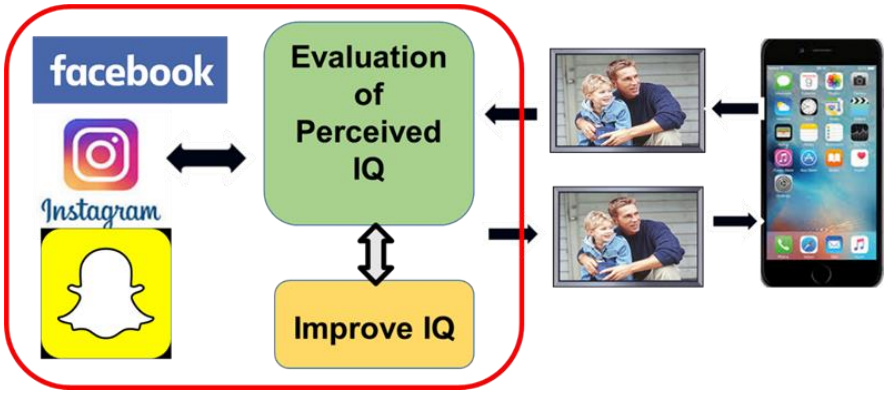


Fig.3.17. Online IQ improvement in social networks.

3.6. Summary

Chapter 3 describes the evaluation process of the new model using mathematical criteria which is the standard performance evaluation procedures of VQEG [6]. The performance of the new model is characterized by three prediction attributes: accuracy, monotonicity and consistency. The proposed new model's was evaluated also by five evaluation metrics as recommended by VQEG[6]:

- Pearson correlation coefficients (PCC);
- Spearman rank order correlation coefficient (SROCC);
- Outlier ratio (OR);
- Mean Average Error (MAE);
- Root Mean Square Error (RMSE);

The performances for every one of the evaluation metrics are sufficient for both the training and the testing datasets.

The outcomes of the new model evaluation were:

- Approval of fully automated NR image quality assessment process, which can assess the quality of processed images without human intervention;
- The proposed method exhibits excellent consistency and compatibility with subjective assessment;
- The new model and framework can blindly estimate and predict the perceived image quality of images of mobile phones in very friendly and simple procedure by using the VIQET application;

Chapter 3 describes also how the new model was evaluated. The new model was evaluated by using mathematical criteria, which is the standard performance evaluation procedures of VQEG [6]. The results of the final HVTs including scores of 35 observers for 40 original natural images and the results obtained by the VIQET were analyzed in order to measure the ability to predict the perceived IQ. The performance of the new model is characterized by three prediction attributes: accuracy, monotonicity and consistency.

The new model was evaluated also by five evaluation metrics which were recommended by VQEG[6]. Pearson correlation coefficient (PCC) and spearman rank order correlation coefficient (SROCC) present the correlation between 2 groups of values (HVTs scores and VIQET scores).

The outlier ratio (OR) presents the ratio between received scores that out of range of the group of overall scores.

Mean average error (MAE) and root mean square error (RMSE) present the error range in the received scores in HVTs and VIQET.

The outcome of this evaluation process indicates that the performances for every one of the evaluation metrics are sufficient for both phase I and phase II testing datasets.

The main conclusions of chapter 3 based on the new model evaluation are: approval of the new NR image quality assessment process, which can assess the quality of processed images without human intervention. The proposed method exhibits consistency and compatibility with subjective assessment. The new model and framework can blindly estimate and predict the perceived image quality of images of smartphones by using the VIQET application and the proposed model formula.

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 results are useful for the mobile phone industry to have a better and efficient process for understanding of the concrete benefit of enhancing the image quality attributes. The proposed quality assessment model can be used by technology magazines while comparing and rating new smartphones models.
- 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.

Future work

The research, described in this study, was focused on still images quality assessment based on four IQ attributes. The continuation of this research will deal with video material in HD content. The perceived image quality of live video is a new challenge in the image quality assessment field. The recommended IQ attributes for future research might be:

- Frame rate conversion quality
- Band width limitations
- Video compression/decompression artifacts
- Motion artifacts

Another application for the research outcomes can be used in social networks (e.g. Facebook, Instagram, Snapchat etc.) domain. Also, since the VIQET tool for image analysis is an open source application it is highly recommended to use the current version as starting point in order to improve it and make it up to date for future IQ attributes and objective image quality assessment.

BIBLIOGRAPHY

(all web references was accessed on 8.01.2018)

1. The statistics portal. https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/
2. Sheikh H. R., Bovik A. C., Cormack L. K., No-reference quality assessment using natural scene statistics: JPEG2000," IEEE Transactions on Image Processing, vol. 14, no. 11, pp. 1918-1927,December2005. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.458.7315&rep=rep1&type=pdf
3. Engeldrum P. G. Image quality modeling: Where are we? In image processing, image quality, image capture systems conference (Savannah, Georgia, Apr. 1999), Society for Imaging Science and Technology, pp. 251–255. 7, 9
4. Gong R., Xu H., Impacts of appearance parameters on perceived image quality for mobile-phone displays, Optik, vol. 125, no. 11, pp. 2554–2559, 2014.
5. Hyesung J., Choon W. K. Perceived image quality assessment for color images on mobile displays, SPIE 9395, 2015.
6. 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/.
7. Fairchild M. D. Image quality measurement and modeling for digital photography, May 2002. 8
8. Yendrikhovskij S. Image quality and color categorization. Color image science: Exploiting Digital Media (2002), 393–420. 8
9. Keelan B. Handbook of image quality: Characterization and prediction. 2009. P. 3.
10. Ciocca G., Corchs S., Gasparini F., et al. Modeling image quality, 2015. 8
11. Zorea P. International conference Mathematics & Information Technologies (MITRE), 2015. Modeling and predicting perceived image quality by smartphone users, ISBN 978-9975-71-678-9, p.102.
12. Pavel M., Sperling G., Riedl T., et al. Limits of visual communication: The effect of signal-to-noise ratio on the intelligibility of American sign language. Optical Society of American 4, 12 (Dec. 1987), 2355–2365. 41, 42
13. Zorea P. Perceived image quality prediction based on image quality attributes extraction. International Journal of Scientific & Engineering Research, Volume 8, Issue 2, February-2017, ISSN 2229-5518, p.1303-1312.

14. Adelson E. H. Lightness perception and lightness illusions. In the new cognitive neurosciences, M. S. Gazzaniga, Ed., second ed., vol. 3. MIT Press, Cambridge, MA, USA, 2000, ch. 24, pp. 339–351. 10
15. Hunt R. W. G. Measuring color, 4 ed. Wiley, 2011. 10
16. Berns R. S. Billmeyer and Saltzman's Principles of color technology, 3 ed. Wiley, New York, NY, USA, 2000. 10
17. Seetzen H., Li H., Ye L., et al. 25.3: Observations of luminance, contrast and amplitude resolution of displays. SID Symposium Digest of Technical Papers 37, 1 (2006), 1229–1233. 41
18. Fairchild M., Wyble D. R. Colorimetric characterization of the apple studio display (flat panel LCD). Tech. rep., Rochester Institute of Technology, New York, NY, USA, July 1998. 3, 10, 30
19. Menu G., Peigne L., Hardeberg J. Y., et al. Correcting projection display non-uniformity using a webcam. In color imaging X: processing, hardcopy, and applications (San Jose, USA, Jan. 2005), R. Eschbach and G. G. Marcu, Eds., International Society for Optics and Photonics, pp. 364–373. 37
20. Ling T. The Assessment of ceiling uniformity for indirect lighting systems. Tech. rep., Rensselaer Polytechnic Institute, 1996. 39
21. Devalois R. L., Devalois K. K. Spatial vision. Oxford University Press, 1988. 11
22. Chu X. Q., Yang C., Li Q. Contrast-sensitivity-function-based clutter metric. Optical Engineering 51, 6 (Apr. 2012), 067003. 13. https://aaltodoc.aalto.fi/bitstream/handle/123456789/18961/isbn9789526065595.pdf?sequence=1
23. Leclaire A., Moisan L. No-reference image quality assessment and blind deblurring with sharpness metrics exploiting Fourier phase information. Journal of Mathematical Imaging and Vision (2014). 45 http://www.math-info.univ-paris5.fr/~moisan/papers/2014-03r.pdf
24. Fei X., Xiao L., Sun Y., et al. Perceptual image quality assessment based on structural similarity and visual masking. Signal Processing: Image Communication 27, 7 (Aug. 2012), 772–783. 12, 15 https://pdfs.semanticscholar.org/2b81/9bef80c02d5d4cb56f27b202535e119df988.pdf
25. Daly S. The visible differences predictor: an algorithm for the assessment of image fidelity. In human vision, visual processing, and digital display III (San Jose, CA, USA, Aug. 1992), B. E. Rogowitz, Ed., vol. 1666, International society for optical engineering, pp. 179–206. 12.

26. Johnson G. M. Measuring images: differences, quality, and appearance. PhD thesis, Rochester Institute of Technology, New York, NY, USA, mar 2003. 12, 16
27. Winkler S. Visual fidelity and perceived quality: Towards comprehensive metrics. In storage and retrieval for image and video databases (San Jose, CA, USA, Jun 2001), B. E. Rogowitz and T. N. Pappas, Eds., SPIE, pp. 114–125. 12
28. Kerren A., Ebert A., Meyer J. Interacting with visualizations in human centered visualization, Volume 4417 of LNCS Tutorial. Springer, 2007, ch. 3, pp. 77–162. 20
29. Van Heesch F., Klompenhouwer M., De Haan G., Masking coding artifacts on large displays. In Digest of Technical Papers - IEEE International Conference on Consumer Electronics (Jan. 2006), vol. 2006, IEEE, pp. 207–208. 15
30. Sheikh H. R., Bovik A. C. Image information and visual quality. IEEE Transactions on Image Processing 15, 2 (Feb. 2008), 430–444. 45
31. Majumder A., Gopi M. Modeling color properties of tiled displays. Computer Graphics Forum 24, 2 (2005), 149–163. 10, 30
32. Rizzi A., Algeri T., Medeghini G., et al. Proposal for contrast measure in digital images. in European Conference on Color in Graphics, Imaging, and Vision (Aachen, Germany, 2004), pp. 187–192. 31, 41, 42
33. Calabria A. J., Fairchild M. D. Perceived image contrast and observer preference I. The effects of lightness, chroma, and sharpness manipulations on contrast perception. Journal of Imaging Science and Technology 47, 6 (2003), 494–508. 15
34. Masaoka K., Berns R.S., Fairchild M.D., et al. The number of discernible object colors is a conundrum, Journal of the Optical Society of America A 30 264-277 (2013).
35. De Ridder H. Naturalness and image quality: Saturation and lightness variation in color images of natural scenes. Journal of Imaging Science and Technology 40, 6 (Jan. 1996), 487–493. 15
36. Barten P. G. J. Contrast sensitivity of the human eye and its effects on image quality, first edit ed. Society for imaging science and technology, 1999. 15
37. Fry E., Triantaphillidou S., Jarvis J., et al. Image quality optimization, via application of contextual contrast sensitivity and discrimination functions. In image quality and system performance XII 93960K, proceedings of electronic imaging (San Francisco, CA, USA, Jan. 2015), M.-C. Larabi and S. Triantaphillidou, Eds., International Society for Optics and Photonics, pp. 93960K–93960K–12. 15

38. Asano T., Takagi Y., Kondo T., et al. Image quality evaluation based on contrast sensitivity function. In IEEE International Conference on Mechatronics and Automation (Beijing, China, Aug. 2011), IEEE, pp. 658–663. 15, 41
39. Liu X., Pedersen M., Hardeberg J. Y. CID:IQ - A new image quality database. In international conference on image and signal processing (Cherbourg, Normandy, France, 2014), A. Elmoataz, O. Lezoray, F. Nouboud, and D. Mammass, Eds., vol. 8509 of Lecture Notes in Computer Science, Springer International Publishing, pp. 193–202. 42, 45, 157
40. Fu W., Gu X., Wang Y. Image quality assessment using edge and contrast similarity. In international joint conference on neural networks (Hong Kong, China, June 2008), IEEE, pp. 852–855. 15
41. Ramesh R., Brown M. S., Yang R., et al. Multi-projector displays using camera-based registration. In proceedings of IEEE Visualization (San Francisco, CA, USA, Oct. 1999), IEEE, pp. 161–168. 3
42. Kim Y. J., Luo M. R., Choe W., et al. Factors affecting the psychophysical image quality evaluation of mobile phone displays: The case of transmissive liquid-crystal displays. Journal of the Optical Society of America. Optics, image science, and vision 25, 9 (Sep. 2008), 2215–2222. 16
43. Eerola T., Lensu L., Kalvi A., et al., Full reference printed image quality: Measurement framework and statistical evaluation. Journal of Imaging Science and Technology 54, 1 (2010), 010201. 26 http://personal.lut.fi/users/tuomas.eerola/files/jist10.pdf
44. Morales J. A. Assessment of Iris reflection artifacts and alignment in fundus images. PhD thesis, Saint Mary's University, San Antonio, TX, USA, 2011. 15 https://www.researchgate.net/profile/Jafet_Morales/publication/311428681_Assessment_of_Iris_Reflection_Artifacts_and_Alignment_in_Fundus_Images/links/58460b5608aeda69681a62c1/Assessment-of-Iris-Reflection-Artifacts-and-Alignment-in-Fundus-Images.pdf
45. Fairchild M. D. Color appearance models, 2nd ed. John Wiley & Sons, Ltd, 2005. 10, 12, 15.
46. Sun P., Chubb C., Wright C., et al. Human attention filters for single colors. Proceedings of the national academy of sciences of the United States of America. 113: E6712-E6720. PMID 27791040 DOI: 10.1073/pnas.1614062113 (2016). http://europepmc.org/backend/ptpmcrender.fcgi?accid=PMC5087007&blobtype=pdf
47. Liang Z. Method for reproducing color images having one color gamut with a device having a different color gamut, Jun. 1994. 18

48. Gong R., Xu H., Wang Q., et al. Investigation of perceptual attributes for mobile display image quality. Optical engineering 52, 8 (Aug. 2013), 083104. 16
49. Kim J. S., Westland S., Luo M. R. Image quality assessment for photographic images. In congress of the international color association (Granada, Spain, may 2005), J. L. Nieves and J. Hernandez-Andres, Eds., pp. 1095–1098. 15
50. Thomas J. B. Colorimetric characterization of displays and multi-display systems. PhD thesis, Universite de Bourgogne, Dijon, France, oct 2009. 10, 16, 18, 30
51. Strand M., Hardeberg J. Y., Nussbaum P. Color image quality in projection displays: A case study. In image quality and system performance II, color imaging conference (San Jose, CA, USA, Jan. 2005), R. Rasmussen and Y. Miyake, Eds., SPIE, pp. 185–195. 10, 16
52. Liu Z., Wu W. The use of the contrast sensitivity function in the perceptual quality assessment of fused image. International journal of Image and Data Fusion 2, 1 (Feb. 2011), 93–103. 15
53. Kuo C., Ng Y. Perceptual color contouring detection and quality evaluation using scanners. Journal of Imaging Science and Technology 49, 1 (Jan. 2005), 41–46. 15
54. Majumder A., Irani S. Contrast enhancement of images using human contrast sensitivity. In the 3rd symposium on applied perception in graphics and visualization (New York, NY, USA, Jun 2006), pp. 69–76. 41
55. ITU-R recommendation BT.500-11. methodology for the subjective assessment of the quality of television pictures, international telecommunication union, Geneva, Switzerland, 2002.
56. VQEG: Final report from the video quality experts group on the validation of objective models of video quality assessment. Available: https://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx
57. Winkler S. Analysis of Public Image and Video Databases for Quality Assessment. Image and video quality resources. https://stefan.winkler.site/Publications/jstsp2012.pdf
58. Zorea P., The 11th Interdisciplinary research conference ORT Braude college. 2015. New approach for modeling perceived image quality by smartphone users., ISBN 978-965-7468-01-2, p.37. http://www.braude.ac.il/files/conferences/2015/ResearchConference15/Program_and_Book_of_Abstracts.pdf

59. Geisler W. S., Banks M. S., Visual performance. In M. Bass (Ed.), handbook of optics. McGraw-Hill, 1995.
60. Cormack L. K., Computational models of early human vision. In A. C. Bovik (Ed.), Handbook of image and video processing, 2nd ed. Elsevier academic press, April 2005.
61. Watson A. B., Digital images and human vision. The MIT press, Cambridge, MA, 1993.
62. Wang Z., Bovik A. C., Modern image quality assessment. Synthesis lectures on image, video, & multimedia processing, Morgan & Claypool, San Refael, Calif, USA, 2006.
63. Sheikh H. R., Sabir M. F., Bovik A. C., A Statistical evaluation of recent full reference image quality assessment algorithms, IEEE Trans. Image Processing, vol. 15, no. 11, pp. 3440-3451, 2006.
64. Yuen M., Wu H. R., A survey of hybrid MC/ DPCM/ DCT video coding distortions, signal processing, vol. 70, no. 3, pp. 247-278, November 1998.
65. Zorea P. New model for evaluation perceived image quality by Smartphone users. STUDIA UNIVERSITATIS MOLDAVIAE, 2015, nr.2 (82) ISSN 1857-2073, ISSN online 2345-1033, p.90-97. http://studiamsu.eu/wp-content/uploads/15.rezumat_82.pdf
66. Zorea P. Prediction of smartphones' perceived image quality by sw evaluation tool. STUDIA UNIVERSITATIS MOLDAVIAE, 2016, nr.7 (97) ISSN 1857-2073, ISSN online 2345-1033, p.170 - 180. http://studiamsu.eu/wp-content/uploads/29.p.170-180_97.pdf
67. Le Callet P., Virad-Gaudin C., Barba D., A convolutional neural network approach for objective video quality assessment, IEEE Transactions on Neural Networks, vol. 17, no. 5, pp. 1316-1327, 2006.
68. Gastaldo P., Zunino R., Neural networks for the no-reference assessment of perceived quality, journal of. Electronic Imaging, 14 (3), 033004, 2005.
69. Babu R. V., Suresh S., Perkis A., No-reference JPEG-image quality assessment using GAP-RBF, signal processing, vol. 87, no.6, pp.1493-1503, 2007.
70. Cvetkovic S., Schirris J., De With P., Non-linear locally-adaptive video contrast enhancement algorithm without artifacts, IEEE Transactions on Consumer Electronics, 2008.
71. Kirenko I. O., Muijs R., Shao L., Coding artifact reduction using no reference block grid visibility measure, in proc. IEEE International Conference. Multimedia expo, Jul. 2006, pp. 469---472.

72. Shao L., Wang J., Kirenko I., et al., Quality adaptive trained filters for compression artifacts removal, IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 897-900, 2008.
73. Puttestein J. G., Heynderickx I., De Haan G., Evaluation of objective quality measures for noise reduction in TV- systems, signal processing: image communication, vol. 19, no. 2, pp. 109-119, 2004.
74. Carnec M., Le Callet P., Barba D., Objective quality assessment of color images based on a generic perceptual reduced reference, signal processing: Image communication, vol. 23 , no. 4, pp. 239-256, 2008.
75. Moorthy A. K., Bovik A. C., Visual importance pooling for image quality assessment, IEEE Journal of Selected Topics in Signal Processing, special issue on visual media quality assessment, vol. 3, no.2, April 2009.
76. Liu H., Heynderickx I., Studying the added value of visual attention in objective image quality metrics based on eye movement data, in Proc. ICIP, 2009.
77. Le Meur O., Ninassi A., Le Callet P. et al., Overt visual attention for free-viewing and quality assessment tasks, Elsevier, Signal processing: Image communication, 2010.
78. Le Meur O., Ninassi A., Le Callet P. et al., Do video coding impairments disturb the visual attention deployment? Signal Processing: Image Communication, vol. 2, no. 8, 2010.
79. Kallio T., Kaikkonen A., Usability testing of mobile applications: a comparison between laboratory and field testing, journal of Usability Studies, vol. 1, pp. 4–16, 2005.
80. Jumisko - Pyykko S., Hannuksela M. M., Does context matter in quality evaluation of mobile television? in proceedings of the 10th International conference on human computer interaction with mobile devices and services, pp. 63–72, ACM, September 2008.
81. De Pessemier T., De Moor K., Verdejo A. J. et al., Exploring the acceptability of the audiovisual quality for a mobile video session based on objectively measured parameters, in proceedings of the 3rd international workshop on quality of multimedia experience (QoMEX ’11), pp. 125–130, Mechelen, Belgium, September 2011.
82. Catellier A., Pinson M., Ingram W. et al., Impact of mobile devices and usage location on perceived multimedia quality, in Proceedings of the 4th international workshop on qality of multimedia experience (QoMEX ’12), pp. 39–44, IEEE, Melbourne, Australia, July 2012.
83. De Pessemier T., De Moor K., Joseph W. et al., Quantifying subjective quality evaluations for mobile video watching in a semi-living lab context, IEEE Transactions on Broadcasting, vol. 58, no. 4, pp. 580–589,

2012.https://pdfs.semanticscholar.org/0e0d/5c7e01d1301fa36638ac179965414d242532.pdf?ga=2.192635498.1763161004.1515359708-2005999852.1515359708

84. Yang F., Wan S., Bit stream-based quality assessment for networked video: a review, IEEE Communications Magazine, vol. 50, no. 11, pp. 203–209, 2012.
85. 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. http://www.itu.int/rec/T-REC-P.913-201401-S
86. Kolmogorov - Smirnov Test, http://www.physics.csbsju.edu/stats/ KS-test.html.
87. Kaikkonen A., Kekalainen, A., et al., (2005): Usability testing of mobile applications: A comparison between laboratory and field testing. In journal of Usability Studies, 1 (1) pp. 4-17. http://www.upassoc.org/upa_publications/jus/2005_november/mobile.pdf
88. Horita Y., Kawai Y., Minami Y. et al., Quality evaluation model of coded stereoscopic color image. In Proceedings of SPIE: Visual communications and image processing (VCIP), volume 4067, pages 389– 398, Perth, Australia, June 2000.
89. Gastaldo P., Zunino R., Heynderickx I. et al., Objective quality assessment of displayed images by using neural networks. Signal processing: Image communication, 20(7):643–661, August 2005.
90. Wan Z., Bovik A. C., Ligang L., Why is image quality assessment so difficult. In international conference on acoustics, speech, and signal processing (ICASSP), Orland, CA, USA, 2002.
91. Sheikh H. R., Wang Z., Cormack L. et al., Live image quality assessment database release 2, 2005, http://live .ece.utexas.edu/research/quality/.
92. ITU. Methodology for the subjective assessment of the quality of television pictures, ITU-R recommendation BT 500-13, ITU radio communication sector, 2012. https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf
93. 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, 2016. https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.913-201603-I!!PDF-E&type=items
94. Clark J. H. The Ishihara test for color blindness. American journal of Physiological Optics 5 (1924), 269–276. 155
95. 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.

96. International Telecommunication Union (ITU). Methodology for the subjective assessment of the quality of television pictures. Recommendation ITU-R BT.500-11. https://pdfs.semanticscholar.org/ea88/28da9d2dad02382678e7dcbb38571986a48f.pdf
97. Engeldrum P. G., Image quality modeling: Where are we? In image processing, image quality, image capture, systems conference, pages 251–255. IS&T, 1999.
98. Gong R., Xu H., Wang B, Luo M., Image quality evaluation for smart-phone displays at lighting levels of indoor and outdoor conditions. Opt. Eng. 0001. 51(8):084001-1-084001-6. doi:10.1117/1.OE.51.8.084001.
99. Xiaoyn L., Shining W., Ding Y., Subjective score predictor: A new evaluation function of distorted image quality, mathematical problems in engineering, vol. 2016, pp. 1–10, 2016. http://downloads.hindawi.com/journals/mpe/2016/1243410.pdf
100. Hyesng J., Choon -W. K., Perceived image quality assessment for color images on mobile displays. Proc. SPIE 9395, Color Imaging XX: displaying, processing, hardcopy, and applications, 93950U (January 8, 2015), doi:10.1117/12.2083824.
101. Zorea P., Paladi F., Bragaru T. Prediction of perceived image quality by mobile devices users. In: Moldavian Journal of the Physical Sciences, Vol. 17, Nr.1-2, 2018.

LIST OF FIGURES

Fig. 1.1. The image quality circle [3] 23
Fig. 1.2. New IQAs image quality assessment process at smartphones vendors' labs [13] 24
Fig. 1.3. An illustration of contrast sensitivity function [22] 27
Fig. 1.4. Image quality assessment models [2] 33
Fig. 1.5. Flow chart of subjective image quality assessment [2] 33
Fig. 1.6. Different contrast levels [100] 38
Fig. 2.1. Image quality assessment flow [13] 46
Fig. 2.2. Subjective visual tests content of original images 50
Fig. 2.3. An example of test material processing ("building") [13] 51
Fig. 2.4. Experiments set-up for IQ visual tests and VIQET analysis [13, 65] 52
Fig. 2.5. *Phase I* procedure flow chart for IQAs rating [13] 54
Fig. 2.6. Staring points ranking [65] 55
Fig. 2.7. Double stimulus impairment scale test format [96] 55
Fig. 2.8. An example of VIQET RGB histogram [13] 56
Fig. 2.9. An example of VIQET Sharpness map [13] 57
Fig. 2.10. The Venn diagram illustrates how the overall IQ is influenced by lightness 57
Fig. 2.11. Phase II image quality assessment with VIQET [13] 60
Fig. 2.12. Process flow chart of the new model examination [13, 66] 61
Fig. 2.13. Relationship between the perceived IQ MOS, MSE and PSNR of "building" 63
Fig. 2.14. Relationship between the perceived IQ MOS, MSE and PSNR of "lake" 64
Fig. 2.15. Relationship between the perceived IQ MOS, MSE and PSNR of "taxi" 64
Fig. 2.16. Relationship between the perceived IQ MOS, MSE and PSNR of "man" 65
Fig. 2.17. Relationship between the perceived IQ MOS, MSE and PSNR of "room" 65
Fig. 2.18. Relationship between the perceived IQ MOS, MSE and PSNR 66
Fig. 2.19. Relationship between the perceived IQ MOS, MSE and PSNR of "hall" 66
Fig. 2.20. Relationship between the perceived IQ MOS, MSE and PSNR of "sunset" 67
Fig. 2.21. Relationship between the perceived IQ MOS, MSE and PSNR of "bar" 67
Fig. 2.22. Relationship between the perceived IQ MOS, MSE and PSNR of "airplane" 68
Fig. 2.23. Relationship between VIQET IQ scores of "building" 69
Fig. 2.24. Relationship between VIQET IQ scores of "lake" 69
Fig. 2.25. Relationship between VIQET IQ scores of "man" 70
Fig. 2.26. Relationship between the VIQET IQ scores of "taxi" 70

Fig. 2.27. Relationship between the VIQET IQ scores of “room” 71
Fig. 2.28. Relationship between the VIQET IQ scores of “king” 71
Fig. 2.29. Relationship between the VIQET IQ scores of “hall” 72
Fig. 2.30. Relationship between the VIQET IQ scores of “sunset” 72
Fig.2.31. Relationship between the VIQET IQ scores of “bar” 73
Fig. 2.32. Relationship between the VIQET IQ scores of “airplane” 73
Fig. 2.33. MOS of ten high quality images with different IQ attributes 76
Fig. 2.34. VIQET scores of ten high quality images with different IQ attributes 76
Fig. 2.35. Illustration of the DMOS of brightness 77
Fig. 2.36. Illustration of the DMOS of contrast 78
Fig. 2.37. Illustration of the DMOS of saturation 78
Fig. 2.38. Illustration of the DMOS of sharpness 79
Fig. 2.39. MOS of ten images with IQ attributes that scored on the HVT 82
Fig. 2.40. Multi-scale edge levels of ten images with IQ attributes by VIQET 83
Fig. 2.41. Noise Signature levels of ten images with IQ attributes by VIQET 83
Fig. 2.42. Saturation levels of ten images with IQ attributes by VIQET 84
Fig.2.43. Illumination levels of ten images with IQ attributes by VIQET 84
Fig. 2.44. Dynamic range levels of ten images with IQ attributes by VIQET 85
Fig. 3.1. An example of perceived image quality and predicted scores [2] 87
Fig. 3.2. Two hypothetical models with different prediction accuracy [95] 88
Fig. 3.3. Two hypothetical models with different prediction monotonicity [95] 89
Fig. 3.4. Two hypothetical models with different prediction consistency [95] 89
Fig. 3.5. VIQET image quality analysis scores and sharpness map 93
Fig. 3.6. VIQET image quality analysis scores and RGB histogram 93
Fig. 3.7. MOSp relation to MOS of outdoor day images [13] 96
Fig. 3.8. Ranked data of MOSp relation to MOS of outdoor day images [13] 97
Fig. 3.9. Outlier Ration values of outdoor day images 98
Fig. 3.10. MOSp relation to MOS of indoor images 100
Fig. 3.11. Ranked data of MOSp relation to MOS of indoor images [13] 101
Fig. 3.12. Outlier ratio values of indoor images 102
Fig. 3.13. MOSp relation to MOS of outdoor night images 104
Fig. 3.14. Outlier Ration values of outdoor night images 105
Fig. 3.15. MOSp relation to MOS of all images 107
Fig. 3.16. Get perceived IQ MOS in four steps 112

LIST OF TABLES

Table 2.1. Subjective test conditions and parameters....56
Table 2.2. VIQET image quality categories....58
Table 2.3. Image quality attributes extraction to VIQET image quality categories....58
Table 2.4. LPCC of each IQ attributes in HVT (Phase I)....62
Table 2.5. Results of the accuracy and signification calculation using RMSE formula....74
Table 2.6. Results of the VIQET image quality analysis....74
Table 2.7. Results of the accuracy and signification calculation using RMSE formula....75
Table 2.8. Comparison between perceived image quality of IQ Attributes....77
Table 2.9. DMOS calculations of IQ attributes....80
Table 3.1. VIQET scores and MOSp of the outdoor day images....94
Table 3.2. VIQET scores and MOSp of the indoor images....94
Table 3.3. VIQET scores and MOSp of the outdoor night image....95
Table 3.4. MOS and MOSp values of the outdoor day images....96
Table 3.5. Pearson correlation coefficients of the outdoor day images....96
Table 3.6. SROCC of the outdoor day images....97
Table 3.7. Outlier ratio of the outdoor day images....98
Table 3.8. AAE of the outdoor day images....99
Table 3.9. RMSE of the Outdoor Day images....99
Table 3.10. MOS and MOSp values of the indoor images....100
Table 3.11. Pearson correlation coefficients of the indoor images....100
Table 3.12. SROCC of the indoor images....101
Table 3.13. Outlier ratio of the indoor images....102
Table 3.14. AAE of the indoor images....102
Table 3.15. RMSE of the indoor images....103
Table 3.16. MOSi and MOSp values of the outdoor night images....103
Table 3.17. Pearson correlation coefficients of the outdoor night images....104
Table 3.18. SROCC of the outdoor night images....105
Table 3.19. Outlier ratio of the outdoor night images....106
Table 3.20. AAE of the outdoor night images....106
Table 3.21. RMSE of the Outdoor Night images....106

Appendix 1. Natural scenes images of test content

Test content was shot by the author and selected for both: the objective and subjective image quality assessment in *phase I* and *phase II*. Forty images of nature scenes were selected in four main categories:

Outdoor Day light



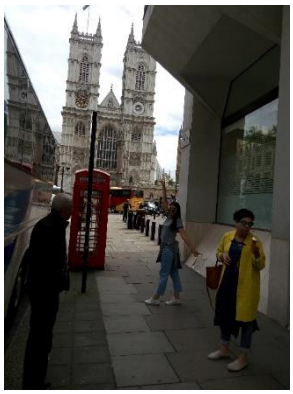
1.jpg



2.jpg



3.jpg



4.jpg



5.jpg



6.jpg



7.jpg



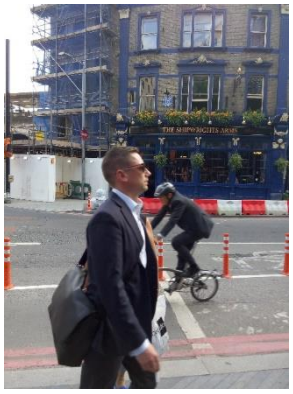
8.jpg



9.jpg



10.jp



11.jp



12.jp



13.jp



14.jp



15.jp



16.jp



17.jp

Indoor Arrangements



18.jp



19.jp



20.jp



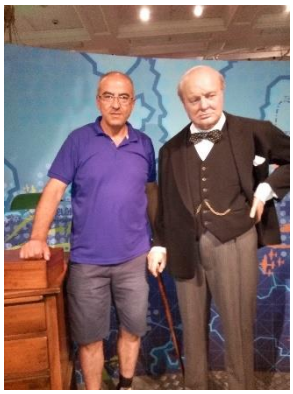
21.jp



22.jp



23.jp



24.jp



25.jp



26.jp

Indoor Wall hangs



27.jp



28.jp

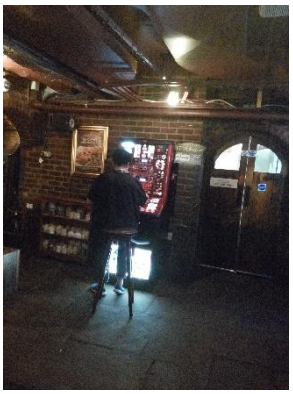


29.jp



30.jp

Outdoor Night



31.jp



32.jp



33.jp



34.jp



35.jp



36.jp



37.jp



38.jp



39.jp



40.jp

VQEG Image Quality Evaluation Tool (VIQET) User Manual

Contents

1. Introduction 1
2. System Requirements 2
3. Download 2
4. Install 2
5. Using VIQET to analyze photos 4
Capture Test Photos 4
General Guidelines 4
VIQET photo categories 5
Outdoor Day - Landscape 5
Indoor – Wall Hanging 6
Indoor – Arrangements 7
Outdoor Night – Landmark 8
Analyze Test Photos 9
6. Support 12
Known Issues 12
7. Uninstall 12
8. Index 12

1. Introduction

VQEG Image Quality Evaluation Tool (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 Mean Opinion Score (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. More information about VIQET and other VQEG projects is available at www.VQEG.org

2. System Requirements

In order to get the best experience with VIQET, ensure that your system meets the following requirements:

- Operating System: Windows® 8.1 or newer
- Processor: Intel® Core® i3 processor or faster
- Memory: 4GB or higher
- Minimum Free Disk Space: 1GB

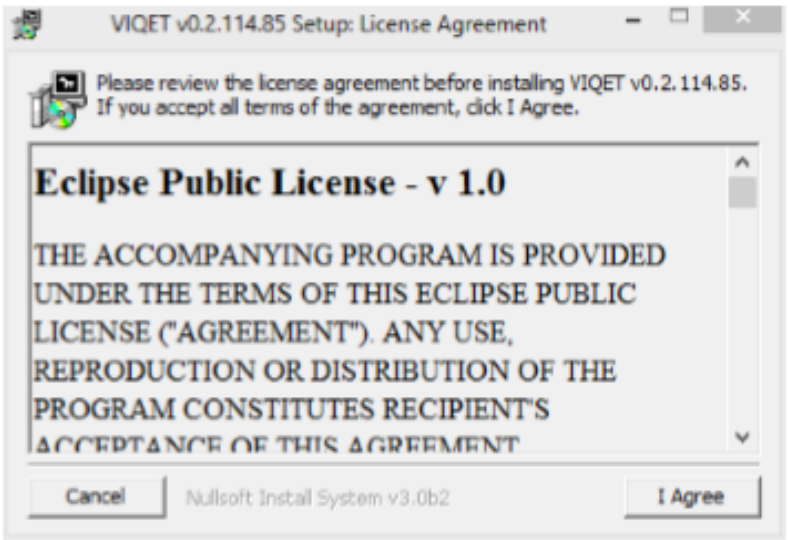
3. Download

The latest version of the VIQET software is available as a free download from https://github.com/VIQET/VIQET-Desktop/releases

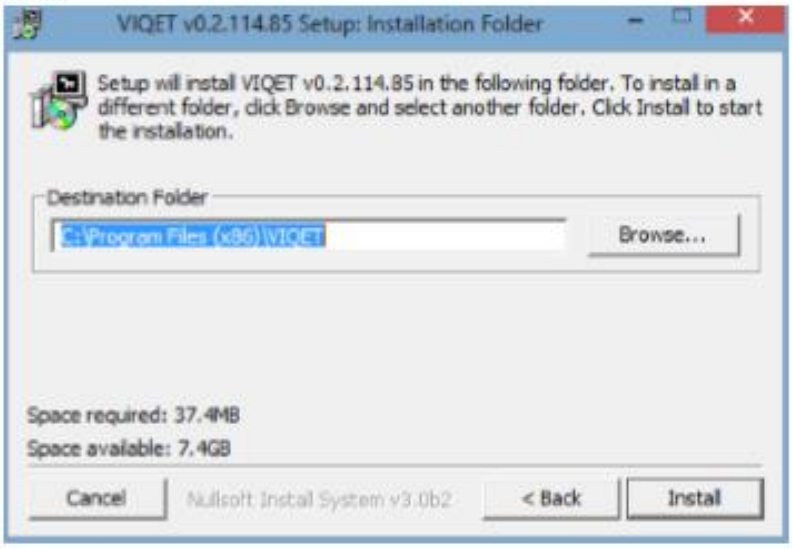
4. Install

In order to install VIQET on your PC, follow the instructions below:

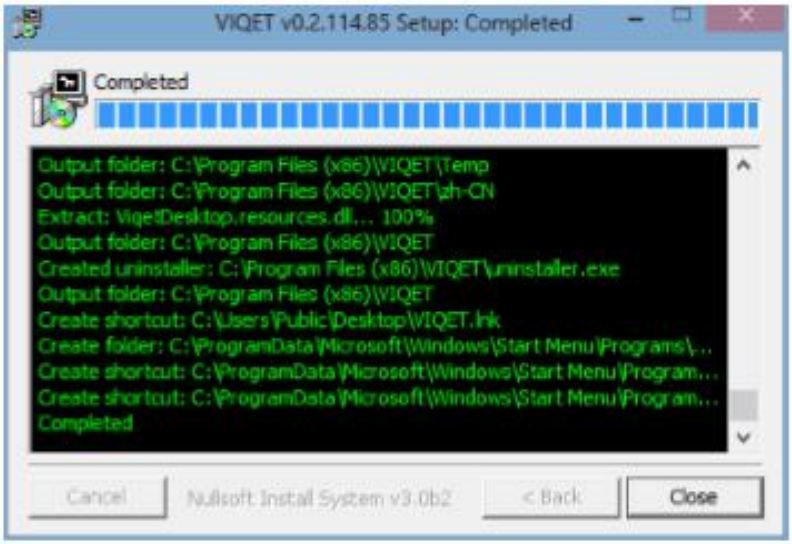
a. Download the setup file VIQET_INSTALLER.exe from the above Download link
b. Double click on the VIQET_INSTALLER.exe file to start the setup.
c. Review the license agreement and click the button that says 'I Agree' to proceed with the installation.



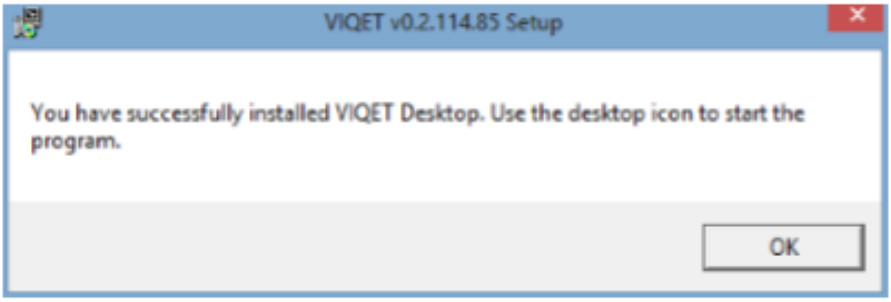
d. You will be prompted with the installation folder window. Select the folder where you want VIQET to be installed. The default is set to your program files folder. When ready, click on the Install button.



e. After the setup is complete, click on the Close button.

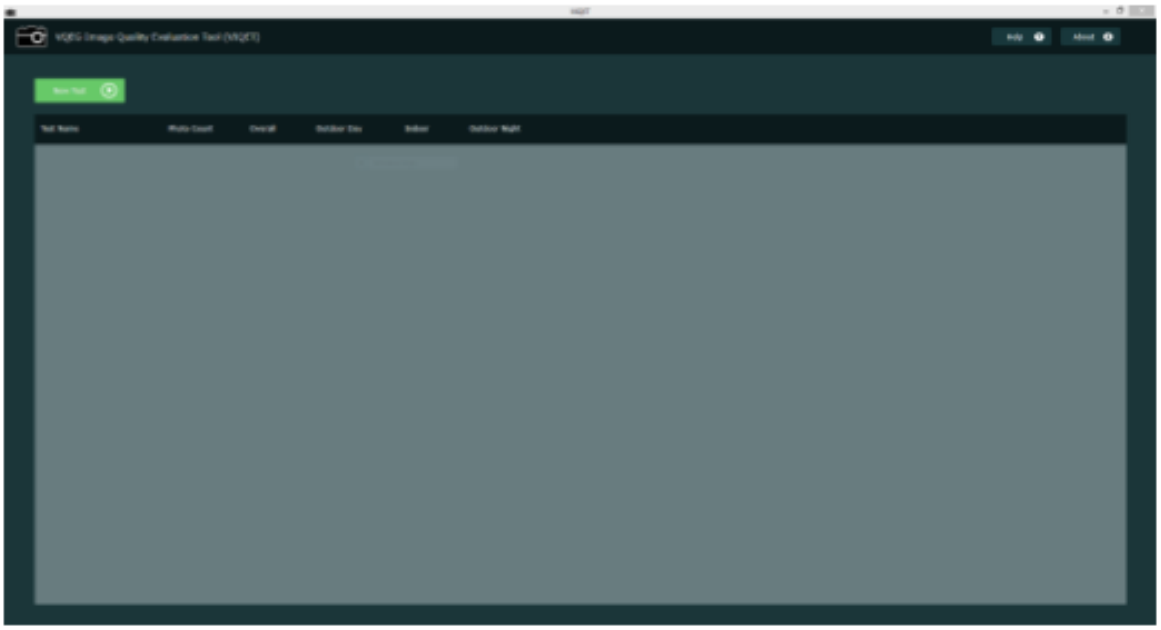


f. From the following dialog, click OK to proceed.



g. Click on the VIQET icon on your Desktop to launch the software. You are now ready to use the tool.





5. Using VIQET to analyze photos

There are 2 steps in this process. First is capturing test photos according VIQET methodology, and second, using the tool to analyze them. Below is a detailed description of these steps:

Capture Test Photos

General Guidelines

The following general guidelines for capturing test photos are crucial to obtain reliable VIQET results. Please ensure that you follow the guidelines below as closely as you can each time you take a photo for any of the four categories explained in the next section.

www.github.com/VIQET

i. Ensure that you capture 5 photos in each of the test photo categories mentioned in the next section
ii. Avoid direct light sources in all scenes
iii. Avoid moving and animated objects such as cars and people
iv. Avoid reflective surfaces such as glass or mirrors
v. Avoid photographer shadow

VIQET photo categories

VIQET requires test photos in four specific input categories. These are:

1. Outdoor Day - Landscape
2. Indoor – Wall Hanging
3. Indoor – Arrangements
4. Outdoor Night - Landmark

Note: If you are using VIQET to compare different devices, please capture the same scene by the devices. That means that you may have to move closer or further from the scene in order to capture the same photo. Also, ensure that you do not show new objects in the photo as you are switching devices to capture the same scene.

Below is a detailed description of each VIQET test photo category:

Outdoor Day - Landscape

Landscape photos should contain trees and grass. When capturing Outdoor landscape photos, you should:

- Avoid a direct light source (like the sun) in the photograph
- Maintain a distance of at least 15 feet away from the objects in your scene
- Outdoors illumination: Outdoor daylight > 300 lux

Below are some examples:







Indoor – Wall Hanging

Wall Hangings are highly detailed flat objects such as maps, posters, and paintings hanging on a non-textured wall. You could also use photos of business cards, receipts, or checks against a flat white background. When capturing wall hanging photos, you should:

- Ensure all corners of the object are showing in your photo
- Stand about 6 feet away from the object
- Surround your object with flat surface on all four sides. Approximately 20% of the image should show the flat background
- The area where the photo is taken should have normal lightning (between 100 – 300 lux).
- Avoid moving objects in the photo

www.github.com/VIQET

- Avoid having shadows in the photo

Below are some examples:



Indoor – Arrangements

An arrangement could be a fruit bowl or dinner plate that sits against a non-textured background. When taking photos of arrangements, you should:

- Stand about 3 feet away from the arrangement
- Surround your object with flat surface on all four sides. Approximately 20-30% of the image should show the flat background
- Recommended indoor illumination less than 100 lux which is equivalent to an indoor light scene.

Below are some examples:



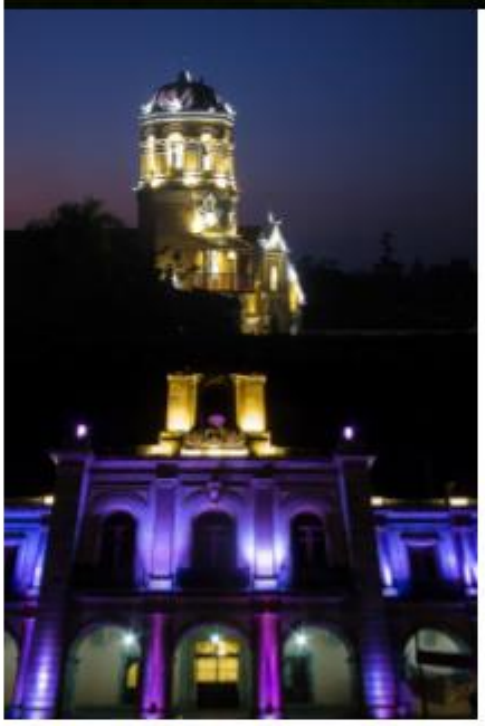
Outdoor Night – Landmark

A good example of a landmark is an outdoors monument or building. When capturing a landmark night photo, you should:

- Maintain a distance of at least 15 feet away from the objects in your scene
- Allow the sky around your landmark to show in about 20-30% of the photo
- Avoid any moving objects in the photo
- Avoid having a direct light source in the scene such as a street light.
- Illumination: around 20 lux

Below are some examples:





Analyze Test Photos

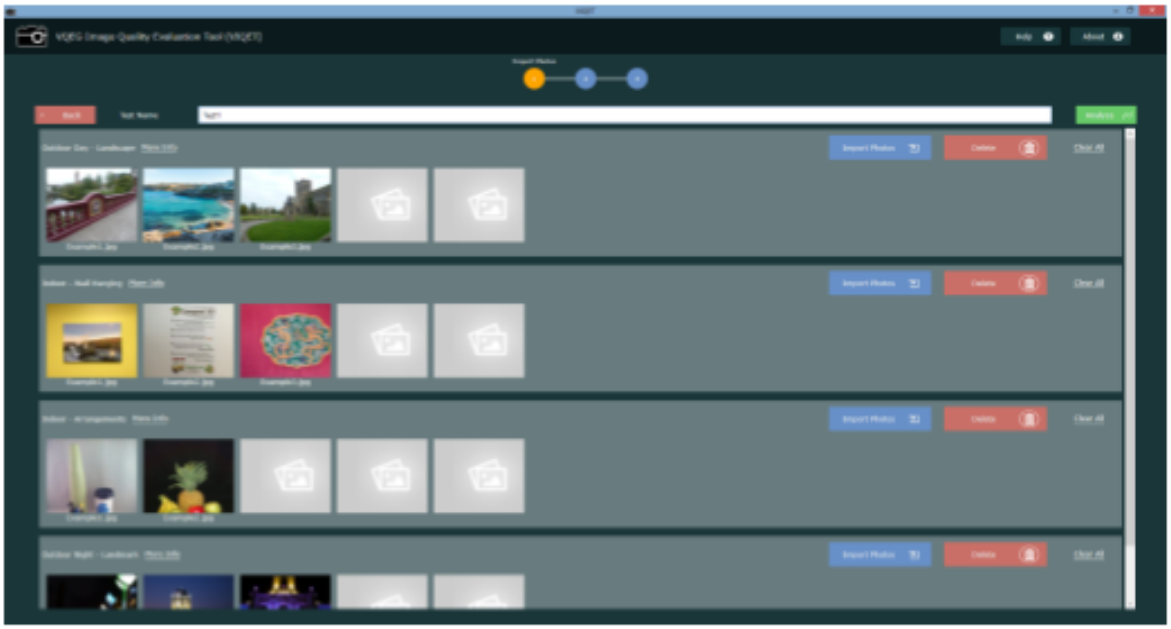
1. This section assumes you have captured photos according to the guidelines mentioned above. It also assumes you have already saved those photos onto your computer.
2. Launch the VIQET desktop tool. You are now on the home screen. To create a new test, click on the New Test button.

Note: *If you have used the tool previously, you will see that there are previous tests saved. To view your previous test, double click on the name. You will now see the results of that test. To go back to the main page, simply click on the Home button.*
To edit or delete a test, click on the name of the test and use the Edit or Delete buttons on the top right corner of the screen.

www.github.com/VIQET

3. You are now on the Import Photos screen. Enter the name of the test in the text box. Use a name that will help you identify the test content.

4. You have to import photos in all 4 photo categories. To do so, click on click on the Import Photos button in each category. You should import 5 photos in each test category for VIQET to compute and display category results.



5. After adding test photos, click on the Analyze button on the top right of the screen.

6. You are now on the photo analysis page. VIQET is now analyzing each photo that you have added to the test. To view a detailed description of each photo, click on the photo's thumbnail. VIQET will automatically display the test results after analyzing all photos in the test.

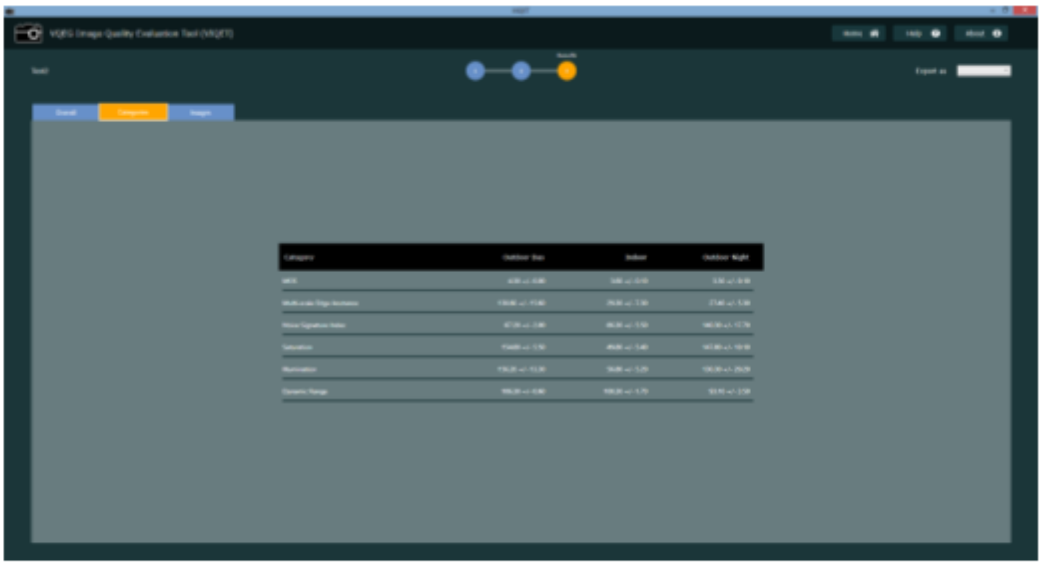


Note: You can choose to view the Sharpness map and an RGB histogram for each photo by choosing one of the options from the drop down menu.

7. You are now on the results page. You could view the Overall results and detailed results for individual Categories and Images by clicking on the corresponding tabs.



www.github.com/VIQET



8. To export this information, click the dropdown menu at the top right corner of the screen. You could export this information in PDF, CSV, or XML formats.

6. Support

Please submit feature requests, tool issues and bugs at https://github.com/VIQET/VIQET-Desktop/issues

Known Issues

1. In cases where VIQET is not able to identify a flat region in the test photo, the estimated MOS may be inaccurate. The tool will display the following message in the results screen - "VIQET was not able to identify flat regions in some test photos. This may be because those test photos do not follow VIQET methodology or might have high noise". Please verify that all test photos adhere to VIQET methodology, replace photos that do not, and repeat the test.

7. Uninstall

Launch uninstaller.exe from the install location to uninstall VIQET from your PC

8. Index

1. **Resolution:** refers to the total number of little squares (pixels) that make up an image. The higher the resolution, the higher the number of pixels.

2. **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.

www.github.com/VIQET

3. **Color Warmth:** refers to the tint of the overall image. Images with a bluish tint are considered to have cool colors. Images with an orange-reddish tint are considered to have warm colors.

4. **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)

5. **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.

6. **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.

7. **Multi-scale texture acutance**: refers to the level of activity and detail in the scene. The higher the level of detail in the scene the higher the feature value.

8. **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.

9. **% Over-exposed:** refers to the percentage of the image area that is covered in white. This feature ranges from 0 to 100. A higher percentage value indicates a larger area of an image is over-exposed.

10. **% Under-exposed**: refers to the percentage of the image area that is covered in black. This feature ranges from 0 to 100. A higher percentage value indicates a larger area of an image is under-exposed.

11. **Lux**: refers to a measure of light intensity that hits or passes through a surface, as perceived by the human eye.

Several test charts were used for different purposes in this research. In advance of subjective experiments, Ishihara Color Plates - 38 Set [94] Figure A3.1, was shown to all invited observers in order to confirm that none of the observers had color deficiency difficulty.

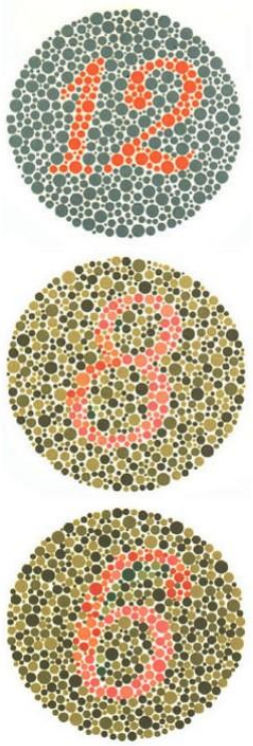


Plate 1

Everyone should see number 12.

Plate 2

Normal view: 8

Red-green deficiency: 3

Plate 3

Normal view: 6

Red-green deficiency: 5

Fig. A3.1 Examples of Ishihara color plates for color deficiency test, reproduced from *unlimitedmemory.tripod.com* [94]

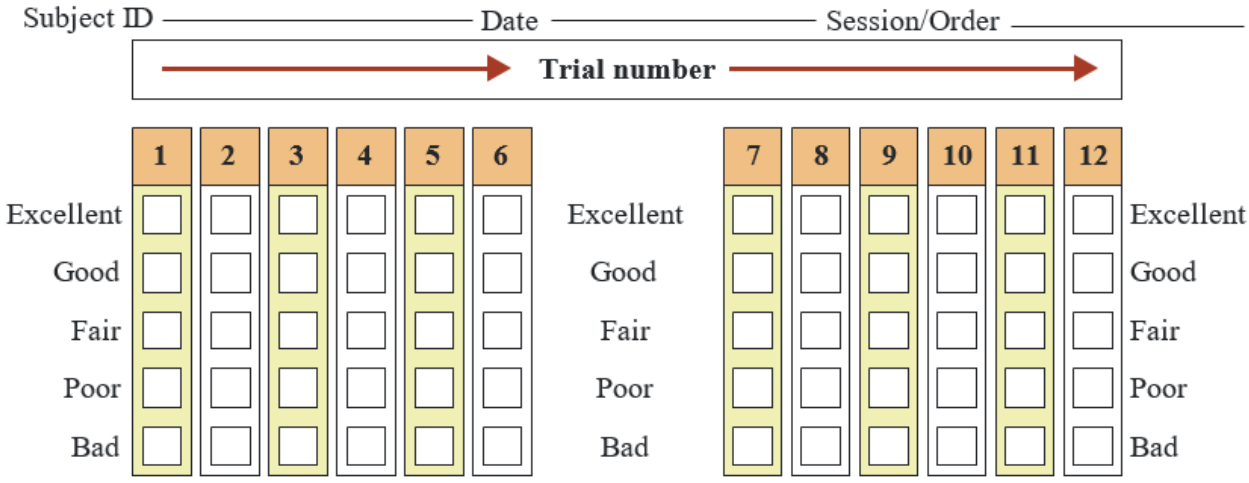


Fig. A4.1 Example paper ballot for the ACR method showing 12 stimuli [93]

אורט בראודה
המכללה האקדמית להנדסה
ORT Braude College

March 8, 2017

To whom it may be of concern:

This is to certify that Pinchas (Pini) Zorea teaches in our Electrical and Electronics Engineering Department at ORT Braude College since the year 2013.

He teaches in Image Processing labs and consults B.Sc. students in their final projects.

It is confirmed that Pini taught the theory of "Perceived image quality prediction" and his students performed successfully the Image Quality assessment experiments by using his new model.
If further information regarding this statement is necessary, please contact me.

Sincerely,

Dr. Nissim Sabag. Head of Electrical and Electronics department
nsabag@braude.ac.il

08 Mar. 17

Dr. Nissim Sabag, Senior Lecturer
Head of Department
Electrical & Electronic Engineering Dept.
ORT Braude College

המכללה האקדמית להנדסה אורט בראודה | רח' סנונית 51, כרמיאל 2161002, טל. 04-9901911 | מרכז מידע 1-700-70-80-05 | פקס. 04-9882016
ORT Braude College | P.O.Box 78, Karmiel 2161002, Israel | Tel: 972-4-9901911 | Fax: 972-4-9882016 | www.braude.ac.il

DECLARATION OF GUARANTOR

The undersigned hereby declares on the personal responsibility that materials presented in a doctoral thesis are the result of its own research and scientific achievements. Rationally understand that, otherwise, is to support the consequences in accordance with the legislation in force.

P.Zorea

Signature: 

Date: 10.01.2018



CURRICULUM VITAE

Personal data: Pinchas ZOREA, date of birth 30.06.1958, Marrakesh, Morocco.

Education:

1980 – 1982 Practical engineer in electronics, Tel - Aviv University, Israel.

1991 – 1994 B.Tech in electronics, academic engineering college, Jerusalem, Israel.

1995 – 1996 Several courses for M.Sc. degree in electronics engineering, University of Wisconsin, Milwaukee, USA.

1997 – 1999 MBA**,** master of business administration, University of Haifa, Israel.

2013 – 2017 Doctorate in engineering, specialty 232.02 – information technologies, products and systems, State University of Moldova, Chisinau, Moldova.

Employment history:

2013 – present lecturer in Braude engineering college in Israel, teaching: electrical engineering, microprocessors, HW and SW (C, C++,Assembly), control systems, analog and digital electronics, consult students on final projects.

2011 – 2013 Intel corporation, Israel development center, manager of image quality & system engineering team (in 2D & 3D digital camera project)**.**

2008 – 2011 Intel corporation, Israel development center, engineering manager of VTG IS (Post Si debug tools)**.**

2005 – 2008 Intel corporation, graphics development group, senior video systems architect & video image quality manager.

1999 – 2005 Oplus technologies, Israel, application engineering & customer support team manager**.**

1985 – 1999 Engineering colleges, Israel, instructor & Electrical Engineering dept. manager. managed study programs for students & instructors.

1984 – 1985 Orbit technologies, Israel, system designer. HW & SW computer control systems for antennae controller based on Motorola and Intel microcontrollers MCS-51 and 6805 series.

1976 – 1980 Army service in IDF, officer (major) in armored corps.

International organizations and forums:

CPIQ - Camera Phone Image Quality (by IEEE) - http://grouper.ieee.org/groups/1858/

VQEG - Video Quality Experts Group - http://www.its.bldrdoc.gov/vqeg/vqeg-home.aspx

IDF (Intel Developers Forum) - the annual conference held in Intel sites worldwide, provided lectures in image processing and image quality assessment.

National conferences:

The 11th Interdisciplinary Research Conference ORT Braude College , 2015

Zorea P., „New approach for modeling perceived Image Quality by Smartphone users”

ISBN 978-965-7468-01-2

International conferences:

International Conference Mathematics & information technologies (MITRE), 2015

Zorea P. „Modeling and predicting perceived image quality by smartphone users”

ISBN 978-9975-71-678-9

The 2-nd International Conference on Information Technologies, Systems and Networks (ITSN-2017), 17-18 October, Chisinau, Moldova: Zorea P., Paladi F., Bragaru T. *Image quality improvement based on the prediction theo*.

International experts forums:

Imaging and Visual Quality Seminar, 2015

Intel Corporation, Santa Clara, CA 95054

IEEE P1858™ CPIQ (Camera Phone Image Quality) – Work group.

Video Quality Experts Group, 2011, Hillsboro, Oregon, USA.

Publications:

- Zorea P. Perceived image quality prediction based on image quality attributes extraction. International Journal of Scientific & Engineering Research, Volume 8, Issue 2, February-2017, ISSN 2229-5518, p.1303-1312.
- Zorea P. New model for evaluation perceived image quality by Smartphone users. In: Studia Universitatis Moldaviae, 2015, nr.2 (82) ISSN 1857-2073, ISSN online 2345-1033, p.90-97.
- Zorea P. Prediction of smartphones’ perceived image quality by SW evaluation tool. In: Studia Universitatis Moldaviae, 2016, nr.7 (97) ISSN 1857-2073, ISSN online 2345-1033, p.170 - 180.
- Zorea P., Paladi F., Bragaru T. Prediction of perceived image quality by mobile devices users. In: Moldavian Journal of the Physical Sciences, Vol. 17, nr.1-2, 2018.