

A STUDY OF FACIAL ELECTROMYOGRAPHY  
FOR IMPROVING IMAGE QUALITY  
ASSESSMENT

by

Gutu Dorin

*A Thesis Submitted for the Degree of  
Doctor of Philosophy*

Graduate School of Science and Engineering for Education

University of Toyama

February 2015

# Acknowledgements

I would like to express the deepest sense of gratitude to my Ph.D. supervisor, Professor Yuukou Horita for welcoming me in the Media Information Processing Laboratory and for his continuous advice, excellent guidance and encouragement throughout the course of my research. I am extremely grateful to Dr. Yasuhiro Inazumi and Dr. Keiji Shibata for their help and enthusiastic support. I am thankful to Mr. Masaharu Sato, Ms. Yang Du and Mr. Takayuki Tamura of our laboratory for their cooperation. I would like to thank all of our laboratory members and all office staff of University of Toyama for their kind assistance and support.

I wish to express my love and gratitude to my wife Mila and daughter Emilia for their caring, patience and continuous support throughout the years. They have been the true source of motivation and inspiration that has guided me to the end. My warmest thanks go to my parents Slavian and Raisa, as well as my brother Marian, for their strong encouragement in pursuing my research and for giving me the strength to succeed. A special thank you to all my relatives and friends who believed in me, providing me vital motivational support.

# Abstract

The long term scope of our research is to develop objective evaluation methods for automatically assessing and improving Quality of Experience (QoE). In order to accomplish this goal, we need to identify the relationship between the activity of facial muscles and perceived image degradation. This crucial relationship is still completely unknown and that is why at this stage we are focusing on understanding and quantifying this relationship by using subjective assessment methods.

In this paper, we demonstrate the relationship between facial muscle activity and the subjective assessment score of image quality, by using Facial Electromyography (fEMG). We physically measure facial muscle activity and then obtain subjective scores of image degradation from subjects through questionnaires. By establishing a relation between these two results, we can assemble a MOS prediction method, by using regression analysis. Stepwise regression analysis has been conducted on an individual basis for each subject from the point of view of QoE. We do not need or attempt to average or deduct statistical models from our experiment, but rather identify specific relations between facial muscle activity and the sensation of image quality degradation.

Results show that if image quality is degraded, the activity of some muscles increases. The regression equation obtained has shown good results. We then advance the model further by introducing Image Entropy - a new set of variables, which focuses on the contents represented in the images, and the different features represented, compared to the general look at all images as equal input sets. This helped to significantly improve the accuracy of the prediction model.

By identifying a consistent relationship between fEMG and subjective assessment scores, we argue that Electromyography as a biometric method can be effectively applied in the field of image quality assessment.

# Table of Contents

Chapter 1. INTRODUCTION .....	1
1.1 Overview.....	1
1.2 Research Scope.....	3
1.3 Organization .....	3
CHAPTER 2. IMAGE QUALITY ASSESSMENT .....	4
2.1 Introduction.....	4
2.2 Quality of Experience and Quality of Service .....	6
2.2.1 Next Generation Network.....	8
2.3 Subjective Image Quality Assessment.....	9
2.3.1 Background.....	9
2.3.2 Image Quality Evaluation Technology Using Questionnaires/Surveys ..	10
2.3.2.1 ACR Rating.....	11
2.3.2.2 DCR Rating.....	12
2.3.3 Single Stimulus Method.....	12
2.3.4 Double Stimulus Method.....	13
2.3.5 Force-Choice Pairwise Method .....	14
2.3.6 Pairwise Similarity Judgements .....	14
2.3.7 Method Comparison .....	14
2.3.8 Application and Shortcomings .....	15
2.4 Objective Image Quality Assessment .....	16
2.5 Comparison and Analysis.....	18
CHAPTER 3. BIOMETRIC APPROACH .....	19
3.1 Image Quality Evaluation Technique Using Biometric Information.....	19
3.2 Cerebral Blood Flow (NIRS Technology).....	20
3.2.1.1 O2 Measurement Using NIRS.....	21
3.2.1.2 Detector Spacing and Photon Paths.....	22
3.2.2 Eye Movement .....	23

3.3	Conclusions .....	23
Chapter 4.	ELECTROMYOGRAPHY.....	24
4.1	Background .....	24
4.2	Electromyograph.....	24
4.3	Facial Electromyography.....	27
4.4	FEMG Advantages and Considerations .....	28
Chapter 5.	EXPERIMENTAL PROCESS .....	29
5.1	Purpose .....	29
5.2	Settings and Preparations.....	29
5.3	Electromyograph.....	30
5.3.1	fEMG Installation.....	30
5.4	Questionnaire Voting.....	32
5.5	Testing Procedure .....	33
5.5.1	Phase 1 – MVC Recording.....	33
5.5.2	Phase 2 – Image Quality Assessment by Questionnaires .....	35
Chapter 6.	RESULTS AND ANALYSIS.....	37
6.1	IEMG .....	37
6.2	Comparison of Muscle Activity .....	39
6.3	Regression Analysis.....	44
6.3.1	Stepwise Regression Analysis.....	45
6.4	Experimental Application .....	47
Chapter 7.	PERFORMANCE IMPROVEMENT USING IMAGE INFORMATION .....	51
7.1	Improvement Goals.....	51
7.2	Image Entropy .....	51
7.3	Multiple Regression Analysis.....	56
7.4	Results and Comparison.....	58
Chapter 8.	CONCLUSIONS.....	64
References.....		67

## Table of Figures

Figure 1. Workflow of MOSp prediction.....	2
Figure 2. Relationship of QoS and QoE.....	8
Figure 3. Presentation method of the ACR method .....	11
Figure 4. Presentation method of the DCR method .....	12
Figure 5. Subjective quality assessment methods, showing the timeline of each method .....	15
Figure 6. Hitachi ETG-4000 Optical Topography System.....	20
Figure 7. EEG-9100 on cart with flash lamp assembly and printer .....	25
Figure 8. EEG disk electrodes.....	26
Figure 9. Electrode junction box .....	31
Figure 10. Electrode map .....	31
Figure 11. Facial muscle groups and electrode placement.....	32
Figure 12. Voting questionnaire.....	33
Figure 13. Experimental Session.....	34
Figure 14(a). Test image example (Quality : 5 – Quality degradation clearly visible).....	35
Figure 14(b). Test image example (Quality : 30 – Quality degradation hardly visible) .....	35
Figure 15. Testing Sequence – subjective image quality assessment .....	36
Figure 16. Integrated Electromyogram .....	38
Figure 17. Relationship between predicted and subjective score for subject A1.....	49
Figure 18. Relationship between predicted and subjective score for subject A2.....	49
Figure 19. Relationship between predicted and subjective score for subject A3.....	49
Figure 20. Relationship between predicted and subjective score for subject A4.....	49
Figure 21. Relationship between predicted and subjective score for subject A5.....	49
Figure 22. Relationship between predicted and subjective score for subject A6.....	49
Figure 23. Relationship between predicted and subjective score for subject A7.....	50
Figure 24. Relationship between predicted and subjective score for subject A8.....	50
Figure 25. Visible degradation in test images .....	52
Figure 26. Test images at QS 30.....	53
Figure 27. Test images at QS 5.....	54
Figure 28. Image Entropy values of test images .....	55

Figure 29. Comparison of relationship between predicted and subjective score for subject A1 .....	60
Figure 30. Comparison of relationship between predicted and subjective score for subject A2 .....	60
Figure 31. Comparison of relationship between predicted and subjective score for subject A3 .....	61
Figure 32. Comparison of relationship between predicted and subjective score for subject A4 .....	61
Figure 33. Comparison of relationship between predicted and subjective score for subject A5 .....	62
Figure 34. Comparison of relationship between predicted and subjective score for subject A6 .....	62
Figure 35. Comparison of relationship between predicted and subjective score for subject A7 .....	63
Figure 36. Comparison of relationship between predicted and subjective score for subject A8 .....	63

## Tables Reference

Table 1. Image quality attributes.....	5
Table 2. Five Stage Interference Scale.....	13
Table 3. The strengths and weaknesses of objective and subjective approaches .....	18
Table 4. Conditions of the experiment .....	30
Table 5. Grading Scale .....	32
Table 6. Results of subject A1 .....	40
Table 7. Results of subject A2 .....	40
Table 8. Results of subject A3 .....	41
Table 9. Results of subject A4 .....	41
Table 10. Results of subject A5 .....	42
Table 11. Results of subject A6 .....	42
Table 12. Results of subject A7 .....	43
Table 13. Results of subject A8 .....	43
Table 14. Results of step-wise regression analysis .....	48
Table 15. Test Images Entropy Results.....	55
Table 16. Results of Multiple Regression using only Muscle Data .....	56
Table 17. Results of Multiple Regression using Muscle Data and Entropy.....	57
Table 18. Results of Multiple Regression using Muscle Data, Entropy and $\Delta$ entropy .....	57
Table 19. Results comparison (Muscle data only versus Muscle, Entropy and $\Delta$ entropy data) .....	58

# Chapter 1

## INTRODUCTION

### 1.1 Overview

In recent years, research in image quality assessment has been focused on the biological information approach [1]. This approach improves on the traditional questionnaire-based method, which suffers from major flaws caused by the difference in response results and the difference in rating scales between subjects. It has been argued that image quality assessment using biological information is less likely to suffer from these issues [2]. We focus our research on Facial Electromyography (fEMG) for biological information extraction.

In previous research [3,4,5], it has been established that facial expressions contain more emotional essence compared to the tone of voice or the content of spoken message [7]. We used this premise to hypothesize a significant relationship between facial muscle activity and image quality assessment and ultimately improving Quality of Experience [8].

The long term scope of our research is to develop objective evaluation methods for automatically assessing and improving Quality of Experience (QoE). In order to accomplish this goal, we need to identify the relationship between the activity of facial muscles and perceived image degradation. This crucial relationship is still completely unknown and that is why at this stage we are focusing on understanding and quantifying this relationship by using subjective assessment methods.

As can be seen from Figure 1, the current stage of our research has been actively focused on determining the reaction of muscle movements while a subject was observing image quality



degradation. We physically measure facial muscle activity using an electromyograph. At the same time, we asked subjects to assess image quality using questionnaires in order to obtain subjective scores of image degradation. By comparing the results of the electromyogram and voting by questionnaires, we can deduce a connection and assemble a MOS prediction method, by using the regression analysis [9].

Stepwise regression analysis [10] have been conducted on an individual basis for each subject from the point of view of QoE. We do not need or attempt to average or deduct statistical models from our experiment, but rather identify specific relations between facial muscle activity and the sensation of image quality degradation [11].

We have also advanced the model a step further, by introducing a new set of variables - Image Entropy, which focuses on the contents represented in the images, and the different features represented, compared to the general look at all images as equal input sets.

The next step in our research will be to solidify the results obtained through subjective testing, in parallel with more data from electromyography.

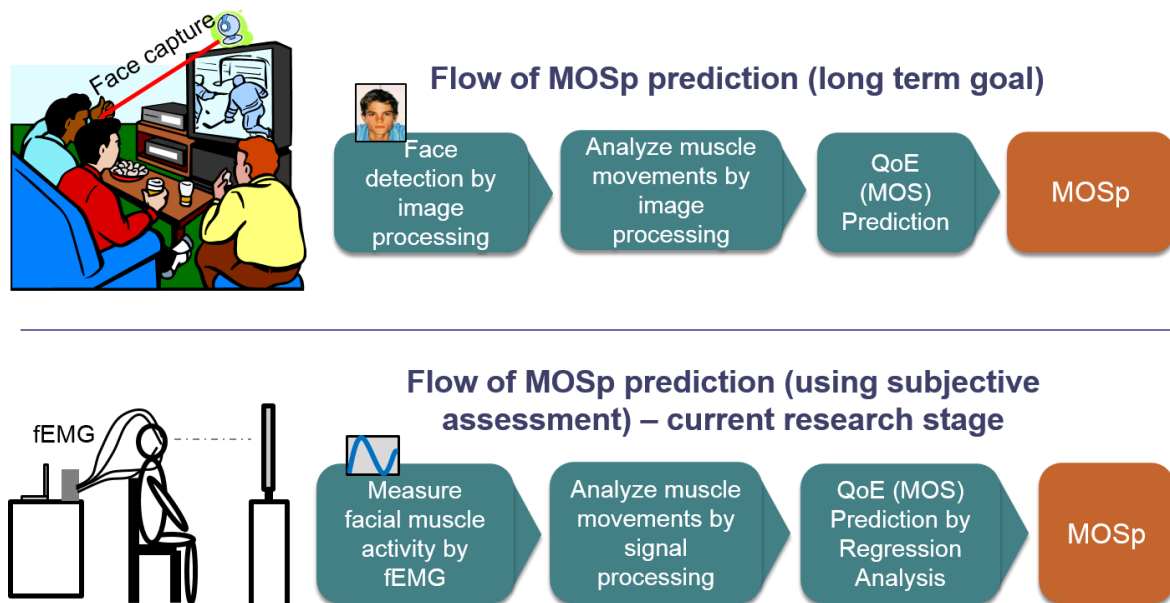


Figure 1. Workflow of MOSp prediction

The findings of this study can serve as a new paradigm in future research based on biological input for Image Quality Assessment. It can be used as a base for a multitude of applications that will replace the intrusive method of Electromyography with face detection and image processing or similar approach. Doing so, will help improve Quality of Experience for the users, which is the ultimate goal of Image Quality Assessment research.

---

## 1.2 Research Scope

The current study and its findings would have not been possible without a thorough examination of Image Quality Assessment (IQA), both subjective and objective, in order to identify the weak points on which research can improve upon. We have started with a wide scope, by scrutinizing current methodology and applications. The vast majority of IQA methodology is based on digital input, by focusing on image and video processing to assess quality, in order to develop quality improvement technology. Several years ago, biological information input has been proposed as a way to cover the lapses of existing technology and perhaps take IQA to a new level. By narrowing down our research on Facial Electromyography, our research has shifted to neuroscience in order to examine the possibility of applying this technology in the field of IQA. After a methodical examination of Facial Electromyography and its current applications, we have debated that it is probable that a connection can indeed be established. We have set out to implement this theory and test the relationship between facial electromyography and IMQ during our research. The main focus of our research was to conduct various experiments, develop algorithms for processing the results and lastly examine test results of by different methods of analysis. We have concluded that the statistical method of Stepwise Regression Analysis would provide the most accurate results relative to the data collected. The results show a promising new technological achievement that can have many useful applications in the field of IQA.

## 1.3 Organization

This paper is organized into eight chapters, each focusing on a specific topic. Besides the Introduction in Chapter 1, which offers a brief overview, we start with Chapter 2 - an in-depth examination of Image Quality Assessment, subjective and objective approaches, which will serve as the basis for further research. Chapter 3 examines the different biological input methodologies discovered so far, introduces the novel approach of biometrics to the field of Image Quality Assessment and prepares the set for our experiments. Chapter 4 is dedicated to the base technology used in this research - Facial Electromyography, and its impact on the methodology examined. In Chapter 5 we describe the experimental process in details. The results and analysis are detailed in Chapter 6, and a performance improvement is proposed in Chapter 7. Chapter 8 provides an insight into the conclusions and future research objective.

## Chapter 2

# IMAGE QUALITY ASSESSMENT

### 2.1 Introduction

Image quality can be defined as the characteristic of an image that evaluates the image degradation perceived by someone when compared to the original or perfect image. The quality here can also be understood as the subjective measure of how precisely or faultlessly something is represented in an image and whether the image contains faulty degradation that flaws the quality of said image. These degradations are caused by some amounts of distortion or artifacts introduced by imaging systems onto the signal.

Image Quality Assessment is a field of study focused on determining and improving the quality of images in order to meet up to the standards for the user quality expectations.

A list of attributes is detailed in Table 1 for image quality analysis, divided into four types (artifactual, preferential, aesthetic or personal) and scores them on a -1, 0, and +1 scale for three characteristics [129]:

1. Amenability to objective description (+1 : straightforward, 0 = difficult, -1 = nearly intractable);
2. similarity of first-party and third-party assessments (+1 = strongly correlated, 0 = partially correlated, -1 = nearly uncorrelated, );

3. degree of dependence on imaging system properties (+1 = strongly influenced, 0 = somewhat influenced, -1 = minimally influenced).

An attribute is specifically important if it has +1 scores for each category, making it objectively tractable, experimentally accessible, and pertinent to imaging system design. As an overall measure of suitability for analysis, 3 scores for each attribute are summed in the Total Score. There is a strong correlation between these attributes that are most amenable to objective description, most easily investigated experimentally, and those influenced by imaging product design [12].

Table 1. Image quality attributes

<b>Attribute by Type</b>	<b>Objective Tractability</b>	<b>1st/3rd Party Correlation</b>	<b>System Dependence</b>	<b>Total Score</b>
<b>Artifactual</b>				
Unsharpness	+1	+1	+1	+3
Graininess	+1	+1	+1	+3
Redeye	+1	+1	+1	+3
Digital Artifacts	+1	+1	+1	+3
<b>Preferential</b>				
Colour balance	+1	+1	+1	+3
Contrast	+1	0	+1	+2
Colourfulness (saturation)	+1	0	+1	+2
Memory Colour Reproduction	+1	0	+1	+2
<b>Aesthetic</b>				
Lighting Quality	0	0	0	0
Composition	0	0	0	0
<b>Personal</b>				
Preserving a cherished memory	-1	-1	-1	-3
Conveying a subject's essence	-1	-1	-1	-3

The scoring in Table 1 is coarsely quantized and may be subject to debate, however the general trend evident in the table is quite compelling, and would only be made stronger by the inclusion of a more extensive list of image quality attributes. Previous research [11] shows that there is a strong correlation between the attributes artifactually or preferentially in nature, those susceptible to objective description, those that could be studied through third-party evaluation, and those that are considerably affected by imaging system design. The attributes in agreement with these criteria assemble near the maximum total score possible. Personal attributes show the opposite behavior, and the total scores are at the low end of the scale. On the other hand, aesthetic attributes goes in a middle tier. This division of attributes into three tiers implies that a slightly restricted working definition of image quality is to be adopted, based on third-party assessment. The quality of an image is defined as an impression of its merit or excellence, as seen by an observer whom is not associated with the act of photography, or closely concerned with the subject matter.

This particular definition of image quality based on third-party assessment [11], captures the artifactual, preferential and aesthetic aspects, but excludes personal attributes. The included attributes are mostly traceable, experimentally accessible and important in system design. With broader concepts of image quality and customer satisfaction, the so defined image quality will correspond well but will be predominantly free of the ambiguities related to the personal attributes. Although these excluded attributes can generate significant discrepancies in individual evaluations, distributions of assessments for observers and collections of images should be similar.

Therefore, while excluding all attributes affecting a photographer's satisfaction with an image, third-party image quality is well-defined and practical. So, taking all these into considerations, image quality might be usefully quantified.

## 2.2 Quality of Experience and Quality of Service

This section aims to provide a better understanding of Quality of Service (QoS) and Quality of Experience (QoE) in the context of Image Quality Analysis, which is our main research field.

Today millions of users are consuming a huge amount of media, over television and computer networks. Social websites like Facebook, Youtube, Instagram and hundreds more are facilitating image and video exchange on an unprecedented scale. In technical terms, any amount of data shared over the Internet will be turned into small packets to enable the transmission of data packages, which is then distributed and processed back for display to the user. In this paper, we focus our attention on still images and moving images (video) and their impact on users and their experience as viewers of images (Quality of Experience). Images and video are compressed to reduce the size in order to ease the transmission of data through networks. Unfortunately, this process will cause degradation of the

image quality in the form of noises because of the loss of information during the compression process, leading to lower Quality of Service.

Previous studies on QoS have that it represents the characteristics that are crucial to actualize the service the users expect it to be. Since there are many layers that contribute to the actual end-to-end level of service [13], there are numerous ways to describe the QoS for each layer [14]. Also, the management components of the architecture maps the QoS so there would be no need for the user to do any mapping of the QoS parameters into the parameters of the layers underlying the whole system. This allows the users to designate the QoS essentials effortlessly.

When too much data is transferred over the network, some areas face serious problems of network congestion and its quality of service decline. This deterioration of the quality of service leads to many networking implications including the loss of packets during data transmission. This also leads to the degradation of the quality of the transferred image data and consequently reduces Quality of Service.

The long term scope of our research is to develop objective evaluation methods for automatically assessing and improving QoE. In order to accomplish this goal, we need to identify the relationship between the activity of facial muscles and perceived image degradation. This crucial relationship is still completely unknown and that is why at this stage we are focusing on understanding and quantifying this relationship by using subjective assessment methods.

Furthermore, based on previous studies [15], QoS plays a crucial role in the case of end-to-end service. Since network functionalities corresponds to each service according to their requirements, the management of network traffic and the differentiation of service are closely related to the QoS structure. The higher the QoS of every layer in the service, the better the overall service itself. The better the overall service is, the better the service provided to the end-user. In other words, in order to ensure that the user can have a pleasurable experience while using the network service, we need to ensure that a good end-to-end QoS can be provided. How the end-user perceive this service quality is defined as QoE [16].

Through the use of this feature, in the case of the service provided to the user who is the final recipient, meaning as in end-to-end, it is possible to control the QoE. Here, the relationship between the QoS and QoE [17, 18] is shown in Figure 2. While QoS is a physical quantity that represents the IP network and the nature of the line, QoE is the quality that the user actually felt towards the provided services.

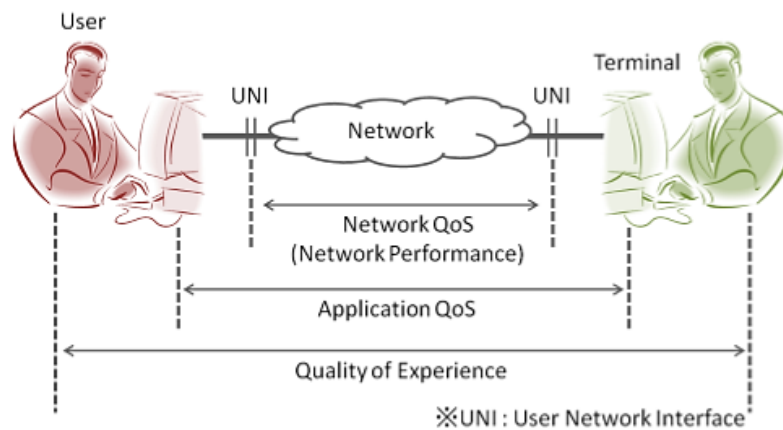


Figure 2. Relationship of QoS and QoE

### 2.2.1 Next Generation Network

To overcome the problem of information transportation, a new network service known as the New Generation Network (NGN) was introduced. According to International Telecommunication Union Telecommunication Standardization Sector, ITU-T [19], the definition of Next Generation Network is the following:

“A Next Generation Network (NGN) is a packet-based network able to provide services including Telecommunication Services and able to make use of multiple broadband, QoS-enabled transport technologies and in which service-related functions are independent from underlying transport-related technologies. It offers unrestricted access by users to different service providers. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.”

From this definition, we derive that NGN architecture is a packet-based transfer network service and has a vital key feature of separating the main functional levels. In other words, NGN separates the service layer, transport layer and control layer on a horizontal platform so that the layers can be technically and commercially be provided by different market players. Thus, to ensure overall service interoperability, we need to make sure that no problems occur in the interconnection between all functional levels and that the linkage is possible in an adequate manner at all times.

In a conventional IP network, the upper limit of the packet loss and the lower limit of the communication bandwidth is not compensated, so the quality of communication may decrease due to network congestion. Meanwhile, NGN has a characteristic feature that is capable of controlling the network QoS. In other words, NGN is a network that can keep the classified bandwidth, and packet loss to a variation within the scope of a certain range.

During the transition to the NGN, the service control of QoE in IP networks' has been studied and it is found that the possibility to control the QoE in the services provided on the NGN future is required. As a mean of controlling the QoE, the degree of coding and coding scheme for each user is changed and the distribution of the content that satisfies the user even a little has been considered. In addition, to deliver the contents that satisfy every each one of the users, the ideal encoding control is the encoding control that is performed by taking into consideration of the individual characteristics of individual users [20]. Therefore, in order to meet the complex demands of a user in the IP network and control the QoE, indicators for assessing the QoE becomes necessary. In order to assess the QoE, previous research on quality evaluation of multimedia delivery services has been carried out, evaluation experiments and the like due to the subjective evaluation method has also been studied [21, 22, 23].

In addition, NGN's most relevant characteristic to our field of study is the NGN's capability to control the quality of service offered to end-users. The NGN service is a service which revolves around the user or customer making it a customer-centered approach. This is different from the conventional internet service approach where the main focus was the technology used in the service. This NGN's major change of approach requires the management to play an important role as important as that of the technology itself to enable a smooth management of the customers.

A study on NGN in 2008 has stated that, in an NGN environment, the management of the network is controlled by the subscription and requested service of the customers [24]. In other words, when the network receives a customer request for a service, the network will assign the service quality. Moreover, in the case of network congestion, the service priority will then be used for traffic engineering to control the service quality provided to the customer.

## 2.3 Subjective Image Quality Assessment

### 2.3.1 Background

The subjective image quality assessment method comes from a large group of psychometric scaling methods that were used to measure psychological attributes [25]. One attribute that describes the inclination for a certain image rendering is image quality. The significance of image and video quality had been focused mainly on video compression and transmission applications, which in turn, leads to some recommendations for the design of quality assessment experiment [26, 27, 28]. These studies describe various experimental procedures, viewing conditions, display calibration parameters and the methods for experimental data processing. The aim of these experimental procedures is to find scalar-valued "quality correlate" that would point out the deterioration level (in case of video quality)



or overall quality. There is a work discussing how to interpret such quality in the context of rating rendering methods [29], and it also focused on a usually neglected topic in the quality assessment literature which is statistical testing. The standards, for example in [26], describe the detailed procedures but lacking wider context, statistical background and are limited to the suggested techniques. One can find more information on the scaling methods used for image quality assessment in [30] and [31]. An example is psychometric methods which is not new to computer graphics. Recent SIGGRAPH courses, such as [32] and [33], indicate the increasing interest in them. This method had been used to scale a light reflection model in a perceptually meaningful space, or to find the best set of parameters for tone mapping [34] or color correction [35].

Another application of experimental methods (the most outstanding), is comparison and validation of the results produced by graphics algorithms. While there are an increasing number of publications that accompany quality comparison studies, there are also work solely on comparison of existing methods, like tone mapping operators [36, 37] or image retargeting methods [38]. However, these studies benefit little from the research that has been dedicated to image quality assessment. Work devoted to comparing subjective quality assessment methods is comparatively small. A few examples, such as Dijk et al. [39]; compared the direct ranking method (category scaling) with similarity judgments (functional measurement) and found that the results of direct ranking can be biased when the evaluated distortions are of a very different nature [40], compared eight direct rating methods and confirmed very high correlation between their results.

There are commonly four types of quality assessment experimental methods that show the broad spectrum of experimental procedures and they are the “Single stimulus categorical rating”, the “Double stimulus categorical rating”, the “Ordering by force-choice pairwise comparison” and lastly the “Pairwise similarity judgments methods” [29], which will be detailed later in this chapter.

### 2.3.2 Image quality evaluation technology using questionnaire/surveys

In the subjective evaluation method using a questionnaire, human subjects are asked to observe image or video contents and evaluate quality according to some criteria. In order to obtain the evaluation value for the image, this subjective evaluation experiment needs to be carried out.

For such methodology, all experiment planning and implementation, including presentation method of the image to be evaluated, the selection of appropriate measures to evaluate the target image and the adjustment of the experimental environment, requires laboratory equipment and expertise in order to conduct effective experiments [41, 42]. All specifications and recommendations for conducting questionnaire-based tests are detailed in in the ITU’s BT.500 [43].

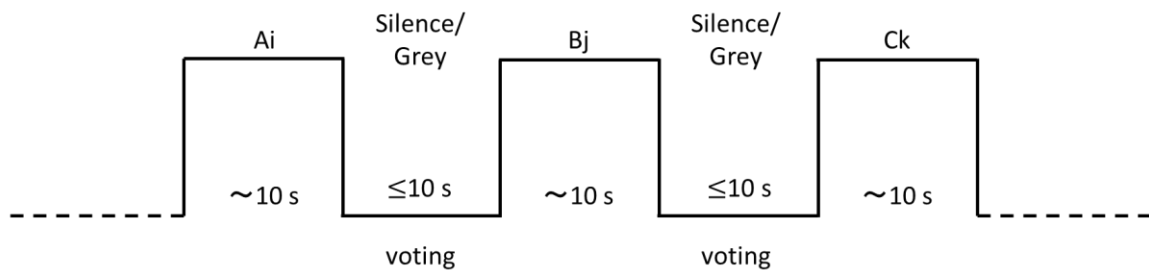
There are two types of subjective evaluation methods that can be broadly divided into relative evaluation and absolute evaluation. The absolute evaluation presents only the content to be evaluated,

and to assess the quality absolutely. In contrast, in relative evaluation, the quality is evaluated relatively by comparing the content to be evaluated and the reference content which is the evaluation criteria.

The two methods are defined as Absolute Evaluation Method, or Absolute Category Rating (ACR) and the Degradation Category Rating (DCR). Both methods will be explained further in this chapter. In our study, the method used is the ACR, which will allow us to synchronize the results of the subjective tests, with the data acquired through Facial Electromyography.

### 2.3.2.1 ACR Rating

ACR method, as shown in Figure 3, presents only the degraded image to a subject. The single stimulation method (described further in Chapter 2.3.3) is applied here to obtain the absolute evaluation of the contents. In ACR, for each content to be evaluated, the subject is required to choose one of the scores from the five-stage rating scale, as shown in Table 2. The scores for the evaluation content are used as the Mean Opinion Score (MOS) in dealing with the average value between the subjects.



- Ai      Sequence A under test condition i
- Bj      Sequence B under test condition j
- Ck      Sequence C under test condition k

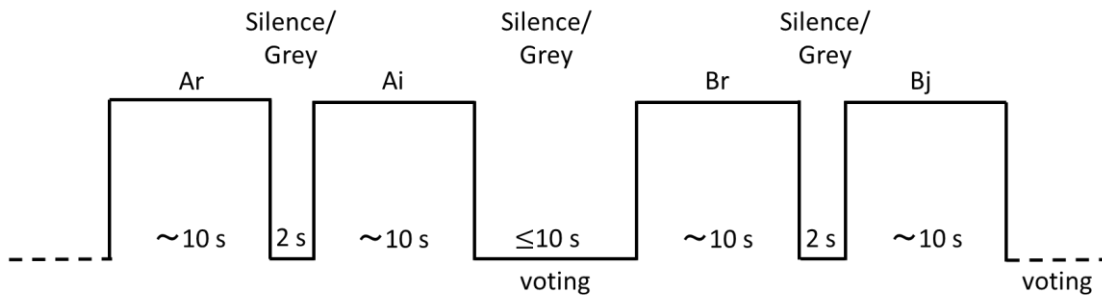
Figure 3. Presentation method of the ACR method

### 2.3.2.2 DCR Rating

As shown in Figure 4, subjects are presented with contents in pairs of the reference contents which is the evaluation criteria and the contents to be evaluated. This one pair of contents will be presented continuously one after another in the duration of 10 seconds. The subjects are then required to evaluate the contents by choosing from the 5-steps interference scale as shown in Table 2.

In DCR, during the evaluations, the subjects are required to look at the reference contents first so that the order effect that happens in ACR can be suppressed.

Eventually, the evaluation scores are dealt the same way as ACR's MOS but in order to distinguish it from the ACR method, the DCR's evaluation scores are called Degradation MOS (DMOS). Since the DCR method uses the comparison of the evaluation of the paired contents, the deterioration can be more sensitively evaluated than the ACR method. Therefore, when using evaluation contents with small coding degradation, DCR method is suitable than the ACR method. However, in DCR, since it is necessary to view a pair of contents to evaluate a single content, it will take 2 times longer than the ACR method to obtain the evaluation value for the same number of contents.



- Ai      Sequence A under test condition i
- Ar,Br    Sequence A and B respectively in the reference source format
- Bj      Sequence B under test condition j

Figure 4. Presentation method of the DCR method

### 2.3.3 Single stimulus method

In the single stimulus categorical rating method, an image is displayed for a short and fixed time. Then, the observer rate the image based on these five categories; excellent, good, fair, poor or bad. Although the five-point Likert-type scale is an approach commonly used for scaling responses, some methods prefer continuous instead of categorical scales to avoid quantization artifacts. All images,

including reference images, are shown in random order one at a time. There is no time limit during voting but no image is shown during that voting time.

### 2.3.4 Double stimulus method

Double stimulus categorical rating is analogous to the single-stimulus method. There are two types of double stimulus categorical scale which are Double Stimulus Continuous Scale (DSCQS) and Double Stimulus Impairment Scale (DSIS) [44], in which we used the latter for our research. In DSCQS, subjects are shown multiple sequence pairs consisting of a ‘reference’ and a ‘test’ sequence. The reference and test sequence are presented twice alternately, and the order of the two are chosen randomly for each cycle. Subjects will not be informed which one is the reference and which one is the test image. They rate each image on a continuous quality scale ranging from ‘bad’ to ‘excellent’. On the other hand, in the DSIS method, the reference is always shown before the test sequence and neither is repeated. Subjects rate the amount of impairment in the test sequence on a five-level scale rating, ranging from ‘very annoying’ to ‘imperceptible’. The DSIS method is relevant for evaluating clearly visible impairments such as artifacts caused by transmission errors [44].

In this method, subjects are presented with contents in pairs of the reference, which is the evaluation criteria and the contents to be evaluated. This one pair of contents will be presented continuously one after another in the duration of 10 seconds. Subjects are then required to evaluate the contents by choosing from the 5-steps interference scale as shown in Table 2.

Table 2. Five Stage Interference Scale

<b>Score</b>	<b>Degradation Scale</b>
<b>5</b>	Imperceptible
<b>4</b>	Perceptible
<b>3</b>	Slightly annoying
<b>2</b>	Annoying
<b>1</b>	Very annoying

This method will be used on our testing, as it translates the results in the closest match with the results provided by Facial Electromyogram.

---

### 2.3.5 Force-choice pairwise method

In ordering by force-choice pairwise comparison method, observers are shown pairs of same images corresponding to different conditions and asked to choose which of each pair is of higher quality than the other. Even if they don't see any difference between the two images shown, they must choose one, thus a forced-choice design. Like single-stimulus categorical rating method, there is no time limit or minimum time when making the choice. And comparing to rating methods, it is more likely to be accurate because of the straightforwardness. However, it also requires more trials to compare each possible pair of conditions.

In the Forced-choice methods images are rated according to quality, but it does not indicate the difference between the images. Compared to the pairwise similarity judgment methods, observers are asked to mark which image they prefer and indicate on a continuous scale how large the difference in quality is between the two images. If they see no difference in quality between the two images, they can choose to put the marker in the '0' position. The ranking of the image is decided by the position of the marker, whether if it is more on the left or right side of '0'. If the position of marker is at '0', the images are ranked randomly.

### 2.3.6 Pairwise similarity judgements

While the forced-choice method orders images according to quality, it does not tell how different the images are. In pairwise similarity judgments, observers are not only asked to mark their preference, but also to indicate on a continuous scale how large the difference in quality is between the two images (see Figure 5). Observers can choose to leave the marker in the '0' position if they see no difference between the pair. The sorting algorithm used for the pairwise comparisons can also be used for the similarity judgments. The position of the marker (on the left or right side of '0') decides on the ranking of the image pair. If '0' is selected, the images are ranked randomly.

### 2.3.7 Method comparison

The four subjective image quality assessment methods are shown side-by-side in Figure 5, in order to picture the timeline and differences between each method. In our experiments described in Chapter 5, we will use the Double Stimulus Method, because it will provide results that are comparable to data acquired through facial electromyography.

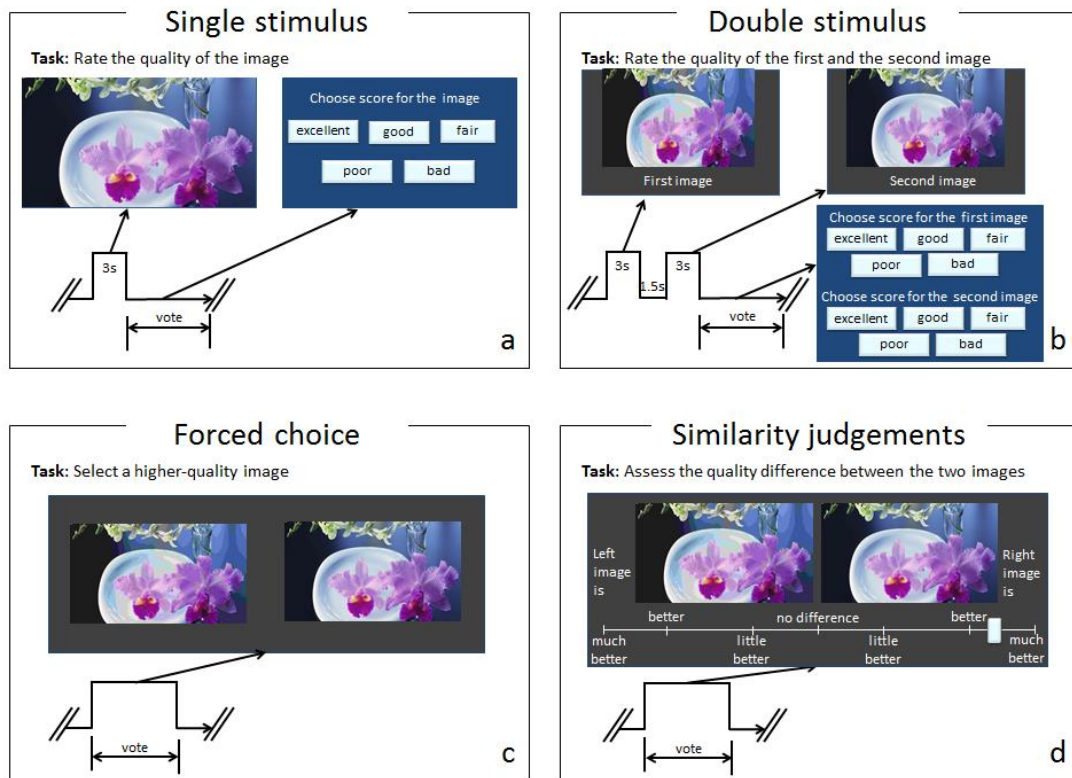


Figure 5. Subjective quality assessment methods, showing the timeline of each method

### 2.3.8 Application and Shortcomings

There are many examples in which the subjective image quality approach is the best solution, as its results are acquired directly from the people, which are the ultimate judges of image or video quality. From this point of view, subjective assessments have been used originally in the development of any other objective or computer-based algorithms for image quality assessment. The efficiency ranking of such methods have to be directly compared with the original assessments taken through questionnaires, or other subjective methods.

The subjective approach is applied, for example, when investigating aerial image quality. It is important to predict the quality potential of an aerial imaging system, produce aerial photographs of the best quality possible, and also to be able to assess the quality of image at hand. Image quality demands have been investigated for different applications of stereoscopic evaluation, namely block triangulation, large and small scale topography map compilation, profiling and DEM measurement, ortho-photo generation and image interpretation. The investigation was made in the form of a questionnaire, directed to stereo operators. The evaluation is conducted in real working situations, with the images adjusted in stereo instruments. Experienced stereo operators judged the quality of image pairs in subjective scales for contrast, sharpness, granularity, interpretability and overall

quality. The results varied with different applications. Subjective assessments are being used to derive frequency weight functions for integration of modulation transfer functions into image quality marks.

Its advantage obvious at a first glance, the subjective approach, one of the shortcomings of subjective assessment is that it is cognitive mediated, which is depended on the observer. For example, users find significantly lower media quality acceptable when a perception of financial cost was attached to the level of quality [45], which in fact has been determined to be unsatisfactory in a number of experimental and field studies. Another finding is that when the task being performed is difficult, the same video quality receives lower ratings. This proves users' subjective assessment of quality can be influenced by contextual variables [45].

Subjective approach is an approach that collects direct responses from users. Because of that, it is the most convincing approach but because of the same reason, it is also inconvenient, expensive and time consuming [46].

## 2.4 Objective Image Quality Assessment

The aim of objective image quality assessment research is to design computational models that can predict perceived image quality accurately and automatically. The term predict is used here, since the numerical measures of quality that an algorithm provides are useless unless they correlate well with human subjectivity. This means that the algorithm should predict the quality of an image that an average human observer will report.

Objective image quality measures has great potential in a wide range of application environments. One of the applications of objective image quality measures is they can be used to monitor image quality in quality control systems [47]. For instance, an image acquisition system can use a quality metric to monitor and automatically adjust itself to obtain the best quality image data. Or a network video server can examine the quality of the digital video transmitted on the network to control and allocate streaming resources [47].

Another application of objective image quality assessment is that they can be selected to benchmark image-processing systems and algorithms. For example, if a number of image denoising [46] and restoration algorithms are available to enhance the quality of images captured using digital cameras, then a quality metric can be used to determine which of them provides the best quality results.

Next, objective image quality assessment can also be applied in image-processing systems and transmission systems to optimize the systems and parameters settings. For instance, in a visual communication system, an image quality measure can assist in the optimal design of the prefiltering

---

and bit assignment algorithms at the encoder and of optional reconstruction, error concealment, and postfiltering algorithms at the decoder.

Existing methods for image quality assessment are mostly based on simple mathematical methods like the mean squared error (MSE) [34], essentially because of lack of knowledge regarding both the Human Visual System (HVS) and the structure and statistics of natural image. It is also convenient in the context of design optimization, owing to the analytic and computational simplicity of these measures.

Through objective approach, most previous researchers [48] evaluate interpolated frame quality using fidelity metrics. For example, Peak-Signal-to-Ratio (PSNR), a normalized Mean-Square-Error (MSE) between original and processed image, is used to measure the interpolation quality or Structure Similarity (SSIM) metric, which combines luminance, contrast and pixel value correlation comparisons as the quality index [48].

The predictive performance of objective quality assessment in relative to subjective human quality assessment has generally been quite poor. While these methods for quality assessment have found ample use as an analytic metrics for theoretical algorithm design, they have long been thought as rather weak for assessing the quality of real images, processed or otherwise. Up until the last decade, the field of image quality assessment remained in a declining state. Research into image quality assessment was nearly nonexistent thanks to a lack of driving forces in the form of new models for human visual perception of images, or of image formation, natural image structure, and natural scene statistics.

Objective image quality assessment methods are designed to emulate the assessment of human observers. To do so, such automated processes must overcome the drawbacks of the MSE and other  $l_p$  norms. There are multiple image quality measures currently available, even though it there is no certain defined classification of all methodologies. In general, there 3 types of information that can be used for designing image quality assessments: information about the “original image”, information about the distortion process, and information about the Human Visual System (HVS).



## 2.5 Comparison and Analysis

In Table 3 we summarize the strengths and weaknesses of the two image quality assessment methodologies, by underlining the most important characteristics relative to each other.

Table 3. The strengths and weaknesses of objective and subjective approaches

	<b>Objective approach</b>	<b>Subjective approach</b>
<b>Strengths</b>	Complexity is reduced	Complexities of the social world can be explored
	Casual connections are more easily made	Deeper meanings may be uncovered
	Suited to the study of behavior rather than strategic intentions	Reluctance of entrepreneurs to report may be overcome
		May uncover the deeper meanings of strategic intentions
<b>Weaknesses</b>	Conclusions may be simplistic	Research may conclude without any clear findings or contribution to practice or policy
	Nuances or explanations outside conceptual framework may be ignored	
	Not suited to finding the deeper meanings of strategic intentions	

For most image-processing applications, the best and most reliable method for assessing image quality is by subjective evaluation, because people are the ultimate receivers in most image-processing applications. The mean opinion score (MOS), a subjective quality measure that requires the direct involvement of a number of human subjects, is still regarded as the best method of image quality assessment. However, the obvious downside of subjective assessment is that MOS method is expensive and relatively slow for many applications.

## Chapter 3

# BIOMETRIC APPROACH

### 3.1 Image Quality Evaluation Techniques Using Biometric Information

Biometric information refers to any signal information output from the human body, such as brain waves, body temperature, electrical signals, fingerprints and others. Biological information is collected and analyzed for various scientific applications in a wide variety of fields. In recent years, biological information input has been proposed as a new paradigm in the field of Image Quality Assessment, previously using image, video or computer-based methods exclusively as the method of input. By analyzing the biological information in parallel with subjective image quality assessment, we propose a new approach in assessing image quality with great prospects for further development.

Other studies have been focused on the assessment and enhancement of fingerprint quality [50], the assessment and enhancement of the quality of biometric data regarding irises [51], and face quality assessment and enhancement [52, 53, 54, 55, 56, 57, 58]. In all researches, the algorithms proposed to enhance and assess the biometric data are different and unique.

During our experiments (Chapter 5), we measure and analyze biometric information in the form of electric signals produced by facial muscles, which will later be compared with results produced by a subjective assessment procedure.

Previous research conducted [59], including by our own team at MICT Laboratory, University of Toyama, have produced good results in using biometric information in the field of image quality assessment, and will be described further in this chapter.

## 3.2 Cerebral blood flow (NIRS Technology)

Cerebral blood flow (CBF), is the blood supply to the brain in a given time. In an adult, CBF is typically 750 milliliters per minute or 15% of the cardiac output and is tightly regulated to meet the brain's metabolic demands. Cerebral blood flow is determined by a number of factors, such as viscosity of blood, how dilated blood vessels are, and the net pressure of the flow of blood into the brain, known as cerebral perfusion pressure, which is determined by the body's blood pressure. Cerebral blood vessels are able to change the flow of blood through them by altering their diameters in a process called autoregulation; they constrict when systemic blood pressure is raised and dilate when it is lowered [60]. Arterioles also constrict and dilate in response to different chemical concentrations. For example, they dilate in response to higher levels of carbon dioxide in the blood and constrict to lower levels of carbon dioxide [60]

Because the amount of blood flow varies according to the state of activity the brain, this is used as an index for observing the brain activity. Among others, a method using Near Infra-Red Spectroscopy (NIRS) is widely used in researches [90]. NIRS uses a special device, like the one depicted in Figure 6 (Hitachi ETG-4000), which is used to record cerebral blood activity, information that can be further used in various applications. In the field of image quality assessment, the relationship of blood flow and the amount of oxygenated hemoglobin has been examined [61].



*Figure 6. Hitachi ETG-4000 Optical Topography System*

---

Near-infrared spectroscopy (NIRS) technology, such as that used in pulse oximetry, has been used and trusted in the world of medicine for decades. Several characteristics contribute to its widespread use, including its noninvasive nature, reliability and safety. Another example of NIRS equipment, is the Somanetics' INVOS System [62] which is used to identify hemoglobin and red-colored oxygenated hemoglobin molecules within red blood cells. It can measure the relative amounts of each to determine whether or not there is adequate oxygenation. Since brain cells and organ tissues die within minutes without proper oxygenation, this measurement provides potentially life-saving or life-changing information. When its regional oxygen saturation (rSO<sub>2</sub>) value shows a change in blood oxygenation toward or beyond threshold levels, the care team can intervene to potentially lessen or prevent complications.

### 3.2.1.1 *O<sub>2</sub> Measurement Using NIRS*

Near infrared (NIR) light photons penetrate the forehead and/or body tissue of interest to the clinician. After being scattered inside the scalp, brain or tissue, some fraction of the injected photons survive, returning to and exiting the skin (a property called "reflectance"). By measuring the quantity of returning photons as a function of wavelength, we can infer the spectral absorption of the underlying tissue and make some conclusions about its average oxygenation [82].

Human tissue is translucent to NIR photons having wavelengths between about 650 and 1100 nm. To see this first-hand, you can conduct your own experiment by illuminating your own tissues in a darkened room with a common red laser "pointer" used for slide presentations. These laser pointers produce light in the near-infrared band (typically 670 nm). You'll see that this light is easily transmitted through thin body parts (e.g., cheek, ear, fingers, etc.) and a "back-scattered" halo of light can be observed from thick tissues.

In the latter case, light can be seen emerging from the skin at distances of a couple of centimeters from the point of injection (even farther for fatty tissue). In the absence of light absorbing materials, some photons will continue penetrating the tissue to considerable depths before meandering back to the surface at the point where a detector is located. Although this light is readily transmitted through human tissue, the scattering prevents it from being useful for imaging.

What is bad for imaging is good for spectroscopy. The long, torturous paths taken by the scattered photons make them exceedingly sensitive to the optical properties of tissue. Even small amounts of colored materials ("chromophores") will cause wavelength-dependent absorption of photons which produces characteristic signatures in the spectrum of the emerging light. As early as 1977, Jöbsis reported measuring the absorption spectrum of NIR light passing through the head of a cat and was also able to get enough light through the human brain from temple to contralateral temple to detect an increase in light transmission during hyperventilation.

Coloration of hemoglobin molecules is directly related to the oxygen they carry. The chromophore with the highest absorption in body tissue is in the 280 million, red-colored hemoglobin molecules found within each of the 1013 red blood cells circulating in the blood. It looks red in white light because it absorbs shorter wavelengths (green and blue).

Hemoglobin is of vital importance to us because it transports oxygen from the lungs to the cells of the body which cannot live without it. The exact shade of red of each hemoglobin molecule depends on the amount of oxygen it is carrying, a property that forms the basis of a number of blood oxygenation measurement devices ("oximeters"). Specialized equipment, like the Somanetics INVOS Cerebral/Somatic Oximeter, is used to measure oxygen in brain or tissues directly beneath the sensor using two wavelengths, 730 and 810 nm, to measure changes in regional oxygen saturation (rSO<sub>2</sub> index). Surface data from the skin and skull is subtracted out, to produce an rSO<sub>2</sub> value for deeper tissues.

### 3.2.1.2 *Detector Spacing and Photon Paths*

Most photons reaching the detector will have taken some optimum course through the tissues of interest. Using sensitive photodiode detectors, light can be measured at considerable distances from the point of injection. The greater the separation of source and detector, the greater the average depth of penetration. Photons that happen to meander close to the surface are very likely to be lost out of the skin before reaching a distant detector. Large source-detector spacings are therefore biased against "shallow" photons except in the tissues directly under the source and detector. On the other hand, geometry and absorption also make it unlikely that very deeply penetrating photons will find their way back to the detector. Therefore, mean photon path is neither deep nor shallow, but a moderate curve like a banana or canoe.

From previous studies [63] we learned that the most likely penetration depth is about a third of the spacing between the light source and the detector. They used a tank filled with a liquid scattering material (intralipid) which approximated tissue and noted the changes in light received by the detector, at various source-detector distances, as they inserted and removed small absorbers (black cylinders 2.5 mm diameter by 1 mm thick) at different depths. They reported the same "banana-shaped" sensitivity distributions found by others using both experiments and computer simulations.

Results have shown been confirmed by Hongo et al [64], in the human forehead by injecting a bolus of infrared absorbing dye (indocyanine green) into the internal carotid artery and observing the transient decrease in signals at various source-detector spacings. The larger signals at increasing source-detector spacings indicated deeper penetration into the head and the very short duration of the signals (~5 seconds) verified cerebral circulation as the source.

---

Dr. Hongo's validation study was performed by injecting infrared absorbing green dye in the internal and external carotid arteries. The deep and shallow signals changed equally when dye was injected in the external carotid which feeds the external structures. The net difference is close to zero. However, when dye was injected into the internal carotid, providing blood to the cerebral cortex, the deep and shallow signals changed differently, corresponding to the amount of light that reaches the brain.

### 3.2.2 Eye Movement

Another area of interest is the eye movement, which has been developed in various applications, like eye-tracking car mirrors, smart helmets and others. It is possible to observe where people are interested in and holding their focus on by keeping track of and analyzing their eye movements. The eye movements can be separated into two types; in the situation that a person thinks that he wants to look at something carefully, the eye movement will come to a stop, this is called the gaze state and the other condition is when the eyeball darts around quickly called as the state.

In the field of image quality assessment that make use of this phenomenon, it is examined whether there are any differences in the effect of different types of multimedia content source on the fixation point of the line of sight [65].

## 3.3 Conclusions

There are several biometric methodologies which have been recently applied towards the field of Image Quality Assessment [81]. Progress is slow, because it requires specialization in two fundamentally different areas (Neuroscience and IT). A lot of focus has been put on assessing and enhancing biometric data so that they can be significant enough to be used in applications. However, research in this direction is still in its infant state, with great potential for new developments in this area. We believe that the biometric approach provides the framework required for the next big leap in the evolution of Quality of Experience.

## Chapter 4

### ELECTROMYOGRAPHY

#### 4.1 Background

Electromyography (EMG) is a technique for evaluating and recording the electrical activity produced by skeletal muscles. It is performed using an instrument called an electromyograph and produces a record called an electromyogram, by detecting electrical potential generated by muscle cells. The signals can be used to analyze biomechanics and muscle activity.

The use of electromyography can be applied in areas such as in neurological departments, neurosurgical clinics or research, in orthopedic clinics during spine operations and many more.

Even though the main application of electromyography is in neuroscience, for the first time we have challenged the possibility of applying this technology in the field of Image Quality Assessment. The testing conducted during our experiments has proven a tangible and feasible application in evaluating muscle electro-responses corresponding to subjects' subjective assessment values.

#### 4.2 Electromyograph

An electromyography device, or electromyograph is able to carry out a great number of neurophysical examinations, they can record, amplify, and measure the bioelectric activity of

---

muscles and nerves under various circumstances, in which each examination has its own stimulation, recording, display, and evaluation characteristics [66]. The various parameter settings are stored in the machine, and the user could change or even design them again. Electromyograph's various amplifiers are very sensitive, have a high input impedance, a large bandwidth, and very low noise.

The principal construction of an electromyograph is shown in Figure 7. The central controller, signal processor and the storage of the EMG machine is operated by a personal computer (PC) or microcontroller (MC) and is equipped with a hard disk as data storage, PC monitor with color display, keyboard, and laser printing. The essential components (amplifier, analog-to-digital converter, stimulation units) are integrated into the PC housing. Software processes functions such as averaging, signal delay, programming of the recording programs, and evaluation of the signals. Triggered signals are averaged and processed also by software and this helps to extract small stimulus-related signals from surrounding signal mixture. The delay line delays signals so that signal components produced before stimulus could be displayed. A loudspeaker evaluates the recorded signals acoustically and this is vital for immediate diagnostics because of the characteristic sound images that myographic signal shapes have. The loudspeaker also provides feedback for the subject, so the subject could judge the strength of his relaxation or voluntary contraction.



*Figure 7. EEG-9100 on cart with flash lamp assembly and printer*



An EMG signal is recorded by electrodes (shown in Figure 8), connected from an electrode junction box, which are attached to the skin overlying a muscle consists of the electrical activity from numerous individual motor units within the detection region and therefore represents the sum of numerous MUAPs, where some are canceled out and others are intensified, as the different motor units are not aligned. The frequency of EMG spans from a few hertz to approximately 500 Hz and spread across a wide range of amplitudes. The characteristics of other electrical signals generated by the body, such as the electrocardiogram and electroencephalogram, and ambient signals, such as 50~60-Hz electrical noise emitted by alternating current (AC)-powered equipment, overlap with the characteristics of the EMG mentioned above.



Shielded EEG disk electrodes,  
BE-911A (24 pcs/set)  
BE-912A (12 pcs/set)

*Figure 8. EEG disk electrodes*

There are several types of surface electrodes, having different advantages and disadvantages [67]. There are generally two types of electrodes. Needle electrodes are inserted into the muscle to be examined, whilst surface electrodes record the electric activity of numerous motor units. The basic difference is that needle electrodes are able to record the activity of a single motor unit or a few units, while the surface electrodes record the electric activity of numerous motor units. The electromyograph contains one or more systems of amplification that can multiply the weak electric potentials recorded from the muscle. For example, if the signal amplitude is 1 microV, 1000-fold amplification reproduces a potential, on the oscilloscope, which has an amplitude of 1 cm. Bipolar Aq/AgCl surface electrodes are most commonly used for the recording of facial EMG in psychological experiments [68]. By using two contact surfaces, bipolar surface electrodes measure the voltage difference between two closely spaced locations. A differential amplifier takes the difference in electrical potential between the two signals and amplifies it. This difference signal is referred to a third signal obtained from a monopolar reference electrode or ground, which is located

at an electrically neutral place (for facial EMG, usually the middle forehead, but earlobes are also an option). By referring the signal from the two measurement electrodes to the reference electrode, the rejection of extraneous electrical signals mentioned above could be carried out, because these signals, are picked up by the reference electrode as well. Then the differential amplifier rejects signals that are common with respect to the reference electrode, and only the difference in potential between the bipolar electrodes is retained. The amplitude of the recorded signal depends on a few things such as the muscle fiber diameter, the distance between the fiber and detection site, and electrode properties. Basically a larger surface area allows better measurement, as independence is lower. But when facial EMG is measured, electrode size be kept small because; one, the larger the surface, the more likely is it that activity from other muscles is recorded as well (“cross-talk”), and two, the larger electrodes may interfere with the movement of the face.

The best way is to place the two electrodes near the middle of the muscle, on a line parallel to muscle fiber [107] to get the best result. The head and face have over 30 bilaterally symmetrical muscle pairs. One challenge is where to place an electrode so as to have a good measure of the muscles’ activity in combination with as little cross-talk as possible from other muscles.

To match numerous purposes, EMG machines are built of system components in a modular way. They vary in relation to the number of channels and the number and type of stimulators, software, and storage medium.

### 4.3 Facial Electromyography

Facial (mimetic) muscles are muscles that control facial expressions and the device used to measure the activity of mimetic muscles is called facial electromyography. Facial electromyography (fEMG) has a few advantages such as it is a precise and sensitive method since it does not depend upon languages nor does it require cognitive effort or memory. fEMG is also capable of registering responses even when subjects were instructed to not show any emotional expression. Moreover, fEMG is less intrusive than other physiological measures like fMRI and EEG [68]. This technique also yields a lot of data, is continuous and scalar (hence more credible).

Striated muscles are groups of bundles of individual muscle fibers. They are packed with myofibrils, with interdigitated sets of the proteins actin and myosin. Muscle contraction happens when small extensions of myosin protein (the “cross-bridges”) pull actin over the myosin, which happens when calcium is released. Groups of activated muscle fibers are called “motor units” [75]. Each motor unit is innervated by a single motor axon. When a motor neuron is activated, all of the muscle fibers in that motor unit are enervated. During muscle contraction, the conduction of action potentials along

---

the muscle fibers (in other words, the motor units) causes electrical potential, and EMG records these changes (motor unit action potentials, or MUAPs).

Facial EMG gives many advantages for the measurement of facial behavior. Especially, because of the high temporal and spatial resolution of facial EMG, the measurement of fleeting and subtle movements, which may easily escape the naked eye, could be detected.

However, the use of facial EMG requires researcher to specify in advance which muscles will be of interest in a given context. For certain applications this requirement is not a problem but if facial EMG is used to assess emotional expressions, the potential for misidentification is higher.

## 4.4 FEMG Advantages and Considerations

Facial EMG has a lot of advantages compared to other biometric methods described in Chapter 3:

- Facial (mimetic) muscles control facial expressions
- fEMG is a precise and sensitive method
- Does not require cognitive effort or memory
- fEMG does not depend upon language and does not require cognitive effort or memory.
- fEMG is capable of registering the response even when subjects were instructed to inhibit their emotional expression.
- Yields a lot of data and is continuous and scalar (hence more credible)
- It is able to measure facial muscle activities to even weakly evocative emotional stimuli
- Less intrusive than other physiological measures like fMRI and EEG.
- Like other physiological measures, facial EMG measurement technique is often the only useful approach when movement is not visible.

It is because of the multitude of benefits of this technology, at the initial stage of our research, we have estimated that Facial EMG will be the best method for understanding the relationship between biometric information and subjective image quality assessment. Following the course of our investigation, the method proved to have better than expected results, which will be explained in Chapter 6.

## Chapter 5

# EXPERIMENTAL PROCESS

### 5.1 Purpose

The goal of our experiment is to test and established a connection between subjective image quality assessment methodology and the results obtained by facial electromyography. In order to accomplish this goal, we conduct a series of questionnaire-based tests, while at the same time capturing subjects' facial muscle responses to images. Testing should be scalable and provide examinable results. The complete experiment procedure will be explained in this chapter.

### 5.2 Settings and Preparations

We have tested a total of 8 subjects during the first phase of our experiment. All subjects participated on a volunteer basis, for the purpose of scientific research. Testing was conducted in the specially-equipped laboratory, at the Media and Information and Communication Technology Laboratory, Faculty of Engineering, University of Toyama, Japan.

Before conducting the experiments, each participant was carefully instructed on the proceeding and safety guidance. Any discomfort, or inability to follow our instructions would allow the subject to

discontinue the experiment. During our testing, we have not received any complaints or premature termination at any phase of the experiments.

Each subject took part in the experiment separately, seated in front of a 32 inch LCD display and followed the testing procedure, which will be detailed below. The viewing distance and test image specifications are detailed in Table 4.

Table 4. Conditions of the experiment

Subjects	8 Adults (male, females)
Luminance	75-150 [lx]
Display	32 inch LCD (1920 x 1080 [pixel])
Viewing distance	3H (H : Height of image)
Image resolution	1920 × 1080 [pixel]
Encoding	JPEG (Quality : 5, 30)
Number of images	Reference: 6, Total: 12

## 5.3 Electromyograph

For the testing platform we have used the EEG-9100J/K (Neurofax  $\mu$ ) Electromyograph (shown in Figure 7, Chapter 4.2), manufactured by Nihon Kohden [69] – a world leading manufacturer of high quality electroencephalographs. The equipment set includes a Windows-based laptop PC, a photo control unit, an electrode junction box, disk electrodes and a flash light assembly (as shown in Figure 7). The included software provided with the Neurofax  $\mu$  can perform a number of EEG monitoring and processing tasks, including 8 channel DSA trendgraph, Digital video option, EEG Examination Support software, Voltage mapping, EEG central monitoring, Automatic photic stimulation and others.

### 5.3.1 fEMG Installation

During the preparation stage, we attached 17 electrodes from the junction box of the EEG9100 electromyograph (see Figure 9) to the skin of the subject's face. The attachment was adjusted to the required positions for recording electrical signals from 6 facial muscle groups (explained further in this chapter). The electrode map (Figure 10) provided specific position and connection requirements. The electrical signals generated by muscle flexing are transmitted through the electrodes back to the

computer for signal analysis, by the EEG software. This allowed us to record muscle activity during the experiment and develop our theory.



Figure 9. Electrode junction box

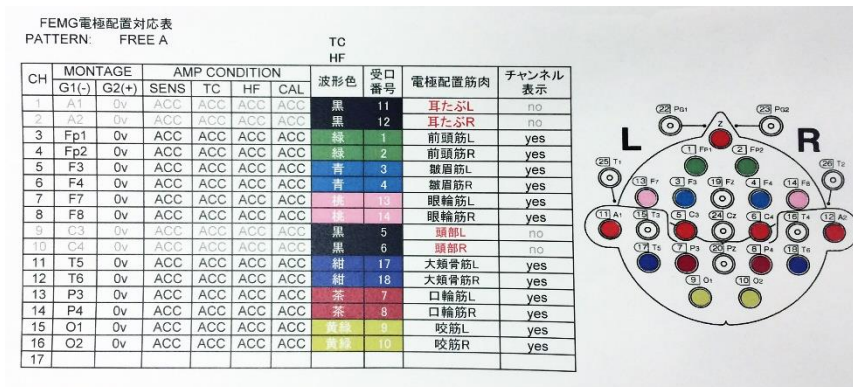


Figure 10. Electrode map

The fEMG measurement area is comprised of six muscle groups, as shown in Figure 13:

1. Venter Frontalis (M1)
2. Corrugator Supercilii (M2)
3. Orbicularis Oculi (M3)
4. Zygomaticus Major (M4)
5. Orbicularis Oris (M5)
6. Masseter (M6).

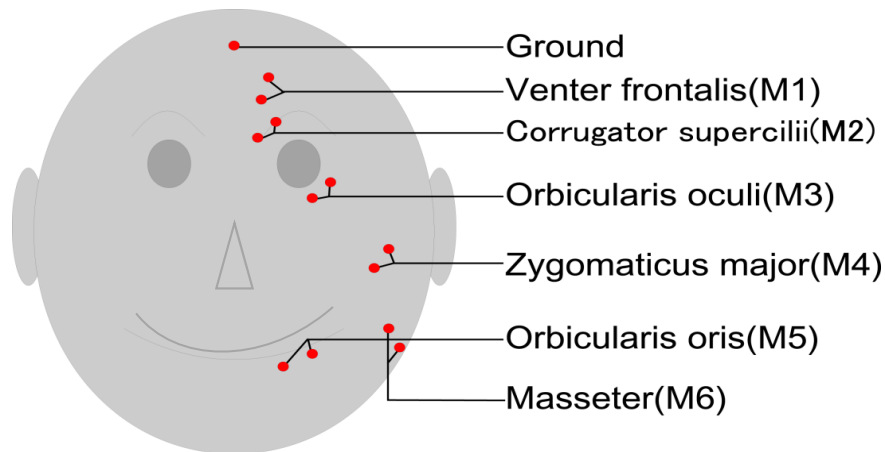


Figure 11. Facial muscle groups and electrode placement

These muscles are closely associated with the formation of facial expressions as described in previous studies [70]. We argue that the examination of these specific muscles' reaction to image quality degradation will provide the factual proof in our theory.

The ground was attached to the forehead area (Figure 11) and fEMG was derived in the bipolar lead. The sampling frequency used was 1 kHz.

## 5.4 Questionnaire Voting

The questionnaires used in the actual experiment are presented in in Figure 12. The grading scale used for voting is detailed in Table 5. We have asked subjects to look at various images and rate the quality and degradation. The results have been used to test our theory.

Table 5. Grading Scale

Score	Image quality assessment vote
5	Good quality, degradation imperceptible
4	Degradation perceptible, but not annoying
3	slightly annoying
2	Annoying
1	very annoying

回答用紙(DSIS) ( - - - )

・評価の各段階

5:劣化がわからない  
 4:劣化がわかるが気にならない  
 3:劣化が気になるが, 邪魔にならない  
 2:劣化が邪魔になる  
 1:劣化が非常に邪魔になる

	1	2	3	4	5
例:	----- ----- ----- -----				
no.1:	----- ----- ----- -----				
no.2:	----- ----- ----- -----				
no.3:	----- ----- ----- -----				
no.4:	----- ----- ----- -----				
no.5:	----- ----- ----- -----				
no.6:	----- ----- ----- -----				

1

Figure 12. Voting questionnaire

## 5.5 Testing Procedure

### 5.5.1 Phase 1 – MVC Recording

After completing the preparation phase, subjects were seated in front of the display and were given the questionnaires for voting. Electrodes have been already attached to facial muscles. The first part of testing included the establishment of the muscles Maximum Voluntary Contraction (MVC) [80]. In medical terms, MVC is defined as “the maximum force which a muscle can produce in a specific isometric exercise” [71]. In practice, MVC is more accurately recorded as the best of three efforts in a single test session.

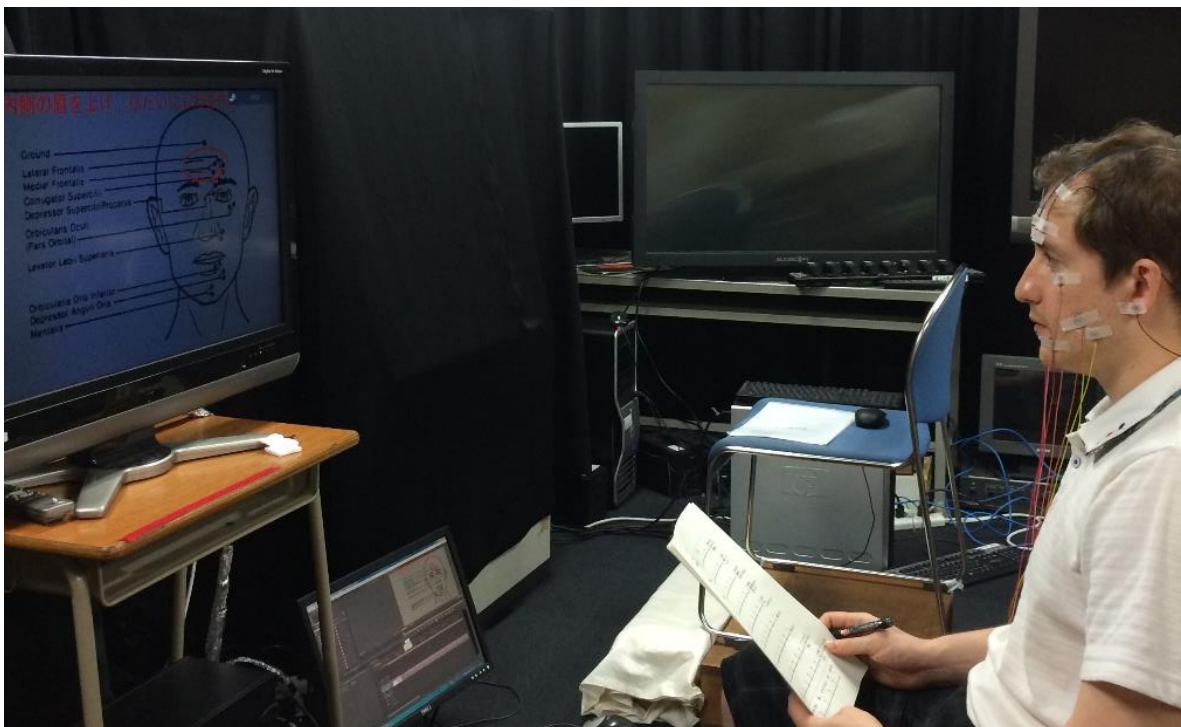


Subjects were asked to flex their facial muscles as much as possible for 10 seconds, while mimicking the following emotions:

- Happy
- Sad
- Angry
- Scared
- Surprised
- Osculation

The cycle for showing and recording these emotions was repeated 3 times, in order to establish an accurate baseline. We measured fEMG during MVC under observation. We set the data as a template standard for the analysis.

The photograph presented in Figure 13 was taken during one of the experiments at phase one and shows a test subject watching the instructions on the screen. During this part, subjects did not vote on the questionnaires.



*Figure 13. Experimental Session*

### 5.5.2 Phase 2 – Image Quality Assessment by Questionnaires

During the second part, room lights were completely shut off, in order to allow maximum focus on the display images and reduce distractions. Following a premade timed presentations, subjects were shown the 12 test images (6 separate images). An example of one test image used in the experiment is shown in Figures 14 (a,b), at two different quality levels respectively.

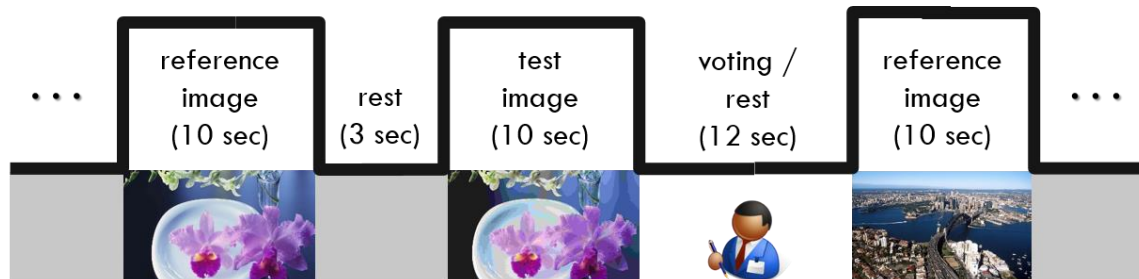


*Figure 14(a). Test image example (Quality : 5 – Quality degradation clearly visible)*



*Figure 14(b). Test image example (Quality : 30 – Quality degradation hardly visible)*

For the part of subjective assessment experiment, subjects assessed still images by the Double-Stimulus Impairment Scale (DSIS, five-grade) method, described in Chapter 2. We showed each subject one reference image for 10 seconds, followed by 3 seconds of gray display, followed by one test image for 10 seconds, followed by a 12 second for voting and rest. This cycle was repeated for all 12 images (see Figure 15).



*Figure 15. Testing Sequence – subjective image quality assessment*

Subjects noted their answers on answer sheets during voting time, based on the gradient scale detailed in Table 5 (Chapter 5.4).

## Chapter 6

# RESULTS AND ANALYSIS

For each subject in the experiment, we have obtained 2 sets of data – the results of the questionnaire voting and the electromyogram results. We set out to process both data sets individually and then compare them against each other. The result of the comparison will allow us to establish the relationship between the two separate methods.

### 6.1 IEMG

Integrated EMG (iEMG) is defined as the area under the curve of the rectified EMG signal, that is, the mathematical integral of the absolute value of the raw EMG signal [73]. When the absolute value of the signal is taken, noise will make the mathematical integral have a constant increase.

iEMG is the best approach to measure total muscular effort and it is important for qualitative EMG relationships such as EMG vs. Work, or EMG vs. Force, etc. There are three different ways to measure iEMG. The first way to measure iEMG is by mathematical integration. In this case, the area under absolute values of EMG time series is calculated. The second way to measure iEMG is by Root-Mean-Square method, in which time duration is similar but does not need to take the absolute value, and the third way to measure iEMG is electronically. Electronical iEMG can be measured in three different ways which are the simple time integration, integration & reset after a particular value is researched, and integration & reset after a fixed time interval [74].

In our example, IEMG has been determined from the integral value of full-wave rectification over a certain time range (as shown in Figure 16). In order to compare waveform values, we converted waveform values to quantifiable values. IEMG represents the amount of total muscle activity for a certain period (between the red lines in Figure 16).

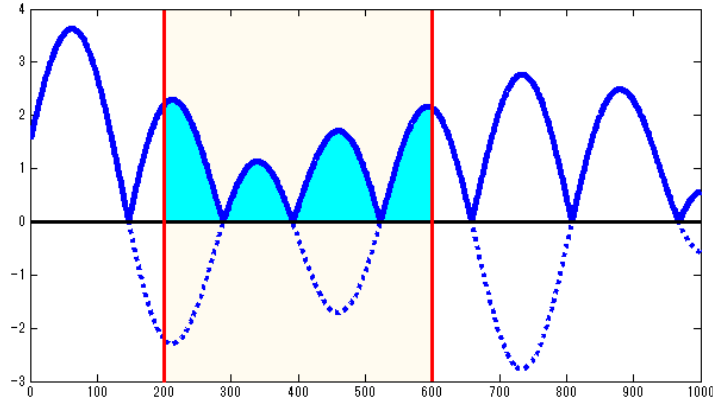


Figure 16. Integrated Electromyogram

Integration is a mathematical process carried out electronically, which calculates the area under a curve. The EMG signals are full-wave rectified before the integration process. The integrated electromyogram (iEMG) represents the total muscle activity and is a function of amplitude, duration, and frequency. The iEMG curve will keep rising, therefore it is reset to zero at a pre-set magnitude or intervals [75]. Signals obtained via surface electrodes are prone to influences such as the thickness of skin layer, cross talk from other body signals, change in electrode position and electrode size. In order to reduce the influence of these physical parameters during each trial of signal collection, signals were normalized with maximum volunteer contraction (MVC) as a reference point [76].

By comparing the amount of muscle activity, all measured waves were cleared of AC noise by band-stop filtering [77]. Next, we calculated IEMG using full-wave rectification for 10 seconds during MVC. For normalization of MVC for all muscle points, we defined  $M_{IEMG}$  as 100% of IEMG.  $M_{IEMG}$  has the maximum value of IEMG from 3 trials of MVC measurement testing for all muscle points.

In the subjective assessment experiment, we obtained the fEMG for each testing session. Then we calculated  $I_{EMG}$  for all image assessment results. We also calculated  $P_{IEMG}$ , which is the percentage of IEMG for all image assessment results, by comparing  $I_{EMG}$  with  $M_{IEMG}$ , and we identify the amount of muscle activity from Equation 1:

$$P_{IEMG} = \frac{I_{EMG}}{M_{IEMG}} \times 100 \quad (1)$$

---

## 6.2 Comparison of muscle activity

We compared the values of the amount of muscle activity and summarized the results in Tables 6 to 13 for each subject. M1 to M6 represent each muscle group (as described in Chapter 5.3.1) and the greater values for each subject are highlighted.

We defined the case where the score value was less or equal to 2 as for subjects that were looking at low-quality images. Special attention was paid to these cases when low-quality were shown.

The results demonstrate the following facts: subjects A1, A2, A6, A7 flexed more the M1 and M2 muscles located around the eyes. Subjects A3, A5, A8 flexed more the M3 and M4 muscles located around the mouth and M3 when looking at low-quality images. From this results we can conclude a connection between the results provided by the fEMG results and the subjective assessment score.

Table 6. Results of subject A1

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	15	13	8	13	6	3
30	5	15	7	8	15	8	3
30	5	24	52	11	16	8	3
30	5	18	10	18	15	13	3
30	4	21	19	11	15	6	3
30	4	18	8	11	16	9	3
5	2	16	6	9	17	10	3
5	2	24	51	12	15	11	2
5	2	22	38	13	15	11	3
5	1	21	19	9	15	7	3
5	1	25	32	45	17	9	3
5	1	26	69	16	18	13	3

Strong activity in the frontalis and corrugator muscles

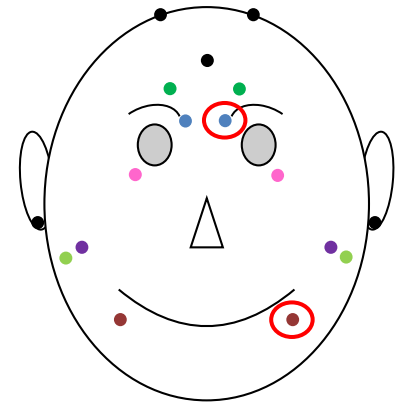


Table 6 shows the results for subject A1. We found a trend where muscles M1 (venter frontalis) and M2 (corrugator supercillii) experienced greater activity at low-quality image observations.

Table 7. Results of subject A2

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	12	28	16	25	26	21
30	4	11	31	19	32	29	24
30	4	10	25	15	24	30	24
30	4	11	28	16	28	29	26
30	4	12	30	19	25	31	26
30	3	12	32	17	24	29	23
5	2	15	38	20	25	26	25
5	2	13	41	19	25	32	23
5	1	12	33	17	27	30	24
5	1	16	45	17	23	30	23
5	1	17	47	21	26	32	26
5	1	13	42	21	27	43	25

Strong muscle activity in the corrugator, orbicularis oris

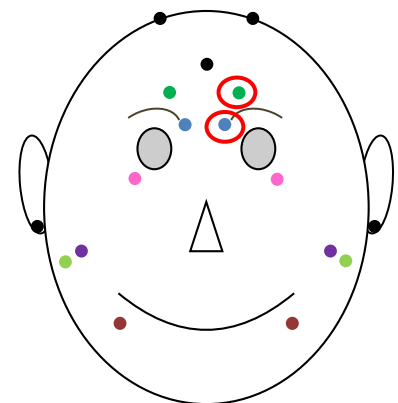
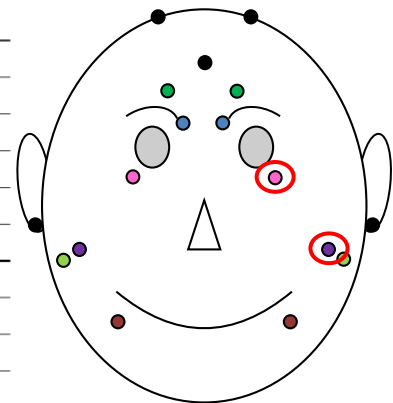


Table 7 shows the results for subject A2. We found a trend where muscles M2 (corrugator supercillii) and M5 (orbicularis oris) showed greater activity at low-quality image observations.

Table 8. Results of subject A3

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	5	10	27	14	20	11
30	5	5	12	32	16	21	9
30	4	5	7	21	13	18	10
30	4	4	10	22	19	23	9
30	3	5	7	22	16	20	10
30	3	4	10	20	20	23	9
5	2	5	19	47	25	24	9
5	1	5	13	55	29	20	11
5	1	5	15	43	24	20	10
5	1	5	15	28	16	21	9
5	1	5	10	44	28	21	10
5	1	5	8	46	31	24	10

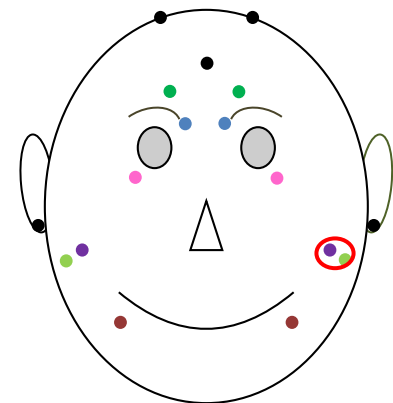


Strong activity in the orbicularis oculi and zygomatic major muscles

Table 8 shows the results for subject A3. We found a trend where muscles M3 (orbicularis oculi) and M4 (zygomaticus major) showed greater activity at low-quality image observations.

Table 9. Results of subject A4

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	10	10	7	18	9	24
30	5	10	8	7	17	8	22
30	5	9	9	8	19	9	23
30	4	10	9	7	17	8	22
30	4	11	13	9	18	8	23
30	4	9	9	9	18	18	21
5	4	10	11	9	18	8	26
5	3	12	10	8	17	8	22
5	3	13	27	9	19	8	22
5	2	12	11	8	18	9	23
5	2	11	9	8	19	13	24
5	2	11	9	7	18	8	22



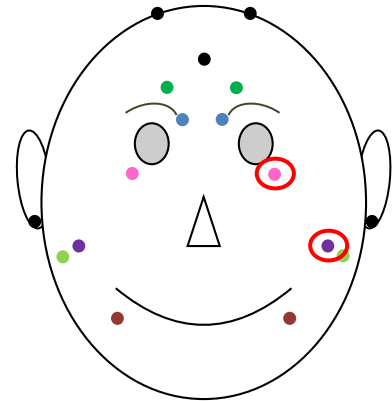
Strong activity in the masseter muscle

Table 9 shows the results for subject A4. We found greater activity in muscle M6 (masseter) at low-quality image observation.



Table 10. Results of subject A5

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	7	10	14	12	9	24
30	5	6	11	11	14	8	22
30	5	7	8	17	13	9	23
30	4	7	10	15	10	8	22
30	4	8	8	16	19	8	23
30	3	6	9	20	19	18	21
5	3	7	10	23	20	8	26
5	2	8	17	49	37	8	22
5	2	8	12	39	35	8	22
5	1	7	13	32	31	9	23
5	1	7	10	56	22	13	24
5	1	6	9	36	28	8	22

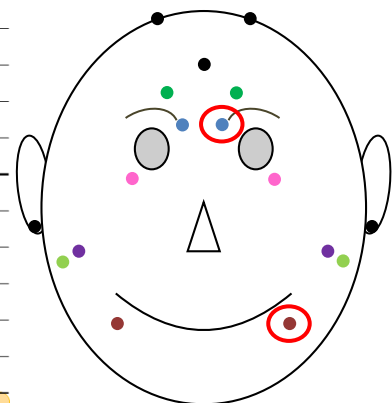


Strong activity in the orbicularis oculi and zygomatic major muscles

Table 10 shows the results for subject A5. We found a trend where muscles M3 (orbicularis oculi) and M4 (zygomaticus major) showed greater activity at low-quality image observations.

Table 11. Results of subject A6

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	4	13	17	11	3	3
30	4	4	14	18	10	6	4
30	4	12	23	23	9	7	4
30	4	8	10	16	13	5	3
30	4	35	41	22	18	10	3
30	4	5	16	13	15	5	3
5	3	37	34	19	20	8	3
5	2	41	37	20	21	7	3
5	2	19	30	17	17	7	3
5	2	42	41	21	19	4	3
5	2	14	12	20	15	5	2
5	1	31	19	13	17	9	3



Strong activity in the frontalis and corrugator muscles

Table 11 shows the results for subject A6. We found a trend where muscles M1 (venter frontalis) and M2 (corrugator supercillii) showed greater activity at low-quality image observations.

Table 12. Results of subject A7

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	9	15	10	20	18	13
30	5	8	13	11	21	18	13
30	4	9	19	14	20	23	15
30	4	8	15	12	23	17	14
30	4	9	33	20	19	20	13
30	4	7	20	14	18	18	14
5	3	10	40	20	24	36	19
5	2	13	43	20	24	38	20
5	2	10	30	14	18	23	18
5	2	13	40	18	22	36	24
5	1	11	32	17	19	32	17
5	1	10	38	15	18	39	20

Strong muscle activity in the corrugator, orbicularis oris

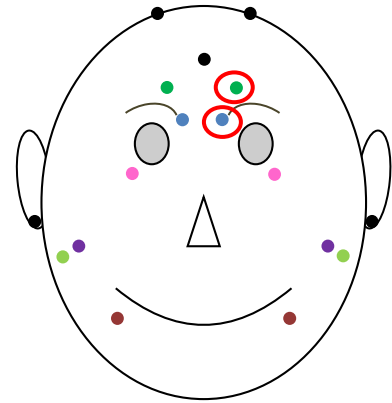


Table 12 shows the results for subject A7. We found a trend where muscles M2 (corrugator supercillii) and M5 (orbicularis oris) showed greater activity at low-quality image observations.

Table 13. Results of subject A8

QS	Evaluation score	Frontalis	Corrugator	Orbic. Oculi	Zygomatic	Orbic. Oris	Masseter
30	5	7	13	13	16	14	8
30	5	7	11	18	21	11	9
30	5	6	14	16	24	14	7
30	4	7	9	12	19	10	9
30	4	7	14	19	29	17	9
30	3	8	10	17	20	15	9
5	3	6	11	25	31	16	9
5	2	5	10	21	21	16	8
5	2	7	14	24	35	17	7
5	1	6	13	33	36	34	7
5	1	6	12	41	42	29	9
5	1	7	10	39	38	23	9

Strong activity in the orbicularis oculi and zygomatic major muscles

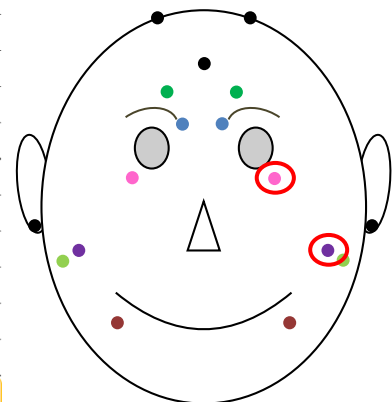


Table 13 shows the results for subject A8. We found a trend where muscles M3 (orbicularis oculi) and M4 (zygomaticus major) showed greater activity at low-quality image observations.

### 6.3 Regression Analysis

Dating back to approximately two hundred years ago, regression analysis is probably one of the oldest topics in the area of mathematical statistics. The first form of linear regression was the least square method, which was published by Legendre (1805), and Gauss (1809). Legendre and Gauss both applied the method to the problem of determining the orbits of bodies around the sun through astronomical observation. The same problem was approached by Euler (1738) but with no success. In 1821, Gauss published a further development of the least squares theory, including today's version of Gauss-Markov theorem, which is a fundamental theorem in the area of general linear models.

A statistical model is a simple description of a state or process. Levins [78] stated that "A model is neither a hypothesis nor a theory. Unlike scientific hypothesis, a model is not verifiable directly by an experiment. For all models of true or false, the validation of a model is not that it is "true" but that it generates good testable hypotheses relevant to important problems."

Linear regression demands that model is linear in regression parameters. Regression analysis is the method to find out the relationship between one or more response variables (also known as dependent variables, explained variables, predicted variables, or regressands, usually represented as  $y$ ) and the predictors (also known as independent variables, explanatory variables, control variables, or regressors, usually represented as  $x_1, x_2, \dots, x_p$ ).

There are three types of regression, namely the simple linear regression, the multiple linear regression and the nonlinear regression. The simple linear regression is used for modeling the linear relationship between two variables, the dependent variable  $y$  and the independent variable  $x$ . The simple regression model is calculated as shown in Equation 2:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (2)$$

where  $y$  is the dependent variable,  $\beta_0$  is  $y$  intercept,  $\beta_1$  is the gradient or the slope of regression line,  $x$  is the independent variable, and  $\varepsilon$  is the random error. It is mostly assumed that error  $\varepsilon$  is normally distributed with  $E(\varepsilon) = 0$  and a constant variance  $Var(\varepsilon) = \sigma^2$  in the simple linear regression. The multiple linear regression model is a linear regression model with one dependent variable and more than one independent variables. This regression assumes that the response variable is a linear function of the model parameters and there are more than one independent variables in the model. The common form of the multiple linear regression model is as in the following Equation 3:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p + \varepsilon \quad (3)$$

Where  $y$  is the dependent variable,  $\beta_0, \beta_1, \dots, \beta_p$  are regression coefficients, and  $x_1, \dots, x_n$  are independent variables in the model. In the classical regression setting, the error term  $\varepsilon$  is usually assumed that it follows the normal distribution with  $E(\varepsilon) = 0$  and a constant variance

$Var(\varepsilon) = \sigma^2$ . The multiple linear regression involves more matters than simple linear regression, for example collinearity, variance inflation, graphical display of regression diagnosis, and detection of regression outlier and influential observation. The third type of regression is the nonlinear regression. This type of regression assumes that the relationship between dependent variable and independent variables is not linear in regression parameters. Nonlinear regression model (growth model) could be written as in the following Equation 4:

$$x = \frac{\alpha}{1+e^{\beta t}} + \varepsilon \quad (4)$$

in which  $y$  is the growth of a particular organism as a function of time  $t$ ,  $\alpha$  and  $\beta$  are model parameters, and  $\varepsilon$  is the random error. Nonlinear regression model is more complicated than linear regression in the sense of estimation of model parameters, model selection, model diagnosis, variable selection, outlier detection, or influential observation identification.

Regression analysis can be applied in many scientific fields such as medicine, biology, agriculture, economics, engineering, sociology, geology, and many more [125]. Basically, the purpose of regression analysis are:

- Establish a causal relationship between response variable  $y$  and regression  $x_1, x_2, \dots, x_n$ .
- Predict  $y$  based on a set values of  $x_1, x_2, \dots, x_n$ .
- Screen variables  $x_1, x_2, \dots, x_n$  to identify which variables are more important than others to explain the response variable  $y$  so that the causal relationship can be determined more efficiently and accurately.

### 6.3.1 Stepwise Regression Analysis

Stepwise regression is a suitable procedure for selecting variables into a model, especially when large numbers of variables are involved, but it also has its setback. Stepwise regression renders hypothesis testing, for example F and t tests. Hypothesis testing is a statistical procedure for accepting or rejecting the null hypothesis on the basis of estimates on a particularized model. Stepwise regression carry out modeling by analyzing large number of variables, then choose those that fit well. Hence, the t-values for the selected variables will probably be important, and hypothesis testing loses its inference capacity. Stepwise regression is not recommended if the aim of modeling is for testing the validity of a relationship between certain variables or to test the meaning of a particular variable. However if the aim or objective is forecasting, it is a convenient way for selecting variables, particularly when large numbers of variables are to be considered.

To be able to use the stepwise procedure, the simple (pair-wise) correlation coefficient and partial correlation coefficient between  $Y$  and each  $X$  variables under consideration need to be calculated.

Simple correlation or pair-wise correlation is the correlation between two variables without other variables influence. The simple correlation between variable  $x$  and variable  $y$  is simply the ratio between their covariance and the product of their respective standard deviations, which is as in the following Equation 5:

$$r_{yx} = \frac{\sum yx}{\sqrt{\sum y^2}\sqrt{\sum x^2}} \quad (5)$$

To simplify the formula,  $t$  subscripts were dropped from the variables and lower case  $y$  and  $x$  were used, which denote the deviations of  $Y$  and  $X$  from their respective means, in the formula. So basically:

$$\begin{aligned} y &= (Y - \bar{Y}) \\ x &= (X - \bar{X}) \end{aligned} \quad (6)$$

Partial correlation coefficient is when the correlation between  $y$  and  $x$  is computed by first eliminating the effect of all other variables. For instance, if there are three variables,  $y$ ,  $x_1$ , and  $x_2$ , and to compute the partial correlation coefficient between  $y$  and  $x_1$ , it would be computed after the impact of  $x_2$  on  $y$  and  $x_1$  is removed. The computed formula is shown in Equation 7:

$$r_{yx_1..x_2} = \frac{r_{yx_1} - r_{yx_2}r_{yx_1x_2}}{\sqrt{(1-r_{yx_2}^2)(1-r_{x_1x_2}^2)}} \quad (7)$$

The computation of partial correlation coefficient becomes troublesome when more and more variables are involved. Nevertheless, thanks to the software for stepwise regression analysis, the task can be achieved quicker.

The stepwise regression procedure is a general method for selecting variables into a model compared to the backward elimination and the forward selection methods. The backward elimination method starts with the inclusion of all the identified variables in the model and then eliminates the insignificant ones from the model one by one. On the other hand, the forward selection procedure takes one variable at a time on the basis of their partial correlation coefficients, for example, it first takes the variable with the highest partial correlation coefficient to enter the model, then the one with the second highest partial correlation coefficient, and so on. After each variable is added, the partial  $F$  test is performed on the last entered variable. The process stops when the partial  $F$  test shows that the last entered variable is insignificant. The variables entered prior to that last variable will remain in the model. The stepwise method is an improvement over the backward and forward selection procedures, and a step-by-step procedure of it is stated below.

---

Step 1 – calculate the simple (pair-wise) correlation between Y and each of the X variables, X1, X2, X3, ..., in accordance to Equation (5), and then the one most correlated with the dependent variable Y as the first variable to enter into the model. Let it be X1.

Step 2 – assuming that X1 is already in the model, calculate the partial correlation coefficients between Y and the remaining X variables in accordance with Equation (7). Select the one that has the highest partial correlation coefficient as the second variable to enter into the model. Let it be X2.

Step 3 – use the partial F of X1 to test the significance of the first variable entered into the model at a pre-determined confidence level. The partial F statistic of X1 is simply the  $t^2$  of the estimated coefficient for X1 after X2 was entered into the model. Variable X1 will be eliminated if the partial F test shows that the coefficient of X1 is insignificant. Assume that X1 is retained.

Step 4 – re-calculate the partial correlation coefficients between Y and the remaining X variables, assuming that X1 and X2 are already in the model. Select the variable with the highest partial correlation coefficient to enter into the model.

Step 5 – Repeat Step 3 and Step 4 to test all remaining variables including those variables eliminated at earlier stages until the best subset of independent variables is selected.

## 6.4 Experimental Application

The regression analysis [10] was conducted on an individual basis for each subject from the point of view of QoE. We do not need or attempt to average or deduct statistical models from our experiment, but rather identify specific relations between facial muscle activity and the sensation of image quality degradation.

In this part, we estimate the subjective assessment scores, as we will apply fEMG to the image quality assessment. We conducted a step-wise regression analysis and formulated the estimated regression equations. The independent variables were the amount of muscle activity of M1 to M6. The dependent variable represents the subjective score value.

The results of the step-wise regression analysis are detailed in Table 14. The estimated regression equations are shown in Equations 8 to 15.

$$\text{Prediction A1} = -0.25 \times M1 + 8.27 \quad (8)$$

$$\text{Prediction A2} = -0.17 \times M2 + 8.78 \quad (9)$$

$$\text{Prediction A3} = -0.20 \times M4 + 6.68 \quad (10)$$

$$\text{Prediction A4} = -0.64 \times M1 + 10.41 \quad (11)$$

$$\text{Prediction A5} = -0.09 \times M3 + 5.49 \quad (12)$$

$$\text{Prediction A6} = -0.20 \times M4 + 6.15 \quad (13)$$

$$\text{Prediction A7} = -0.13 \times M5 + 6.59 \quad (14)$$

$$\text{Prediction A8} = -0.14 \times M3 + 6.27 \quad (15)$$

Table 14. Results of step-wise regression analysis

<b>Subjects</b>	<b>Entry variable</b>	<b>Coefficient of correlation</b>
<b>A1</b>	M1	0.58
<b>A2</b>	M2	0.85
<b>A3</b>	M4	0.76
<b>A4</b>	M1	0.68
<b>A5</b>	M3	0.85
<b>A6</b>	M4	0.64
<b>A7</b>	M5	0.82
<b>A8</b>	M3	0.86

The graphs are shown in Figures 17 to 24. According to the results gathered in the experiment, we have generated scatter charts of the predicted score and the subjective assessment score. The horizontal axis represents the score, while the longitudinal axis represents the predicted score. As can be seen in the graphs, we found a nearly linear relationship between the predicted score and the subjective assessment score for all subjects. From these facts, we can conclude that the use of fEMG is feasible technology in estimating subjective assessment scores.

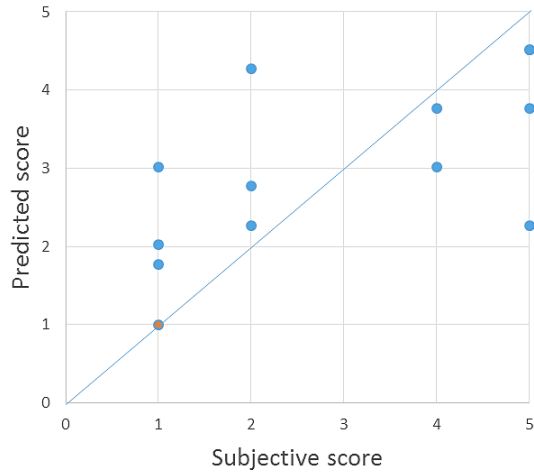


Figure 17. Relationship between predicted and subjective score for subject A1.

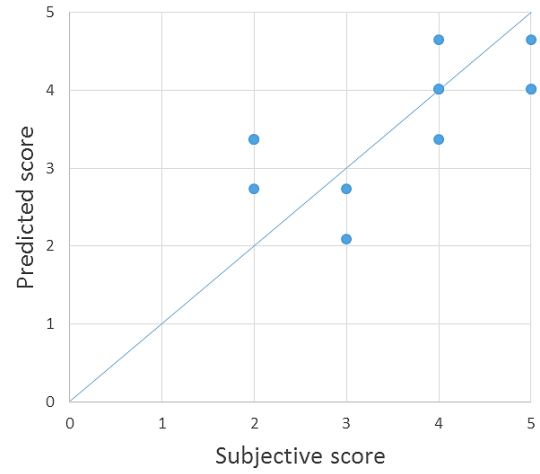


Figure 20. Relationship between predicted and subjective score for subject A4.

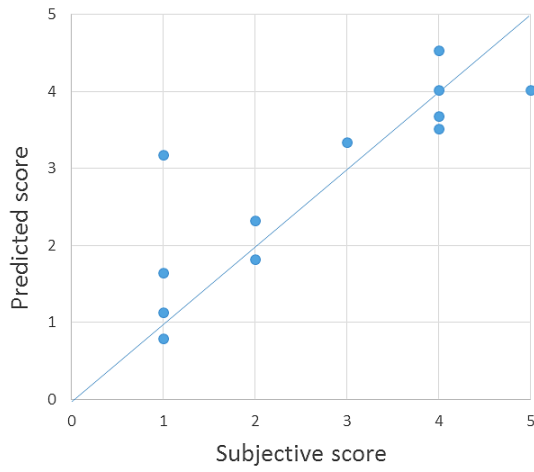


Figure 18. Relationship between predicted and subjective score for subject A2.

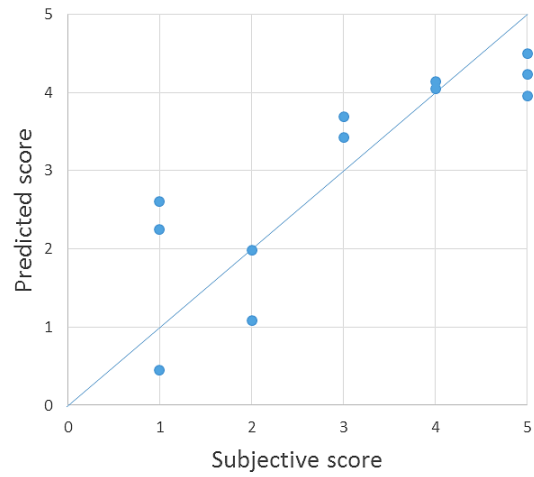


Figure 21. Relationship between predicted and subjective score for subject A5.

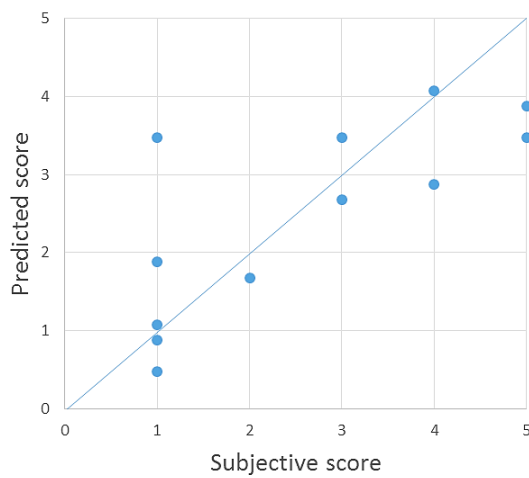


Figure 19. Relationship between predicted and subjective score for subject A3.

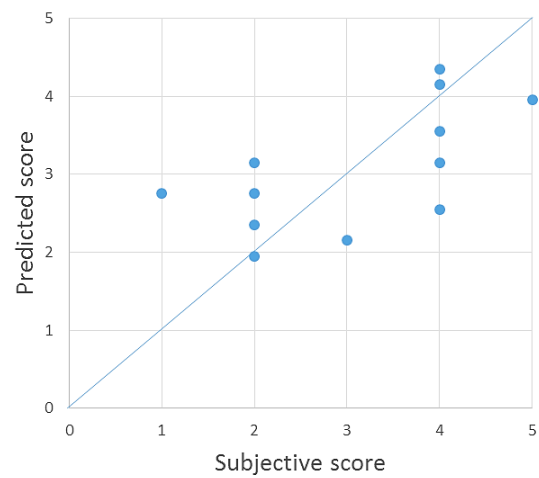


Figure 22. Relationship between predicted and subjective score for subject A6.



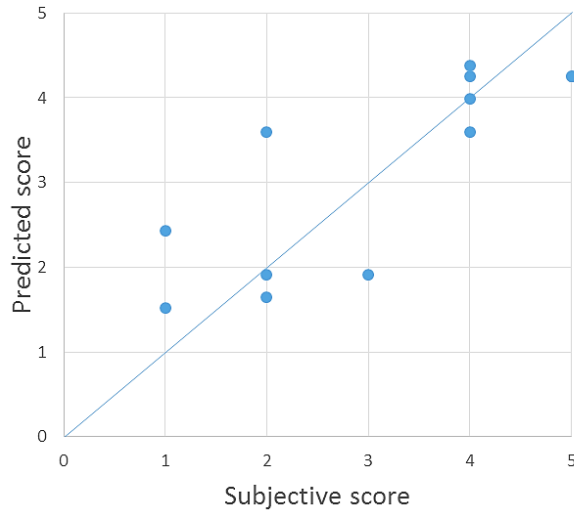


Figure 23. Relationship between predicted and subjective score for subject A7.

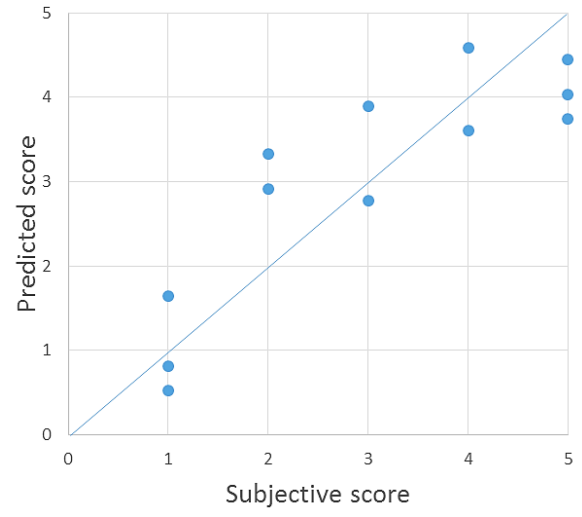


Figure 24. Relationship between predicted and subjective score for subject A8.

## Chapter 7

# PERFORMANCE IMPROVEMENT USING IMAGE INFORMATION

### 7.1 Improvement Goals

Now that we have established a definite connection between image contents and facial muscle activity, we examine further aspects that might improve the prediction algorithm. Specifically we look at image contents and how that affects the prediction results. We argue that a person is judging image quality by focusing on the degradation parts of coded images (as seen in the example of Figure 27). Therefore, for improving the prediction performance of previously proposed model, we must take into account the image features related to the coarse-fine information of spatial frequency characteristics [140]. The key to improving the model performance is finding the appropriate methodology for factoring in the image features related to muscle activity.

### 7.2 Image Entropy

Image entropy is a quantity used to describe the amount of information which must be coded by a compression algorithm. Low entropy images, i.e. containing plain color areas, have very little contrast and large number of pixels with the same or similar DN values [141]. An image that is completely flat will have therefore an entropy of zero and can be compressed to a relatively small

size. On the other hand, high entropy images having a lot of variation in content contain a lot of contrast between adjacent pixels, and as a result cannot be compressed as much as low entropy images. The formula for calculating image entropy is given in Equation 16:

$$\text{Entropy} = - \sum_i P_j \text{Log}_2 P_j \quad (16)$$

In the above expression,  $P_j$  is the probability of  $i$ -th gray lever for monochrome still image (8 bits/pixel) converted from color image (24 bits/pixel). Therefore, the maximum value of the entropy for gray level image is 8 bits/pixel [141].

One of the important features is the entropy of image, the value of which is strongly connected with image content. If an image contains a lot of detail information, its entropy value is very close to 8 bits/pixel (for a monochrome image). On the other hand, if the image contains mostly flat areas, then the entropy value is much lower than 8 bit/pixel. This feature is useful for expressing image contents.



Figure 25. Visible degradation in test images

A high entropy value represents the presence of maximum information, containing high-detail areas. An image with very low entropy value represents coded noise with tiling based on the block size and false contours (like images 9 and 11 shown in Figure 25) .

In Figures 26 and 27, we can see the test images used in the experiment. In this study, we have added the additional variables of Entropy and  $\Delta$ entropy. Entropy was calculated by converting images to grayscale, which allows us to quantify image contents, as explained above. For the purpose of obtaining  $\Delta$ entropy, we have also calculated Original Entropy, using the “perfect” images and deducted Entropy from Original Entropy, as detailed in Table 15.  $\Delta$ entropy represents the difference

between the original image and the coded image, and serves a good indicator for image content analysis in our calculations.



Image 1

QS:30

Entropy = 7.748

$\Delta$ entropy = 0.008



Image 2

QS:30

Entropy = 7.416

$\Delta$ entropy = 0.220



Image 3

QS:30

Entropy = 7.233

$\Delta$ entropy = 0.347



Image 4

QS:30

Entropy = 7.412

$\Delta$ entropy = 0.081



Image 5

QS:30

Entropy = 6.922

$\Delta$ entropy = 0.628



Image 6

QS:30

Entropy = 7.328

$\Delta$ entropy = 0.019

*Figure 26. Test images at QS 30.*



Image 7

QS:5

Entropy = 7.106

$\Delta$ entropy = 0.650



Image 8

QS:5

Entropy = 5.746

$\Delta$ entropy = 1.890



Image 9

QS:5

Entropy = 5.898

$\Delta$ entropy = 1.682



Image 10

QS:5

Entropy = 5.406

$\Delta$ entropy = 2.087



Image 11

QS:5

Entropy = 4.390

$\Delta$ entropy = 3.159



Image 12

QS:5

Entropy = 6.591

$\Delta$ entropy = 0.756

*Figure 27. Test images at QS 5.*

Entropy is a measure of graylevel distribution (disorder or randomness) in the histogram. Therefore we have converted all images in grayscale, before computing Entropy Information. In Figure 28 we have sorted the test images in descending order of entropy level. The image with the highest entropy level (7.106) indicates the highest level of details and contrast, while the image with the lowest entropy (4.390) indicates the presence of large areas of same color.

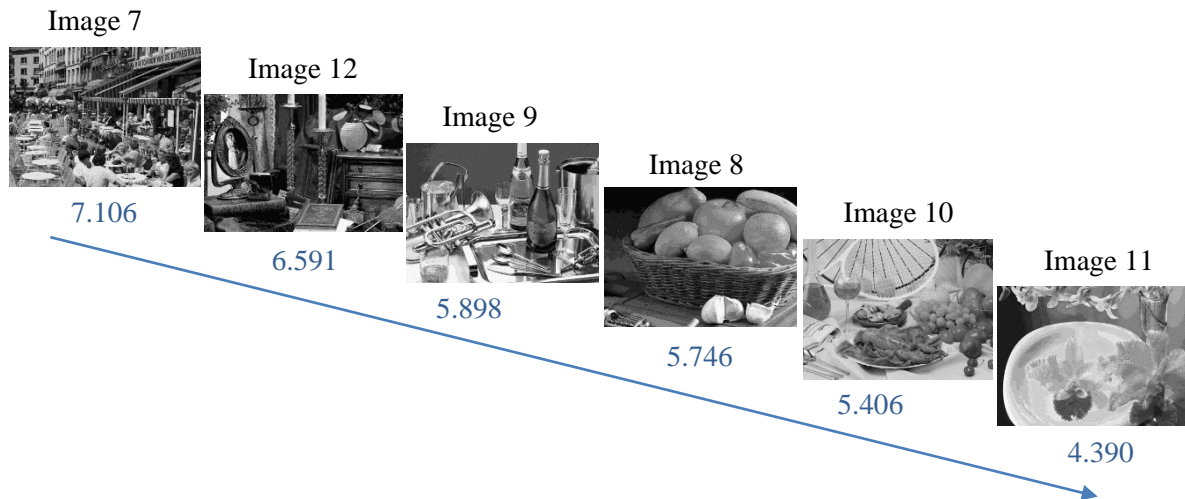


Figure 28. Image Entropy values of test images

From Table 15, it can be observed that  $\Delta$ entropy of images 8-11 is fairly high, compared with the others, because these images contain large flat areas like the background (see Fig. 27).

Table 15. Test Images Entropy Results

Image	Entropy	$\Delta$ entropy	Original Entropy
1 (QS:30)	7.748	0.008	7.756
2 (QS:30)	7.416	0.220	7.636
3 (QS:30)	7.233	0.347	7.580
4 (QS:30)	7.412	0.081	7.493
5 (QS:30)	6.922	0.628	7.550
6 (QS:30)	7.328	0.019	7.347
7 (QS:5)	7.106	0.650	7.756
8 (QS:5)	5.746	<b>1.890</b>	7.636
9 (QS:5)	5.898	<b>1.682</b>	7.580
10 (QS:5)	5.406	<b>2.087</b>	7.493
11 (QS:5)	4.390	<b>3.159</b>	7.550
12 (QS:5)	6.591	0.756	7.347

As a result, the degradations of these images (like blocking artifacts and contours [142]) are more visible than in other images. This explains why it is easy to judge the quality assessment by 5-grade scale without looking or seeking the degradation parts specifically.

### 7.3 Multiple Regression Analysis

In order to factor in the additional variables of Image Entropy, we have conducted a multiple regression analysis [3] using 3 sets of variables (muscle activity, entropy and  $\Delta$ entropy), in order to established the most suitable pattern for quality prediction. The results are indicated as Multiple R, Adjusted R (obtained as the square root of Adjusted R Squared) and Maximum (Prediction) Error.

As in the case of regression analysis, the subjective scores for each image served as the dependent variables.

The first analysis was conducted using only muscle activity data (M1-M6) as independent variables. The results are detailed in Table 16.

Table 16. Results of Multiple Regression using only Muscle Data

<b>Subject</b>	<b>Multiple R</b>	<b>Adjusted R</b>	<b>Max Error</b>
<b>A1</b>	0.579	0.518	2.827
<b>A2</b>	0.854	0.838	2.016
<b>A3</b>	0.740	0.709	2.538
<b>A4</b>	0.677	0.635	1.370
<b>A5</b>	0.847	0.831	1.575
<b>A6</b>	0.636	0.588	1.768
<b>A7</b>	0.817	0.796	1.546
<b>A8</b>	0.857	0.842	1.305

The second analysis was conducted using muscle data, plus Entropy, as independent variables. The results are detailed in Table 17.

Table 17. Results of Multiple Regression using Muscle Data and Entropy

<b>Subject</b>	<b>Multiple R</b>	<b>Adjusted R</b>	<b>Max Error</b>
<b>A1</b>	0.813	0.765	1.886
<b>A2</b>	0.870	0.839	1.606
<b>A3</b>	0.881	0.853	1.190
<b>A4</b>	0.851	0.814	1.516
<b>A5</b>	0.856	0.820	1.451
<b>A6</b>	0.818	0.771	1.906
<b>A7</b>	0.923	0.905	1.095
<b>A8</b>	0.892	0.867	1.030

The third analysis was conducted using muscle data, plus Entropy, plus  $\Delta$ entropy as independent variables. The results are detailed in Table 18.

Table 18. Results of Multiple Regression using Muscle Data, Entropy and  $\Delta$ entropy

<b>Subject</b>	<b>Multiple R</b>	<b>Adjusted R</b>	<b>Max Error</b>
<b>A1</b>	0.816	0.735	2.016
<b>A2</b>	0.885	0.838	1.622
<b>A3</b>	0.893	0.849	1.055
<b>A4</b>	0.930	0.902	0.768
<b>A5</b>	0.911	0.875	1.217
<b>A6</b>	0.850	0.786	1.435
<b>A7</b>	0.953	0.934	0.925
<b>A8</b>	0.918	0.885	0.985



## 7.4 Results and Comparison

From the results of the regression analysis in Table 16-18, we can conclude that the best performance is the use of all 3 variables (muscle activity, entropy and  $\Delta$ entropy). For this case, the estimated regression equations are represented in Equations 17-24:

$$\text{Prediction A1} = -0.01 \times M1 + 2.40 \times \text{Ent} + 1.08 \times \Delta\text{ent} - 13.54 \quad (17)$$

$$\text{Prediction A2} = -0.12 \times M2 + 2.31 \times \text{Ent} + 1.86 \times \Delta\text{ent} - 10.27 \quad (18)$$

$$\text{Prediction A3} = -0.10 \times M4 + 2.50 \times \text{Ent} + 1.68 \times \Delta\text{ent} - 13.39 \quad (19)$$

$$\text{Prediction A4} = -0.26 \times M1 + 4.08 \times \text{Ent} + 3.40 \times \Delta\text{ent} - 23.82 \quad (20)$$

$$\text{Prediction A5} = -0.06 \times M3 + 4.23 \times \text{Ent} + 3.83 \times \Delta\text{ent} - 26.90 \quad (21)$$

$$\text{Prediction A6} = -0.11 \times M4 + 2.89 \times \text{Ent} + 2.21 \times \Delta\text{ent} - 16.42 \quad (22)$$

$$\text{Prediction A7} = -0.08 \times M5 + 3.33 \times \text{Ent} + 2.62 \times \Delta\text{ent} - 19.24 \quad (23)$$

$$\text{Prediction A8} = -0.07 \times M3 + 3.48 \times \text{Ent} + 2.77 \times \Delta\text{ent} - 20.94 \quad (24)$$

Table 19. Results comparison (Muscle data only versus Muscle, Entropy and  $\Delta$ entropy data)

Subject	Muscle Information only			Muscle Data, Entropy and $\Delta$ entropy		
	Multiple R	Adjusted R	Max Error	Multiple R	Adjusted R	Max Error
A1	0.579	0.518	2.827	0.816	0.735	2.016
A2	0.854	0.838	2.016	0.885	0.838	1.622
A3	0.740	0.709	2.538	0.893	<b>0.849</b>	<b>1.055</b>
A4	0.677	0.635	1.370	0.930	<b>0.902</b>	<b>0.768</b>
A5	0.847	0.831	1.575	0.911	0.875	1.217
A6	0.636	0.588	1.768	0.850	0.786	1.435
A7	0.817	0.796	1.546	0.953	<b>0.934</b>	<b>0.925</b>
A8	0.857	0.842	1.305	0.918	<b>0.885</b>	<b>0.985</b>

From the results comparison detailed in Table 19 we can see that by introducing entropy and  $\Delta$ Entropy in the prediction model, the regression coefficient R, the adjusted regression coefficient R, and the maximum prediction error have all returned improved results. Multiple and Adjusted R values have increased by an average of 16%, while Maximum Error was decreased by an average of 54%, which is also an important indicator of improved prediction. Therefore, we can deduce that Entropy and  $\Delta$ Entropy are effective indicators for predicting the overall image quality for image quality assessment methodology.

The scatter charts of predicted model are presented in Figures 29 to 36, showing a side-by-side visualization of the two scenarios (muscle information only versus Muscle, entropy and  $\Delta$ Entropy information). The (a) charts of the figures are based on the relationship between the predicted scores and the subjective scores for the case of using only muscle information. The (b) charts of the figures are based on the relationship between the predicted scores and the subjective scores for the case of using muscle information, entropy and  $\Delta$ Entropy. As in the case of stepwise regression analysis, we found a near linear relationship, further improved by the addition of entropy and  $\Delta$ Entropy. As we can see from the charts, accuracy is significantly improved for all cases.

The results of the Multiple Regression Analysis using the additional variables of Entropy and  $\Delta$ Entropy have proven that image contents is a crucial factor in establishing a robust image quality prediction model. As can be seen by looking at the test images, some contain large portions of obvious image degradation (i.e. image 8, 12). However, compared with the human eye, it is very difficult to express these indications in quantifiable measurements, unless we use content-dependent variables like Entropy. The closer an image's Entropy level is to 8 bits/pixel, the more image contains a higher level of details, which makes subjective assessment higher. The lower an image's Entropy level is to 8 bits/pixel, the more image contains a lower level of details, which makes degradation less perceptible to the human eye. These are very important factors, which have helped improve the accuracy of the prediction model.

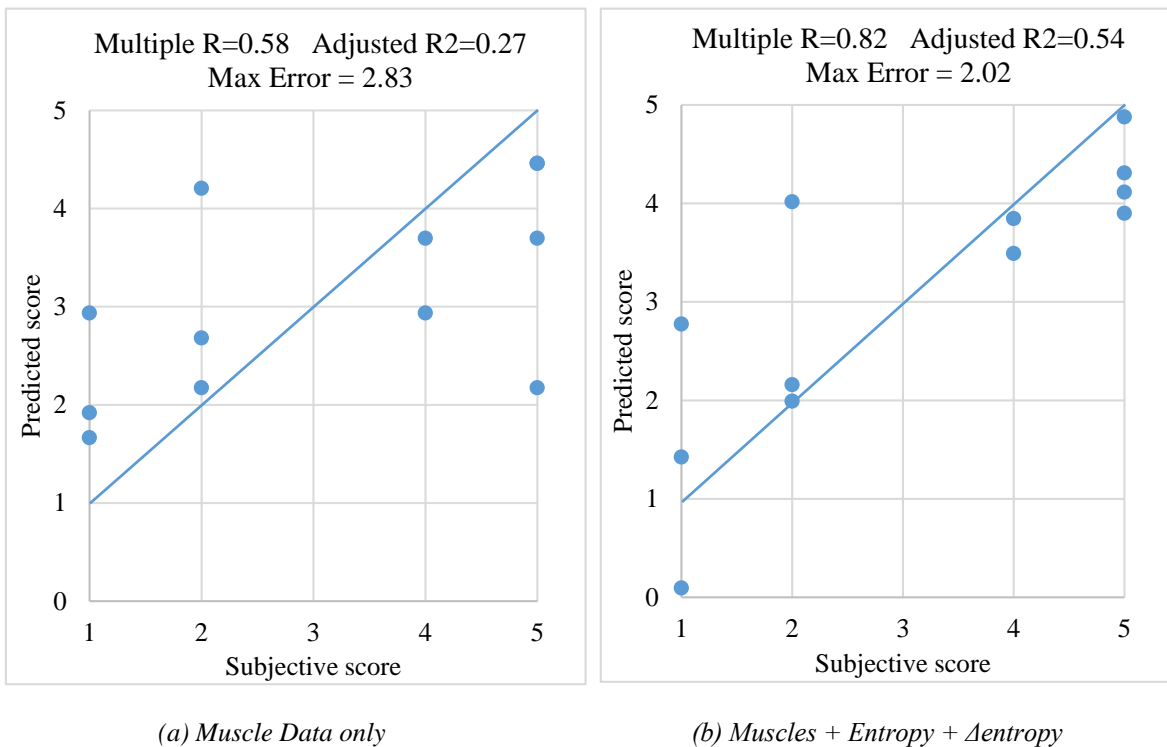


Figure 29. Comparison of relationship between predicted and subjective score for subject A1

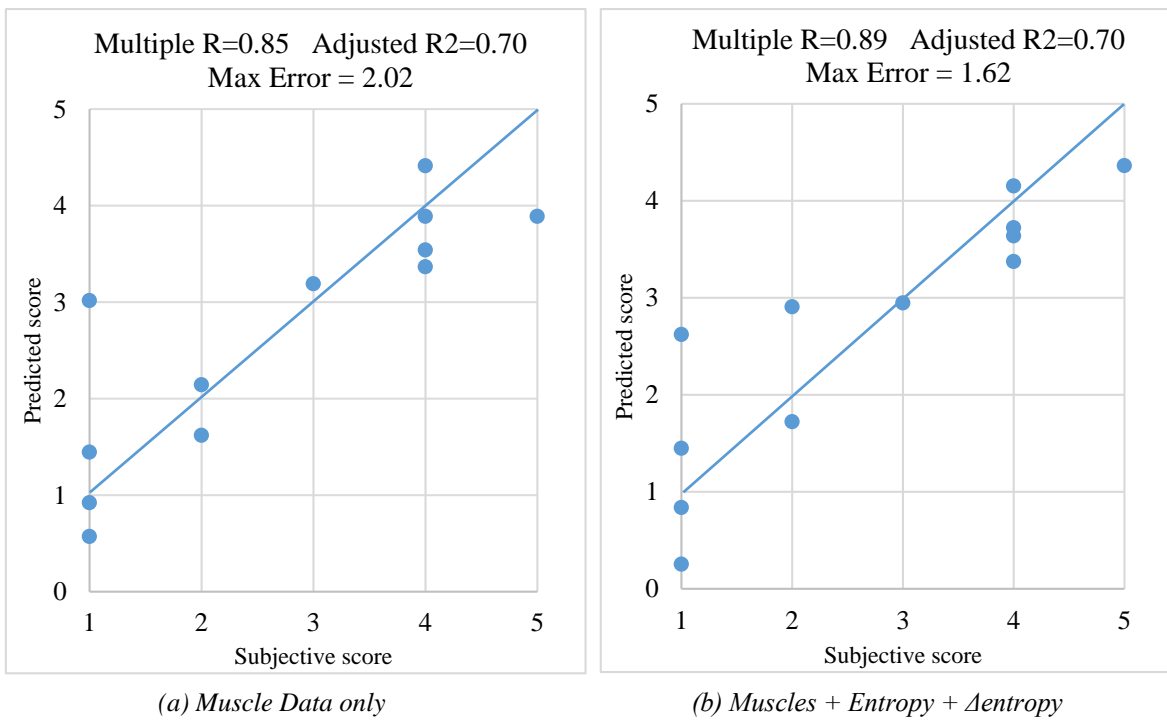


Figure 30. Comparison of relationship between predicted and subjective score for subject A2

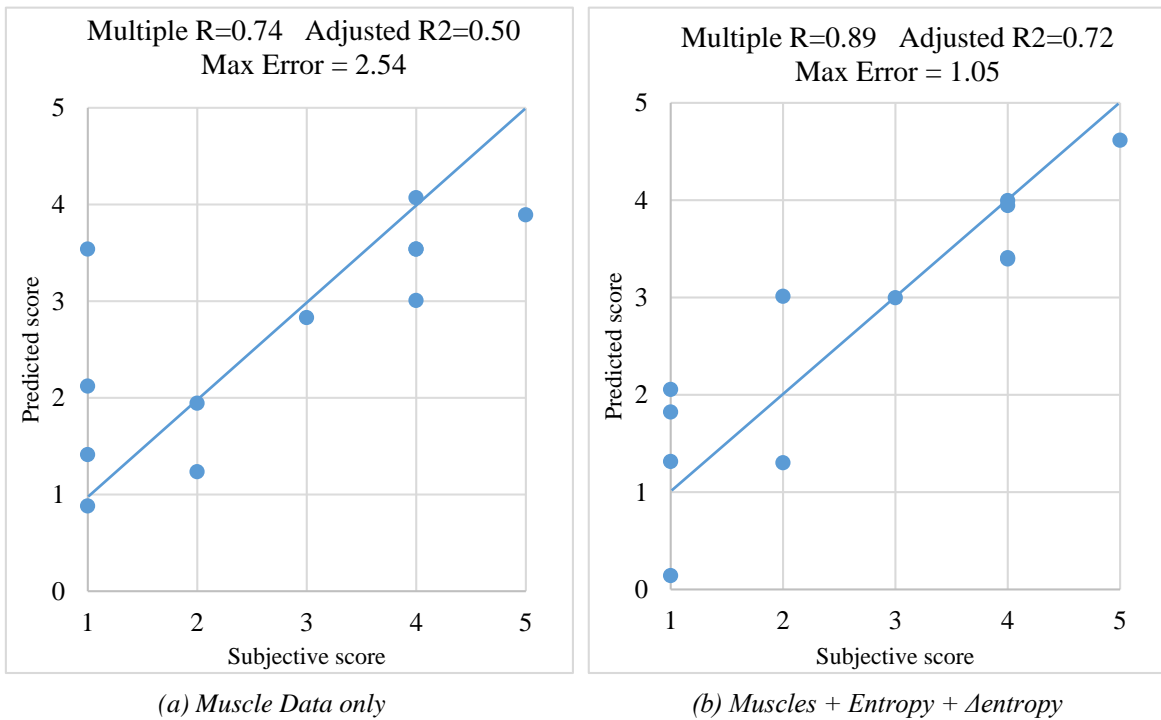


Figure 31. Comparison of relationship between predicted and subjective score for subject A3

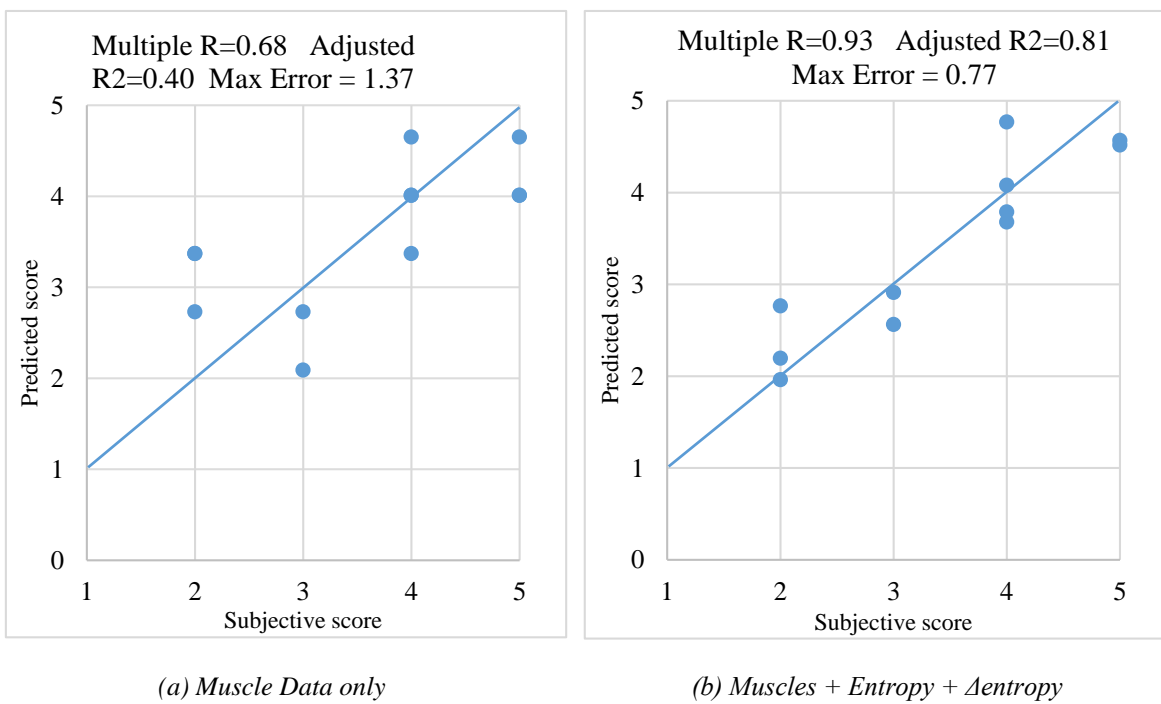
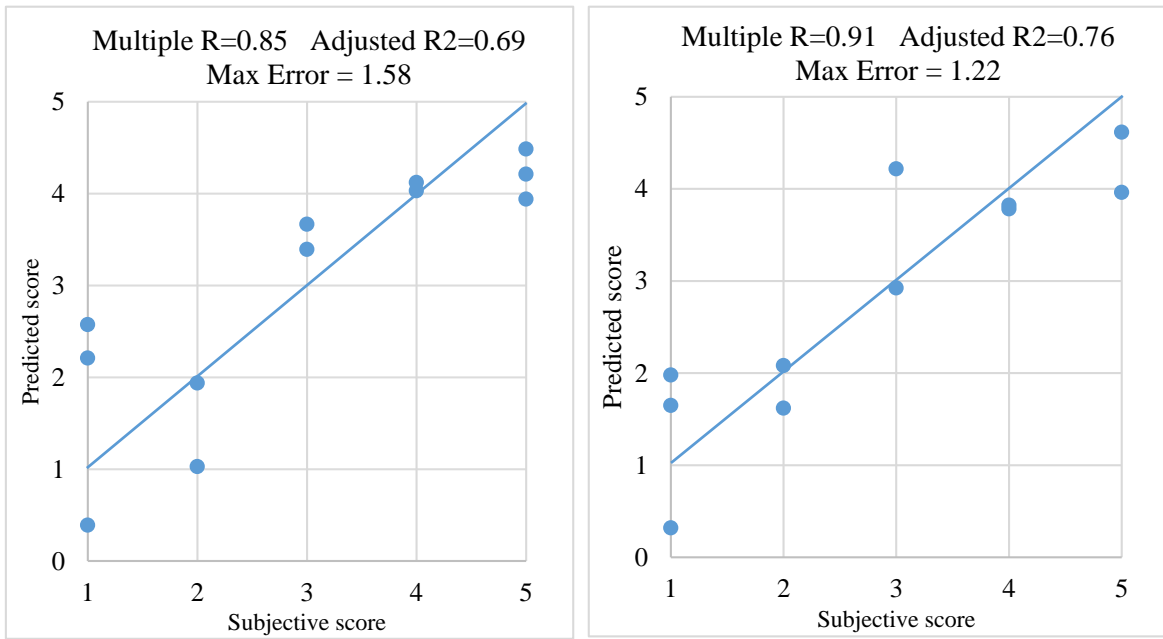


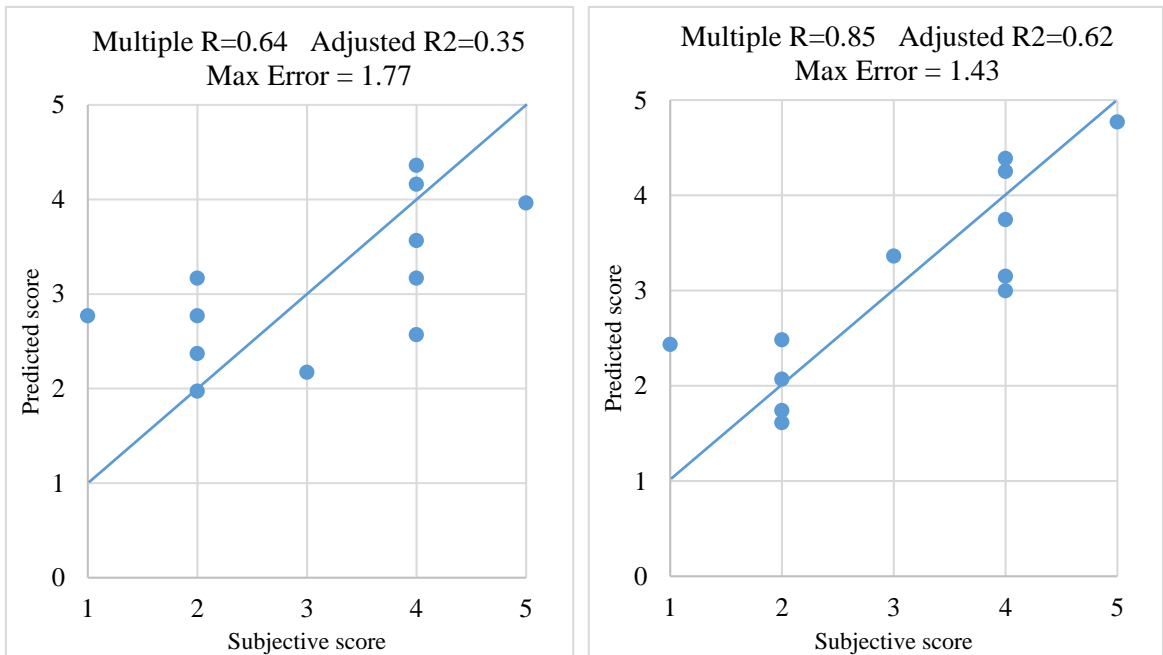
Figure 32. Comparison of relationship between predicted and subjective score for subject A4



(a) Muscle Data only

(b) Muscles + Entropy + Δentropy

Figure 33. Comparison of relationship between predicted and subjective score for subject A5



(a) Muscle Data only

(b) Muscles + Entropy + Δentropy

Figure 34. Comparison of relationship between predicted and subjective score for subject A6

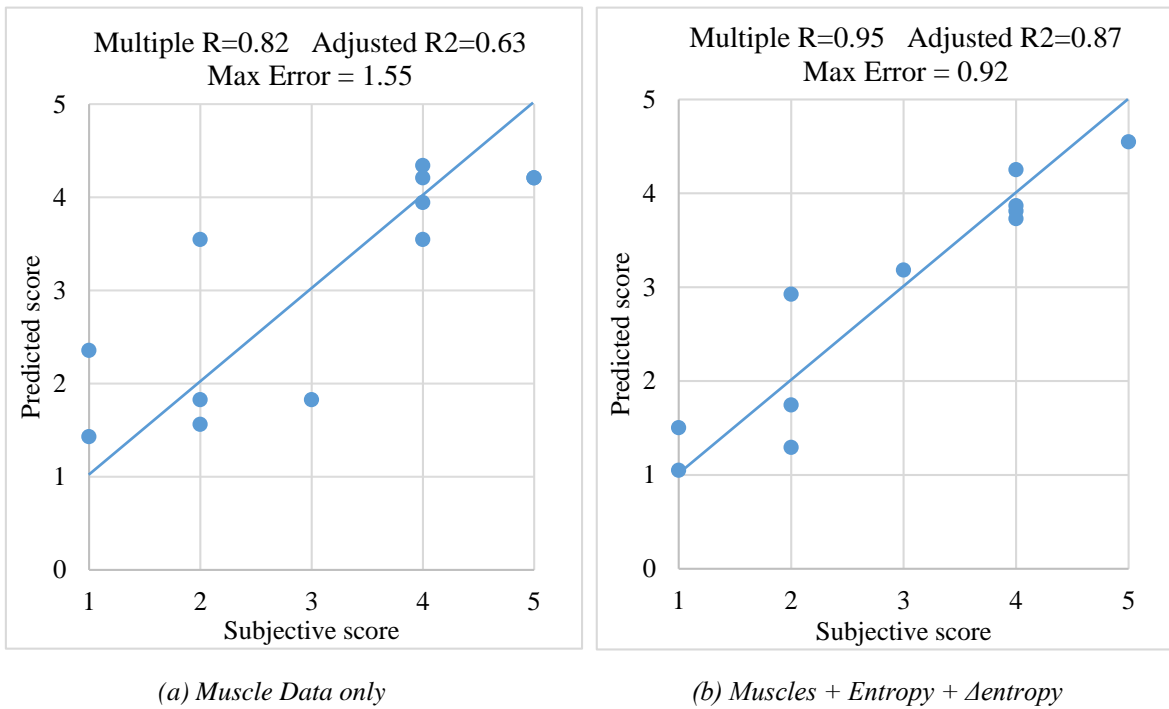


Figure 35. Comparison of relationship between predicted and subjective score for subject A7

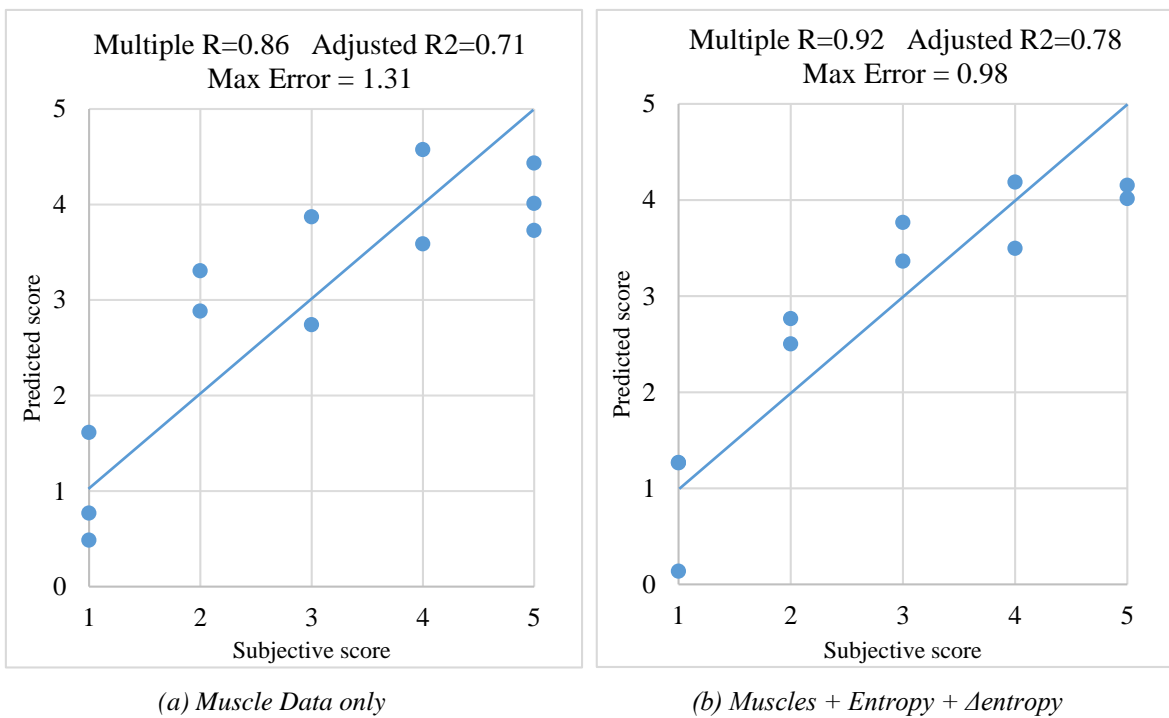


Figure 36. Comparison of relationship between predicted and subjective score for subject A8

## Chapter 8

# CONCLUSIONS

It is undeniable that image quality assessment is a very important research topics nowadays, taking into account the unprecedented engagement of people with technology, especially in visual aspects, like image and video information. There is a multitude of algorithms and methodologies in the image-processing field, with the goal of sustaining and enhancing the quality of media to be presented to users. Until now, the only subjects that are able to evaluate the quality of images have been the people themselves and this is has proven the most reliable method so far. Unfortunately, this method is too costly and too slow to be applied on the large scale expected.

If we think about it, future research will undoubtedly evolve along the lines of offering other methods than human assessment of the image quality and predict perceived quality of the image automatically. In other words, a breakthrough research on objective image quality assessment is needed in the field in order to facilitate progress to a new level, a level envisioned by many researchers and IT companies. We see this progress happening every day around us, with newer and better technology emerging every day, taking image quality towards the envisioned level.

Previous research have provided significant improvements in objective assessment methodologies. Nevertheless, the field of image quality assessment has not yet seen “the big leap”. If we were to take a look at the number of studies conducted on “image quality assessment” when compared to “image

---

enhancement”, the number is discouragingly limited and insufficient [79]. Is the field of image quality assessment such an intimidating, insoluble problem? Is it not possible to produce technology that may come close to a human’s ability to perceive image quality by utilizing today’s software and computing technology? The biggest obstacle to overcome at this point is to understand the human perception of quality in and use this knowledge to develop better technologies for objective image quality assessment. Another approach is to obtain biometric information from the human observer himself during the process of image quality assessment, examine and use this information to build truly efficient objective image quality assessment models.

Through the research described in this thesis, it is the first time that a conclusive relationship between biometric information, specifically the activity of facial muscles of human beings, and image quality has been established. This approach improves on the traditional questionnaire-based method, which suffers from major flaws caused by the difference in response results and the difference in rating scales between subjects. It has been argued that image quality assessment using biological information is less likely to suffer from these issues. Our research has been focused on Facial Electromyography (fEMG) for biological information extraction. Facial expressions contain more emotional essence compared to the tone of voice or the content of spoken message, therefore making it suitable for our goal. We used this premise to hypothesize a noteworthy relationship between facial muscle activity and image quality assessment and ultimately improving Quality of Experience. The long term objective of such an approach is to develop and implement evaluation methods for automatically assessing and improving Quality of Experience (QoE). In order to accomplish this goal, we set out to set the premises that will not only identify the relationship between the activity of facial muscles and perceived image degradation, but also quantify this relationship and provide a robust base for further development.

During our experiments, we measured electrical responses of facial muscles in parallel with a subjective assessment test using still images. We studied the relationship between fEMG and the subjective assessment score. As a result, we found a trend where subjects’ muscles around the eyes and mouth areas show greater activity when subjects were assessing low-quality images. We also developed estimated regression equations for each subject using the step-wise regression analysis, as well as the possibility of estimating the subjective assessment score. From experimental results we can conclude a connection between the results provided by the fEMG results and the subjective assessment score. This demonstrates that fEMG, which uses part of biological information, can be applied in the evaluation of image quality assessment under certain conditions.

The results of the Multiple Regression Analysis using the additional variables of Entropy and  $\Delta$ entropy have proven that image contents is also a crucial factor in establishing a robust image quality prediction model. Some images contain obvious (for the human eye) image degradations. However, compared with the human eye, it is very difficult to express these indications in



---

quantifiable measurements, unless we use content-dependent variables like Entropy. The closer an image's Entropy level is to 8 bits/pixel, the more image contains a higher level of details, which makes subjective assessment higher. The lower an image's Entropy level is to 8 bits/pixel, the more image contains a lower level of details, which makes degradation less perceptible to the human eye. These are very important factors, which have helped improve the accuracy of the prediction model.

With this first step of introducing the usage of facial muscle activity, we can further develop our work towards non-intrusive objective methods for assessing video quality and improving image quality assessment technology using biological information.

# References:

- [1]. Kayvan Najarian, Robert Splinter, "Biomedical Signal and Image Processing, Second Edition", CRC Press, May 4, 2012. P.271-272
- [2]. Jacques Blanc-Talon, Wilfried Philips et al, "Advanced Concepts for Intelligent Vision Systems: 11th International Conference, ACIVS 2009 Bordeaux, France, September 28--October 2, 2009 Proceedings", Springer Science & Business Media, Sep 15, 2009. PP.156-166
- [3]. H. Julia Hannay, "Experimental Techniques in Human Neuropsychology", Oxford University Press, 1988, p.457-478
- [4]. Aimé Lay-Ekuakille, "Wearable and Autonomous Biomedical Devices and Systems for Smart Environment", Issues and Characterization, 2010, p.354
- [5]. Marco Barbero, Roberto Merletti, Alberto Rainoldi, "Atlas of Muscle Innervation Zones: Understanding Surface Electromyography and Its Applications", Springer, 2012
- [6]. Gary Kamen, David A. Gabriel, "Essentials of Electromyography", Human Kinetics", 2010
- [7]. Himabindu Yamana, "Embedded System for Monitoring Human Activities Using 3-axis Accelerometer", ProQuest, 2007
- [8]. Zhou Wang, Alan Conrad Bovik, "Modern Image Quality Assessment", Morgan & Claypool Publishers, Jan 1, 2006
- [9]. David Kleinbaum, Lawrence Kupper et al, "Applied Regression Analysis and Other Multivariable Methods", Cengage Learning, Aug 30, 2013
- [10]. George A. F. Seber, Alan J. Lee, "Linear Regression Analysis", John Wiley & Sons, Jan 20, 2012. pp 391-427
- [11]. H.R. Wu, K.R. Rao, "Digital Video Image Quality and Perceptual Coding", CRC Press, Nov 18, 2005. PP.559-567a
- [12]. Zhou Wang, Alan Conrad Bovik, "Modern Image Quality Assessment", Morgan and Claypool Pulishers, January 2006
- [13]. Syed Mahbubur Rahman, "Multimedia Networking: Technology, Management and Applications", Idea Group Inc (IGI), 2002
- [14]. Marco Alfano, Rolf Sigle, "Controlling QoS in a Collaborative Multimedia Environment", HPDC, Proceedings of 5th IEEE International Symposium on High Performance Distributed Computing, pp. 340, 1996, doi:10.1109/HPDC.1996.546204
- [15]. 3GPP TSG 44.060. "Radio Link Control/Medium Access Control (RLC/MAC) protocol".
- [16]. Gerardo Gomez, Rafael Sanchez et al, "End-to-End Quality of Service over Cellular Networks: Data Services Performance Optimization in 2G/3G", John Wiley & Sons, pp 117, October 2005
- [17]. Abe Iro , Yutaka Ishibashi , Hideaki Yoshino , " Trends of the Next Generation on Quality of Service Technology" , Journal of Electronics, Information and Communication Society , vol. 91, no. 2, pp. 82-86,2008.
- [18]. Alimuzzaman MD., Takamichi Miyata et al, "Proposal of the receiving side prediction retransmission control scheme for real-time media QoE improvement" , IEICE technical report IN, information network , vol. 108, no. 289, pp. 1-6, 2008.
- [19]. The Study Groups of ITU's Telecommunication Standardization Sector (ITU-T) ([www.itu.int](http://www.itu.int))
- [20]. Tatsuya Yamazaki , " the future of service quality in the next generation network era" , Institute of Electronics, Information and Communication Engineers Journal , vol. 91, no. 2, pp. 113-116, 2008.
- [21]. ITU-T Recommendation : G. 1080, "Quality of experience requirements for IPTV services" , 2008.
- [22]. Tanaka, Saori et al, "Extraction of quality elements suitable for QoE measurement in IPTV", IEICE General Conference Proceedings 2010 Sciences , pp. 41 -42, 2010.
- [23]. Eguchi Masato , Takumi Miyoshi et al, "Construction of the user quality of experience (QoE) model that takes into account the selection good for the delay condition and content", IEICE technical report CQ, vol. 108, no. 137, pp. 43-48,2008.
- [24]. Jingming Li Salina, Pascal Salina, "Next Generation Networks: Perspectives and Potentials", John Wiley & Sons, February 2008, pp 33
- [25]. Torgerson W. S.: Theory and methods of scaling. Wiley, 1985.
- [26]. ITU-R.REC.BT.500-11: Methodology for the subjective assessment of the quality for television pictures, 2002.
- [27]. Keelan B. W.: ISO 20462: a psychophysical image quality measurement standard. In Proceedings of SPIE (2003), vol. 5294, SPIE, pp. 181-189.
- [28]. ITU-T.REC.P.910: Subjective audiovisual quality assessment methods for multimedia applications, 2008.
- [29]. Rafał K. Mantiuk, Anna Tomaszewska, Radosław Mantiuk, "Comparison of four subjective methods for image quality assessment", The Eurographics Association and Blackwell Publishing Ltd, 2012, PP 2-4
- [30]. Engeldrum P.: Psychometric scaling: a toolkit for imaging systems development. Imcotek Press, 2000.
- [31]. Cunningham D., Wallraven C.: Experimental Design: From User Studies to Psychophysics, vol. 2011. Taylor & Francis, 2011.

- 
- [32]. Ferwerda J. A.: Psychophysics 101: how to run perception experiments in computer graphics. In ACM SIGGRAPH 2008 classes (2008), ACM, pp. 1–60.
- [33]. Sundstedt V., Whitton M., Bloj M.: The whys, how tos and pitfalls of user studies. In ACM SIGGRAPH 2009 Courses (2009), no. 25, ACM. 2
- [34]. Yoshida A., Mantiuk R., Myszkowski K., Seidel H.P., “Analysis of Reproducing Real-World Appearance on Displays of Varying Dynamic Range”, *Computer Graphics Forum* 25, 3 (2006), 415–426.
- [35]. Mantiuk R., Tomaszewska A., Heidrich W.: Color correction for tone mapping. *Computer Graphics Forum* 28, 2 (2009), 193–202.
- [36]. Ledda P., Chalmers A., Troscianko T., Seetzen H.: Evaluation of tone mapping operators using a high dynamic range display. *ACM Trans. Graph.* 24, 3 (2005), 640–648.
- [37]. C Adik M., Wimmer M., Neumann L., Artusi A.: Evaluation of HDR tone mapping methods using essential perceptual attributes. *Computers & Graphics* 32, 3 (2008), 330–349.
- [38]. Rubinstein M., Gutierrez D., Sorkine O., Shamir A.: A comparative study of image retargeting. *ACM Trans. on Graph.* 29, 6 (2010), 160.
- [39]. Van Dijk A. M., Martens J. B., Watson A. B., “Quality assessment of coded images using numerical category scaling”. *Proc. SPIE* 2451 (1995), 90–101.
- [40]. Tominaga T., Hayashi T., Okamoto J., Takahashi A.: Performance comparisons of subjective quality assessment methods for mobile video. In 2nd Int. Workshop on Quality of Multimedia Experience (QoMEX) (2010), pp. 82–87.
- [41]. Takanori Hayashi, Akira Takahashi et al, "Trends related to QoE Evaluation Technology of Multimedia Communication Services and Challenges", *Institute of Electronics, Information and Communication Engineers Journal* , vol. J91-A, no. 6, pp. 600-612 , 2008
- [42]. Takashi Hasegawa, "Psychological Evaluation Technology ", *ITE Journal* , vol. 54, no. 9, pp. 1259-1264,2000.
- [43]. ITU-R Recommendation : BT. 500-13, “Methodology for the subjective assessment of the quality of television pictures”, 2012.
- [44]. Stefan Winkler, “Digital Video Quality: Vision Models and Metrics”, John Wiley & Sons, May 28, 2013
- [45]. Yvonne Waern, Sharon McDonald et al, “People and Computer XIV – Usability or Else!: Proceedings of Hci 2000”, Springer Science & Business Media, Aug 16, 2000. PP. 327-329
- [46]. Kai-Chieh Yang, “Perceptual Quality Assessment for Compressed Video”, ProQuest, 2007, PP. 121-122
- [47]. Zhou Wang, Alan Conrad Bovik, "Modern Image Quality Assessment", Morgan and Claypool Publishers, January 2006
- [48]. Wang Z., Bovik A. C.: Modern Image Quality Assessment. Morgan & Claypool Publishers, 2006.
- [49]. VQEG Final report from the video quality experts group on the validation of objective models of video quality assessment (March 2004). <http://www.vqeg.org/> Ryszard S. Choras, "Image Processing and Communications Challenges 5", Springer Science & Business Media, Jul 19, 2013
- [50]. Chikkerur S., Cartwright A. et al, “Fingerprint Image Enhancement Using STFT Analysis, International Workshop on Pattern Recognition for Crime Prevention, Security and Surveillance, Lecture Notes in Computer Science”, Springer Berlin/Heidelberg, Volume 3687/2005, pp 20-29
- [51]. Daugman J., “Statistical richness of visual phase information: update on recognizing persons by iris patterns. In: International Journal on Computer Vision 45:25-38”, 2001
- [52]. Basri R., Jacobs D., “Lambertian reflectance and linear subspaces. NEC Research Institute Technical Report”, 2000, 172R
- [53]. Zhang L., Samarasinghe D., “Face recognition under variable lighting using harmonic image exemplars”, Proceedings of International Conference on Computer Vision on Pattern Recognition, 2003, pp 1:19-25
- [54]. Shan S., Gao W. et al, “Illumination normalization for robust face recognition against varying lighting conditions”, Proceedings of International Workshop on Analysis and Modeling of Faces and Gestures, 2003, pp 157-164.
- [55]. Shashua A., Riklin-Raviv T., “The quotient image: class-based re-rendering and recognition with varying illuminations”, *Transaction on Pattern Analysis and Machine Intelligence*, 2001, 23(2):129-139
- [56]. Fronthaler H., Kollreider K. et al, “Automatic image quality assessment with application in biometrics”, *Computer Vision and Pattern Recognition Workshop*, 2006, pp 30
- [57]. Levine M. D., Bhattacharya J., “Detecting and removing specularities in facial images”, *Computer Vision and Image Understanding* 100, 2005, pp 330-356
- [58]. Gross R., Brajovic V., “An image pre-processing algorithm for illumination invariant face recognition”, Proceedings of Audio- and Video-Based Biometric Person Authentication 10-18, 2003
- [59]. Kela N., Rattani A. et al, “Illumination invariant elastic bunch graph matching for efficient face recognition.”, Proceedings of International Conference on Computer Vision and Pattern Recognition Workshop 42-42, 2006
- [60]. Kandel E.R., Schwartz, J.H., Jessell, T.M. 2000. Principles of Neural Science, 4th edition., McGraw-Hill, New York. p.1305
- [61]. Manami Ueda, Keiji Shibata et al, "Relevance of the image content and image quality degradation and  $\delta$ oxy-Hb", *IEICE technical report IMQ*, vol.112, no. 161, pp. 15 -20, 2012
- [62]. Covidien Systems <http://www.covidien.com/rms/brands/invos>
- [63]. Zhang G., Salganicoff M., “Method of measuring the focus of close-up image of eyes”, United States Patent Number 5953440, 1999
-

- [64]. Hongo M, Itoi E, Sinaki M, Miyakoshi N, et al., "Effect of low-intensity back exercise on quality of life and back extensor strength in patients with osteoporosis: a randomized controlled trial", 2007
- [65]. Hiroki Katayama, Keiji Shibata et al, "Impact on the considerations depending on the type of sound source of multimedia content", IEICE technical report IMQ, vol. 111, no. 192, pp. 27-30, 2011
- [66]. Sazzad Hussain, Mamun Bin Ibne Reaz, "Electromyography Signal Analysis and Characterization: Based on Wavelet Transform and Higher Order Statistics", LAP Lambert Academic Publishing, 2012
- [67]. Loeb, G. E., & Gans, C., "Electromyography for experimentalists", University of Chicago Press, 1986
- [68]. Alan J. Fridlund and John T. Cacioppo, "Guidelines for human electromyographic research," *Psychophysiology*, vol. 23(5), pp. 567–589, September 1986.
- [69]. Neurofax  $\mu$  Electroencephalograph EEG-9100J/K, [http://www.nihonkohden.com/products\\_en/brochure/pdf/CAT\\_EEG9100.pdf](http://www.nihonkohden.com/products_en/brochure/pdf/CAT_EEG9100.pdf)
- [70]. J. Clerk Maxwell, "A Treatise on Electricity and Magnetism", 3rd ed., vol. 2. Oxford: Clarendon, 1992, pp.68-73.
- [71]. Dictionary of Sport and Exercise Science and Medicine by Churchill Livingstone © 2008 Elsevier Limited.
- [72]. BMES 642: EMG Analysis, [http://www.biomed.drexel.edu/labs/biomechanics/emg\\_analy.htm](http://www.biomed.drexel.edu/labs/biomechanics/emg_analy.htm)
- [73]. Rudiger Kramme, Klaus-Peter Hoffmann et al, "Springer Handbook of Medical Technology", Springer Science & Business Media, October 2011, p.9
- [74]. Dionysios-Dimitrios Koutsouris, Athina A. Lazakidou, "Concepts and Trends in Healthcare Information Systems", Springer, 2014, PP.222
- [75]. Shrawan KUMar, A. Mital, "Electromyography in Ergonomics", CRC Press, Mar 20, 1996. PP.31
- [76]. Reaz, M.B.I, Hussain, M.S., "A non-MVC EMG normalization technique for the trunk musculature: Part 1 method development", *Journal of Electromyography and Kinesiology* 11(1), 2001, 1-9
- [77]. André Galagowicz, Wilfried Philips, "Computer Vision/Computer Graphics Collaboration Techniques: 5th International Conference, MIRAGE 2011, Rocquencourt, France, October 10-11, 2011. Proceedings". PP.274-284
- [78]. R. Levins, *Am. Scientist* 54, 1966, PP 421
- [79]. Sugimoto Osamu, "The video quality element and its evaluation in terms of broadcasting and communication", *Journal of Electronics, Information and Communication Society*, vol. 96, no. 4, pp. 242-248, 2013.
- [80]. [b73] Lindsley DB., "Electrical activity of human motor units during voluntary contraction", *Am J Physiol*, 1935; pp.90-99
- [81]. De-Shuang Huang, Yong Gan et al, "Bio-Inspired Computing and Applications : 7th International Conference on Intelligent Computing, ICIC2011, Zhengzhou, China, August 11-14. 2011, Revised Papers", Springer Science & Business Media, January 2012 pp 368
- [82]. Daubechies I., "Ten lectures on wavelets", Society for Industrial and Applied Mathematics, 1992
- [83]. Parvez Sazzad, Z., Kawayoke, Y., Horita, Y., "Image quality evaluation database", 2000, <http://mict.eng.u-toyama.ac.jp/mictdb.html>
- [84]. Le Callet, P., Atrousseau, F., "Subjective quality assessment IRCCyN/IVC database", 2005, <http://www.irccyn.ec.nantes.fr/ivcdb/>
- [85]. Sheikh, H., Wang, Z., Cormack, L., Bovik, A., "LIVE image quality assessment database release 2, 2005, <http://live.ece.utexas.edu/research/quality>
- [86]. Louis G. Tassinary and John T. Cacioppo, "Unobservable facial actions and emotion," *Psychological Science*, vol.3, no. 1, pp. 28–33, January 1992.
- [87]. Guoping Qiu, Kin Man Lam, "Advances in Multimedia Information Processing -- PCM 2010, Part I: 11th Pacific Rim Conference on Multimedia, Shanghai, China, September 21-24, 2010, Proceedings", Springer Science & Business Media, Sep 3, 2010
- [88]. Tarek Sobh, Khaled Elleithy, "Innovations in Computing Sciences and Software Engineering", Springer Netherlands, Aug 25, 2010
- [89]. Roberto Merletti, Philip A. Parker, "Electromyography: Physiology, Engineering, and Non-Invasive Applications", John Wiley & Sons, Jul 26, 2004
- [90]. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
- [91]. Shanbao Tong, Nitish Vyomesh Thakor, "Quantitative EEG Analysis Methods and Clinical Applications", Artech House, 2009
- [92]. Stefan Winkler, "Digital Video Quality: Vision Models and Metrics", John Wiley & Sons, May 28, 2013
- [93]. Issues in Diagnostics and Imaging: 2013 Edition, Scholarly Editions, May 1, 2013. P.479-495
- [94]. Thierry Dutoit, Ferran Marques, "Applied Signal Processing: A MATLAB™-Based Proof of Concept", Springer Science & Business Media, Jun 10, 2010
- [95]. George C. S. Wang, Chaman L. Jain, "Regression Analysis: Modeling & Forecasting", Institute of Business Forec, 2003. Page 57-70
- [96]. André Galagowicz, Wilfried Philips, "Computer Analysis of Images and Patterns: 11th International Conference, CAIP 2005, Versailles, France, September 5-8, 2005

- 
- [97]. Ajith Abraham, Jaime Lloret Mauri et al, "Advances in Computing and Communications, Part II", Springer Science & Business Media, Jul 8, 2011. pp.
- [98]. Ashish Ghosh, Rajat K. De, Sankar K Pal, "Pattern Recognition and Machine Intelligence: Second International Conference, PRMI 2007, Kolkata, India, December 18-22, 2007. PP.87-93
- [99]. Ramón Pallás-Areny, John G. Webster, "Analog Signal Processing", John Wiley & Sons, 1999. PP.234-245
- [100]. André Gagalowics, Wilfried Philips, "Computer Vision/Computer Graphics Collaboration Techniques: 5th International Conference, MIRAGE 2011, Rocquencourt, France, October 10-11, 2011. Proceedings", Springer Science & Business Media, Sep 23, 2011. PP. 276
- [101]. PEDERSEN M., HARDEBERG JON Y.: Full-Reference Image Quality Metrics: Classification and Evaluation. *Foundations and Trends in Computer Graphics and Vision* 7, 1 (2011), 1–80.
- [102]. [b25] SHEIKH H., SABIR M., BOVIK A.: A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing* 15, 11 (2006), 3441–3452.
- [103]. [b26] PONOMARENKO N., LUKIN V., ZELENSKY A., EGIAZARIAN K., CARLI M., BATTISTI F.: TID2008 – A database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics* 10 (2009), 30–45.
- [104]. Gary Kamen, David Gabriel, "Essentials of Electromyography", *Human Kinetics*, Sep 28, 2009. PP. 2-3
- [105]. Merletti, R., D. Farina, M. Gazzoni, A. Merlo, P. Ossola and A. Rainoldi, "Surface electromyography: A window on the muscle, a glimpse on the central nervous system", *Europa Medicophysica* 37, 2001, PP. 57-68
- [106]. Wolf, W., "The EMG as a window to the brain: signal processing tools to enhance the view", in I. Gath, and G. Inbar, eds., *Advances in processing and pattern analysis of biological signals*, Plenum Press, New York, 1996, PP. 339-356
- [107]. Matthews, B. H. C., "A special purpose amplifier", *J Physiol (Lond)* 81, 1934, PP. 28-29
- [108]. Stalberg E, J. V. Trontelj, *Single fibre electromyography*, Mirvalle Press, UK, 1979
- [109]. McGill, K. C., K. L. Cummins, and L. J. Dorfman, "Automatic decomposition of the clinical electromyogram", *IEEE Trans BME* 32 (7), 1985, PP. 470-477
- [110]. Roberto Merletti, Philip A. Parker, "Electromyography: Physiology, Engineering, and Non-Invasive Applications", John Wiley & Sons, Jul 26, 2004, PP. xv-xvi
- [111]. Lindsley DB., "Myographic and Electromyographic studies of myasthenia gravis", *Brain*, 1935; 58:470-82
- [112]. Denny-Brown D., Nevin S., "The phenomenon of myotonia", *Brain*, 1941; 64:1-16
- [113]. Dawson GD., "Cerebral responses to nerve stimulation in man", *Br Med Bull*, 1950; 6:326-29
- [114]. William F. Brown, "The Physiological and Technical Basis of Electromyography", Butterworth-Heinemann, Jan 1, 1984. PP. v
- [115]. Waldemar Karwowski, "International Encyclopedia of Ergonomics and Human Factors, Second Edition - 3 Volume Set, CRC Press, Mar 15, 2006. PP. 3107
- [116]. Franco Postacchini, "Lumbar Disc Herniation", Springer Science & Business Media, 1999, PP.279
- [117]. Eddie Harmon-Jones, Jennifer S. Beer, "Methods in Social Neuroscience", Guilford Press, May 9, 2012. PP. 70-73, 85
- [118]. Zhou Wang, Alan Conrad Bovik, "Modern Image Quality Assessment", Morgan & Claypool Publishers, Jan 1, 2006. PP. 2-3, 11-15
- [119]. Forczmański, P., Okarma, K., "Application of the central weighted Structural Similarity index for the estimation of the face recognition accuracy", *Annales UMCS-Informatica* 9(1), 2009, 153-163
- [120]. Ryszard S. Chora, "Image Processing and Communications Challenges 4", Springer Science & Business Media, Aug 15, 2012. PP.188
- [121]. Rüdiger Kramme, Klaus-Peter Hoffmann, Robert Steven Pozos, "Springer Handbook of Medical Technology", Springer Science & Business Media, Oct 2, 2011. PP.138-140
- [122]. Ponomarenko, N., Lukin, V., Zelensky, A., Egiazarian, K., Carli, M., Battisti, F., "TID2008 - a database for evaluation of full-reference visual quality assessment metrics", *Advances of Modern Radioelectronics* 10, 2009, 30-45
- [123]. Robert Burduk, Marek Kurzynski, Michal Wozniak, Andrzej Zolnierak, "Computer Recognition System 4", Springer Science & Business Media, Apr 21, 2011. PP.289-290
- [124]. George C. S. Wang, Chaman L. Jain, "Regression Analysis: Modeling and Forecasting", Institute of Business Forec, 2003, PP. 6, 57-59
- [125]. Xin Yan, Xiaogang Su, "Linear Regression Analysis: Theory and Computing", World Scientific, 2009, PP.1-3
- [126]. Ajith Abraham, Jaime Lloret Mauri et al, "Advances in Computing and Communications, Part II: First International Conference, ACC 2011, Kochi, India, July 22-24, 2011. Proceedings", Springer Science & Business Media, Ju 8, 2011. PP.122
- [127]. Ipke Wachsmuth, Timo Sowa, "Gesture and Sign Languages in Human-Computer Interaction: International Gesture Workshop, GW 2001, London, UK, April 18-20, 2001. Revised Papers", Springer Science & Business Media, Apr 24, 2002. PP.304
- [128]. Tinku Acharya, Ajoy K. Ray, "Image Processing: Principles and Applications", John Wiley & Sons, October 2005
- [129]. Brian Keelan, "Handbook of Image Quality: Characterization and Prediction", CRC Press, March 2002
-

- 
- [130]. Andreas Vogel, Brigitte Kerherve et al, "Distributed Multimedia and QoS: A Survey", IEEE MultiMedia vol. 2 Issue 2, pp. 10-19, June 1995
- [131]. Rei Takahashi, "Standardization of Multimedia QoE Evaluation Technology", IEICE technical report CS, vol. 107, no. 18, pp. 25-30 / CQ, vol. 107, no. 19, pp. 49-54, 2007.
- [132]. Keishirou Watanabe, Hitoshi Aoki et al, "Proposal of QoE evaluation framework for quality management of IPTV services", IEICE technical report CQ, vol. 107, no. 134, pp. 119-122, 2007.
- [133]. ITU-T Recommendation : P. 911, " Subjective audiovisual quality assessment methods for multimedia applications ", 1998.
- [134]. Jeng-Shyang Pan, Ching-Nung Yang et al, "Advances in Intelligent Systems and Applications – Volume 2 : Proceedings of the International Computer Symposium ICS 2012 Held at Hualien, Taiwan, December 12-14, 2012", Springer Science & Business Media, December 2012, pp 333
- [135]. Suresh Chandra Satapathy, P S Avadhani et al, "Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012 (India 2012) held in Visakhapatnam, India, January 2012", Springer Science & Business Media, December 2012
- [136]. Bhanu Prasad, S. R. Mahadeva Prasanna, "Speech, Audio, Image and Biomedical Signal Processing Using Neural Networks", Springer Science & Business Media, January 2008, pp 351 – 356
- [137]. King S., Tian G. Y. et al, "Cross-channel histogram equalization for colour face recognition", Proceedings of International Conference on Audio- and Video-Based Biometric Person Authentication 454-461, 2003
- [138]. Chen P.-H., Lin C.-J. et al, "A tutorial on nu-support vector machines", Applied Stochastic Models in Business and Industry 21:111-136., 2005
- [139]. Chew H. G., Lim C. C. et al., "An implementation of training dual-nu support vector machines", Qi, Teo, and ang, (Eds.) Optimization and Control with Applications, Kluwer, Dordrecht, 2005.
- [140]. Pramod K. Varshney, "Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data", Springer Science & Business Media, Aug 12, 2004
- [141]. Jacob Cohen, Patricia Cohen, Stephen G. West, Leona S. Aiken, "Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences", Routledge, Jun 17, 2013
- [142]. H. Julia Hannay, "Experimental Techniques in Human Neuropsychology", Oxford University Press, 1988, p.457-478