Standardizing the Evaluation of QoE by Users

A methodology to estimate the acceptance of interactive services

Graziella Cardoso Bonadia
Cláudia de Andrade Tambascia
Ewerton Martins de Menezes
Service Technology Corporate Department
CPqD Foundation
Campinas, Brazil
{bonadia, claudiat, emenezes}@cpqd.com.br

***Abstract*— Every potential service user has had some experience while using a wide variety of interactive services. Therefore, a pre-established concept is carved of what is considered to be a good service when using it. This interferes directly on how the user evaluates the service globally and, consequently, the score that will be given during a test. Being a primarily subjective evaluation, the methodology that estimates the quality of experience requires a method that could translate all this experience into consistent evaluations regarding service acceptance. The objective of this paper is to shape a methodology to estimate the QoE considering the evaluation adjustment of a group of potential users in order to achieve a reference model for future periodic evaluations for a given interactive service.**

Keywords-Quality of Experience; methodology; usability; command response time; interactive service.

I. INTRODUCTION

Quality of Experience (QoE) can be defined as the acceptance of an application or service subjectively perceived by the user regarding performance and usefulness, including system components (terminal, network, service infrastructure, etc.), as well as context of use and end-user expectations [1]. The concept of Quality of Experience has being considered for many kinds of services and it is used for a variety of issues ([2], [3] and [4]). Nevertheless, the assessment of interactive services (services that allow users to interact with the provider through some network) that takes into account the user's opinion is largely based on a single score ([5] and [6]). The most common method is capturing a score based on a discrete scale (for example, 5: Excellent, 4: Good, 3: Fair, 2: Poor and 1: Bad, [7]) or a continuous quality scale and then calculate de Mean Opinion Score [8]. In this case, the authors use a methodology based on a Difference Mean Opinion Score (DMOS), which takes into account the presentation of a sequence of videos and among them there are videos with a specific and high quality format. By comparing the scores of low quality videos with the high quality ones (see also [9]), it is possible to overcome the fact that people have distinct notions about scoring. When assessing a service, it must take into consideration that people have different expectations and they translate these expectations in distinct manner for each distinct service. But it is not rare that developers are not able to simulate and test high quality reference format hidden among other formats of the interactive service. For interface navigation testing, for example, developers have usually just one option of interface and it is not known *a priori* whether it is already on a high quality basis or not.

For this reason, it is necessary to consider another way to create a reference and then compute the acceptance based on a score. This paper will propose a methodology that combines two answers from users: The score on a well known five-point scale and the binary answer about whether the service is acceptable for use on a routine basis. This approach will provide a reference curve of score and acceptability which accounts for the target market and the assessed service.

The user QoE evaluation regarding an interactive service should consider that a given group of users may have different expectations on the use of the technology, taking into account previous experiences and distinct contexts of use. Being mainly a subjective evaluation, some mechanisms must be established to examine the differences during the evaluation, both on site and in the laboratory. In general, the test structure includes five layers, assembled according to the features and types of measures associated:

- User: target audience using the service defined by factors such as class, age, etc.
- Terminal: TV receiver (fixed or mobile), cellular phone, notebook or any physical device used to receive the service. The devices can be fixed, nomadic or mobile, which include features, such as, video resolution, size, processing capacity and control buttons.
- Service: services evaluated by this methodology include: VoD, IPTV, HDTV, chat and telephone.
- Application: data that compose the service, such as: video, audio, voice and data.
- Transport network: physical means through which the service reaches the end-user. They are networks such as: IP, RF, GPRS, 3G, etc.

These aspects are taken into account in order to structure the experiments that will consider, also, critical factors perceived by the users. These are key factors and they are described as follows:

Usability: established parameters that include browsing, presentation, authentication, remote control, usage facility, visual appearance and interface. Each of these contains aspects that must be observed during interaction as a method to assure quality of experience. Examples of usability parameters are listed in Table I.

TABLE I. EXAMPLES OF PARAMETERS REGARDING USABILITY

Usability		
ID	*Parameter*	*Characteristic*
1	Navigation	Numbers of steps in order to conclude the task.
2	Presentation	Organization of text and graphic data on the screen.
3	Authentication	Ability to perform authentication so that it is neither a discouraging factor nor a hindrance to perform the task.
4	Remote control	Adaptation of keys to the task's main functions.
5	Interface	Icon legibility considering size, definition and colors. Text legibility considering size, type and colors of the used font. Functions mapping assuring easy learning and low impact when accumulating functions.

Accessibility: parameters related to usage facility during interaction so that they meet exceptional/specific requirements as well as people with low literacy. Examples of parameters are listed in Table II.

TABLE II. EXAMPLES OF PARAMETERS REGARDING ACCESSIBILITY

Accessibility		
ID	*Parameter*	*Characteristic*
1	Text reading	Program that helps people with hearing and visual disabilities to understand the text.
2	Change of contrast	Interface that allows changing the contrast for people with partial visual disability to read texts.

Intelligibility: parameters are established for browsing through content, display of content, use of iconography, suitable language, facility to understand and the information to be displayed. Examples of parameters are listed in Table III.

TABLE III. EXAMPLES OF PARAMETERS REGARDING INTELLIGIBILITY

Intelligibility		
ID	*Parameter*	*Aspects to be evaluated*
1	Browsing	Demonstration of steps on how to browse through application contents.
2	Iconography	Size and colors of icons, representative and intuitive icons.
3	Presentation	Icon type, size and color.
4	Suitable language	No use of technical terms and accessible language.
5	Intelligibility	Suitable language and no use of technical terms and accessible language.
6	Information	Level of interest, level of understanding and amount of data displayed.

Command response time: parameters related to access time used for an application as well as the system startup. Examples of parameters are listed in Table IV.

TABLE IV. EXAMPLE OF PARAMETERS REGARDING COMMAND RESPONSE TIME

Command Response Time		
ID	*Parameter*	*Aspects to be evaluated*
1	Response time for channel switching	Acceptable time between performing the action and its result.
2	Initialization system	Time needed for the application to reinitialize.

Audio and video: parameters for audio, video and synchronization are established. Examples of parameters are listed in Table V.

TABLE V. EXAMPLE OF PARAMETERS FOR AUDIO AND VIDEO

Audio and Video		
ID	*Parameter*	*Aspects to be evaluated*
1	Video	Display quality, including verifying distortions (video quality in terms of parameters such as bit rate, encoding type, etc.)
2	Audio	Quality, including verifying distortions (audio quality in terms of parameters such as bit rate, encoding type, etc.)
3	Synchronization	Media synchronization.

For every key factor originated from the human-computer interaction, there are several studies on the test operations performed in laboratories [10].

In general, during the development of the product or service, tests are performed until a final format is achieved, ready to be launched. However, for QoE evaluation, an already tested and reformatted product is used according to the tests performed during the development phase (for usability, see [11]). Tests performed to assess the QoE will focus mainly on the user's sensation when watching a video or performing some task, which will be evaluated according to its acceptance level. The following section describes the methodology whose objective is to shape the parameters that influence the user's QoE resulting in a QoE estimate. After presenting the methodological aspects, Section 3 presents the test conditions that should be considered to perform the estimation of QoE, explained in detail in Section 4. We will also present, in Section 5, a case study in which a QoE estimate is calculated for a T-Commerce interactive service (commerce service that could be provided by broadband networks). Finally, the conclusion and future work are presented in Section 6.

II. PRESENTING THE METHODOLOGY

The QoE analysis is connected to the users' perception on the use of technology. Thus, the purpose of the QoE estimate methodology is to provide a value that would quantify the subjectivity of the user evaluation as a form of acceptance probability threshold. Generally, users are encouraged to combine the evaluations regarding the specific use of a given service in two ways: the first is to provide a score according to a pre-established scale, and the second is to determine whether or not the service is acceptable regarding its usage only (market aspects are not included). The combination of both will be used to adjust the scale and provide an answer suitable for the target audience.

Based on the model adjustment analysis, the score and the acceptance status of the service are submitted to a heuristic procedure that results in the service usage acceptance probability threshold (given its characteristics). This value, which necessarily varies between 0 and 1, is the *a priori* QoE estimate for interactive services in the after development phase. This value structures market aspects and provides clear data on the number of people within the target market that would be ready to use the service according to the evaluated format. However, the same method can be applied throughout the use of the service when already settled in the market. Therefore, a score compilation algorithm and analysis can be used as a recurring evaluation to improve the service during its life cycle.

It is worth mentioning that by calculating this estimate, a preparation is required to establish the evaluation parameters, given the complexity of some types of services. Hence, the methodology includes a brief description of the service as well as the target audience. From the service description, key factors are selected for evaluation, and for each key factor, essential tasks are determined. This selection follows a set of rules, but it continues to be subjective depending on the analysts' viewpoint. The tasks, which are arranged in groups according to the key factor, are provided to one or more groups of potential users previously selected. These groups will evaluate the service according to the laboratory environment, during the *a priori* phase of the methodology. Then, the results will be analyzed according to the procedures and statistical analysis described throughout this paper. Figure 1 illustrates the methodology plan.

Each stage relies on a set of procedures that must be followed according to the rules and the analysts' critical evaluations. After the *a priori* evaluation of the service regarding QoE, two paths emerge: i) the service is forwarded for market analysis and deployed as tested, or ii) the service returns for adjustments and new laboratory tests are performed. In the latter case, the same methodology is applied following the described structure until it achieves an acceptable level of QoE and fit to be submitted to market analysis. If required, after the service deployment, QoE levels can be controlled in real time.

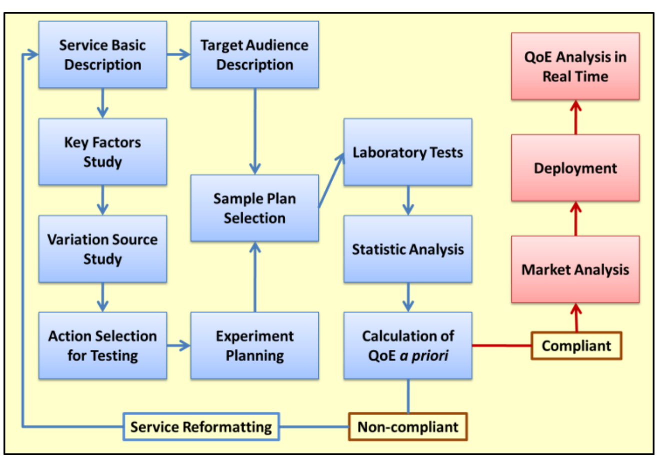

Figure 1. Methodology for QoE estimate

The next section will address the aspects related to conducting the tests in order to gather informations and users' perceptions about a general interactive service.

III. ESTABLISHING TEST CONDITIONS

The service key factors establish the laboratory test types that will be performed to estimate the QoE. By describing the service, it is possible to identify which key factors will be the centerpiece of the evaluation. So, the first step is to enumerate these factors according to the service. Based on this definition, the specialists should critically analyze the importance of each factor regarding the global use of the service. The definition of this importance, as a unit to calculate the final estimate, can adopt a multi-criteria analysis method such as, for instance, AHP ([12] and [13]) or simply the weight distribution according to the factors, so that the sum of all weights is equal to 1 (one). This importance can also be analyzed by the users that shall perform the tests. Then finally, the values are used to calculate the QoE estimate.

Regardless of the value assigned to each key factor, breakpoints can occur. Breakpoints are variable values that affect the use of a service, but they are not included in the essential setup. To state an example, the quality of the telecom network to provide a service that depends on the communication between a server and a user terminal. These breakpoints, when they occur while using the service, can affect the QoE even if the value of a given key factor is null. Therefore, the QoE estimate depends on the conditions of such variables herein named as variation sources. Some examples of variation sources that are important in the eyes of the user are:

- Network problems: network conditions can affect significantly the user's perception regarding the QoE. When structuring a service, it is expected that, even with variations in the network conditions, this fluctuation does not exceed a limit to cause the user dissatisfaction. This limit is considered the breakpoint. For QoE estimate, this variation is considered only if the probability for a breakpoint to occur is from average to high.
- Server overload problems: as more users start to use the service, a server overload is very likely to happen, resulting in a downgrade of the service. Similar to the network, the service is designed to support a maximum number of users simultaneously. For instance, in this case, the breakpoint is the number of simultaneous users that lead to the server overload. Another example is the maintenance, which reduces performance and may even interrupt the service momentarily. Only cases involving average and high probability of occurrence should be considered in the QoE evaluation.
- Corrupt service file: in general, this variation source is caused by small problems with service codes or by a type of virus. This type of variation source risk is handled throughout the testing process with developing the product. However, in some cases, it

is assumed that some service files can be deteriorated and they can be fixed though maintenance. In this case, this variation source should only be considered for QoE calculation when its probability is from average to high. When unknown *a priori*, only the QoE evaluation in real time would provide useful data for maintenance.

- Terminal used and network access service contracted by the user: these variation sources are not included in the service structure. However, when launching a service, the entrepreneur must consider the innumerous terminals with specific characteristics. When the impact is significant depending on the terminal in use, the QoE calculation method should include tests on different terminals. For the network access service contracted by the user, the service specification must inform the potential user on the minimum network service conditions required to use the service or consider the item as "problems with the network".
- Renewable content that changes the user perception: if the service depends on the addition of new content, the user perception can vary significantly when content quality specifications are changed. The QoE is estimated according to the service structure when being tested in the laboratory. If the content quality control is not performed before the item becomes a complete service, the QoE calculation method may consider several contents from which the average is drawn. Again, this must be taken into consideration only when the probability is from average to high.

In case a variation source should occur, laboratory tests must include the simulations related to the effect of the source variation. The criteria definition to simulate the variation sources depends on the average condition and its standard deviation. Then, the simulation criteria are estimated for a final setup of the laboratory tests to be performed.

Therefore, within each key factor, actions to be performed by the users in order to use the service are described. It is not unusual for the number of possible actions to be so high that it becomes unfeasible to perform the laboratory test. It is at this moment that the specialists extract the most relevant actions related to the service being evaluated. Several actions are similar, and therefore, there is no need to evaluate all of them. Similar to the laboratory test, only one user sample is used, as well as an action sample that strictly represents the core of the service. The addition of this set of actions to be performed by the service users (part of the target audience) to the different conditions from the variation sources results in the final setup of the laboratory tests.

The actions of each key factor are tested in specific laboratories using appropriate equipment and physical conditions according to the methodology objective. The amount of test conditions depends on the number of actions to be evaluated and the conditions to simulate the variation source. For this reason, to optimize the QoE estimate, the use of condition randomization methods based on the Design of Experiments theory [14] is recommended.

Thus, every test session will include a given number of test conditions and, for each test condition, the potential user will perform two types of evaluations:

- Score for every action regarding the use of service, which can be translated into image quality, easy usage, visual comfort, audio quality, etc. This score follows the scale described next: 1 : very <negative characteristic>; 2: <negative characteristic>; 3: neither <negative characteristic> nor <positive characteristic>; 4: <positive characteristic>; and 5: very <positive characteristic>

  The description that replaces <characteristic> depends on the type of task or action that is being evaluated, for instance: 1 to perform a "painstaking" task or 4 for a good image.
- Acceptance condition, i.e., if the user thinks it is acceptable to use a service with the characteristics similar to the tested one. Only two answers are possible: yes or no.

The purpose of both questions is to fine-tune a priori what is an acceptable score for the target audience and, as a result, in real time estimates, this score would be the direct outcome of the QoE estimate. After obtaining the laboratory test results, the data are consolidated and the QoE estimate methodology is performed.

IV. ESTIMATING QOE

The first analysis to be performed is the adjustment of a curve with the following variables: Score and Acceptance. The adjustment method to be used is the logistic regression [15] that takes the form of equation (1). The Score variable is independent, while the Acceptance is dependent. For the Acceptance, the value 0 (zero) is used for a negative opinion, whereas 1 (one) is used for a positive one.

$$ln\,(p/(1-p)) = a + b*score(1) + c*score(2) + \\ + d*score(3) + e*score(4) + f*score(5) \qquad (1)$$

where:

ln p/(1-p) is the acceptance or non-acceptance of the image quality;

p is the acceptance probability, for a given *score(x)*;

a, b, c, d, e, f: are the parameters adjusted by the logistic regression;

score(x): is equal to 1 if the score is *x* or *0*, otherwise

When performing the adjustment, the result will be parameters *a* and *b*. Then, to resolve the equation, parameters *a* and *b* are replaced with their corresponding adjustment values (stated here with the symbol ^). Thus, with equation (2) the value of *p* is obtained for each score value.

$$p = e^{\hat{a}+\hat{b}x}/(1+e^{\hat{a}+\hat{b}x}) \qquad (2)$$

The objective of this adjustment is to fine-tune the scores that are considered acceptable for a given set of actions. Every user has a different tolerance level that may vary when translating the perception into a score. By adjusting the acceptance regarding the score, the variation is tuned and, as a result, a reference curve is obtained. The curve is obtained by adjusting the linear regression [16] between the acceptance probability and the scores, as shown in equation (3).

$$p_{calc} = c + d * scores \quad (3)$$

As stated previously, the purpose of the QoE estimate methodology is to provide a value that would quantify the subjectivity of the user evaluation as a form of acceptance probability threshold. The acceptance threshold is associated with the lowest acceptance probability for a given X% from a sample amongst those that received the highest acceptance score. That is, the samples are organized from the highest acceptance score to the lowest acceptance score. The score to be considered is the one given by the last tester that sums X% of the total sample. Then, the final score is compared with the results from equation (3), and the QoE of the key factor being analyzed will be its corresponding calculated *p*. In this methodology, the percentage considered is 50% of the sample with the highest acceptance (average estimate). However, this value varies according to the specialists or service representative. Figure 2 illustrates this procedure.

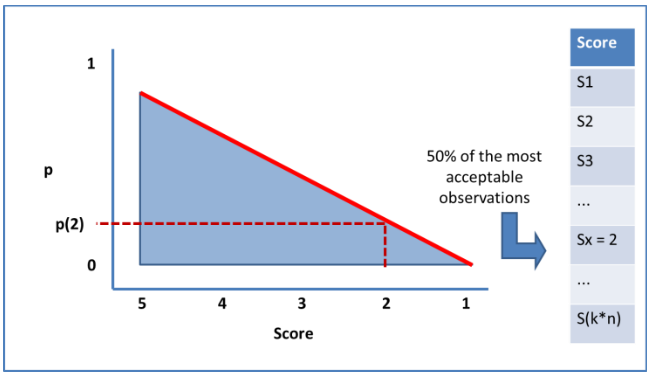

Figure 2. Acceptance probability threshold according to the best 50% score

This procedure is replicated for every key factor. To estimate the service QoE final value, the weights assigned to each factor are used. The final estimate will be calculated using equation (4):

$$QoE = w_1 * p(limite)_1 + \ldots + w_s * p(limite)_s \quad (4)$$

where:

w_i, from $i = 1$ to s, is the weight of each key factor;

s is the amount of the service key factors.

At this stage, the *a priori* methodology to estimate the QoE is completed. However, the reference curve, obtained from the linear regression, can be used to evaluate periodically the service in a real environment of usage, once it is feasible to question the user remotely, throughout its use. Consequently, it will be possible to ask the user to provide a single score.

V. APPLYING THE METHODOLOGY TO ESTIMATE QOE

As an example and in order to perform the methodology proof of concept, a service developed by CPqD and supported by FUNTTEL (Telecommunications Technology Development Fund) was selected: T-Commerce. The objective of this service is to sell products through interactive TV using the remote control. By using an interactive Digital TV terminal the user can browse through the T-Commerce window to search for products and services or check the listed ads. The key factors considered for this study are the usability and the command response time, with weights of 0.45 and 0.55, respectively. The variation source to be evaluated refers solely to the command response time factor, which assigns three server overload conditions, simulating different day-times of the service simultaneous usage. For the usability factor, a group of potential users was asked to evaluate the tasks in a laboratory environment, and for each task two questions would be answered (score and acceptance). For the command response time factor, the potential user was asked to evaluate two tasks with three server overload conditions (simulated in laboratory). Each task was evaluated regarding the command response time, from the command up to the return of the screen. Both the score and the acceptance were requested for this key factor. After obtaining the two laboratory test results, the logistic regression was performed for both key factors: usability and command response time. Probabilities for each score value were calculated according to parameters *a* and *b* resulting from the logistic regression (see the list in Table VI), as established in equations (1) and (2).

TABLE VI. ACCEPTANCE PROBABILITIES

Score	Acceptance probability – Usability	Acceptance probability – Command response time
1	0.40	0.00
2	0.86	0.60
3	0.82	0.92
4	0.92	1.00
5	1.00	1.00

After calculating the probabilities and scores, a linear regression was performed on each case being studied, as established in equation (3). The result parameters for usability factor were: c = 0.4205 and d = 0.1261. The result parameters for command response time factor were: c = -0.0159 and d = 0.24. Then, both curves are used as a reference to finally estimate the QoE. Figure 3 illustrates the reference curve for usability and command response time parameters.

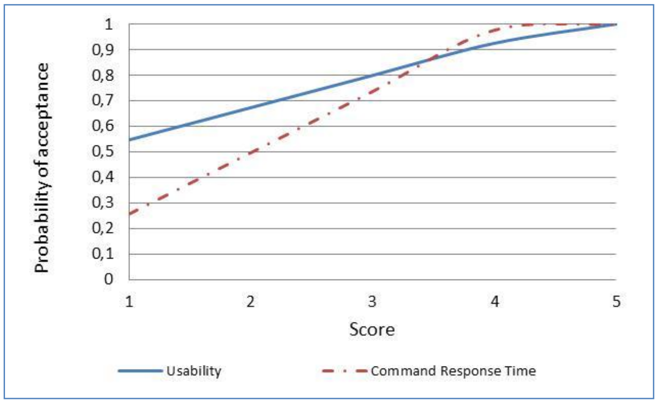

Figure 3. Acceptance probability threshold according to the best 50% score

As stated previously, the purpose of the QoE estimate methodology is to provide a value that would quantify the subjectivity of the user evaluation as a form of acceptance probability threshold. Based on the reference curve, the acceptance probability threshold is achieved considering the average observation, i.e., after organizing the data from the best to the worst scores, the score representing the observation completing 50% of all observations is selected. The score is chosen and linked to its acceptance probability, according to the reference curve. The average score of the usability factor was 4, with a probability threshold provided by the reference curve equal to 0.93. For the command response time factor, the average score was 3, which means a probability threshold of 0.70. By using the weights from each key factor, QoE *a priori* of the T-Commerce service is calculated, as established in equation (5).

$$QoE_{T\text{-}Commerce} = 0.45*0.93 + 0.55*0.70 = 0.80 \qquad (5)$$

The QoE value for the T-Commerce service shows that its current format is well accepted by the potential users. However, it is crucial to perform a follow-up of this indicator, once the command response time can change according to the number of users accessing the server simultaneously. Thus, this QoE can change significantly according to the diffusion curve regarding the service usage.

VI. CONCLUSION AND FUTURE WORK

Fine-tuning is used to provide a feeling very close to the reality of potential users, considering that the assignment of scores and words have different meanings depending on each individual. Based on the sample tuning that represents the target audience profile, the reference curve will be used as the basis for real time estimate of QoE.

The application of the methodology for QoE estimate, although presented here as a mode for the product's final evaluation, it may be used also as a mean to evaluate in development services when building the accessibility and usability features, for example. In addition, it can be used to find the lack of technical evaluation threshold, which may be of inconvenience to the user. The variation sources for each service can be evaluated in two ways. One method is to separate the estimates distinctively for comparison. The other, for the T-Commerce service, is to evaluate the situation that changes throughout the usage. The overall situations will provide the service characteristics. In this case, the number of the test environment situations must be equal to the number of real usage situations, even by estimate.

By evaluating QoE data in different communication services, it is possible to compare the results and take more accurate decisions on the design, marketing and sales. This user perception can be used to improve technical performance and achieve a satisfactory QoE. This can reduce the number of telecommunication services rejected by potential consumers.

Finally, an additional application of the methodology is introduced, which includes the periodic evaluation of QoE along the service life cycle. This application is crucial to predict the service downfall due to dissatisfaction. Therefore, improvements by adjusting QoE can be made to avoid churning. This application can be included within the service interface and its data can be extracted in real time, allowing for easy product decision making.

The methodology presented provides a mean to estimate in which level a certain interactive service would be accepted by potential user consedering technical aspects. This approach brings a probability reference curve on acceptability rather than a single score outcome. Therefore, it is usefull, especially because the user expectations about some new service is normally unknown *a priori*.

The future steps of the present work are to apply the methodology to other interactive services and to validate the outcomes in a real market basis.

REFERENCES

[1] ITU-T, Recommendation P. 10/G. 100 Amendment 1, "New Appendix I – Definition of Quality of Experience (QoE)", International Telecommunication Union, 2007.

[2] I. Wechsung, K. P. Engelbrecht, C. Kühnel, S. Möller, and B. Weiss, "Measuring the Quality of Service and Quality of Experience of Multimodal Human-Machine Interaction", Journal on Multimodal User Interfaces, vol. 6, Issue 1-2, July 2012, pp. 73-85.

[3] H. A. Tran and A. Mellouk, "QoE Model Driven for Network Services", Wired/Wireless Internet Communications – Lecture Notes in Computer Science, vol. 6074, 2010, pp. 264-277.

[4] D. K. Moore *et al*., "Linking Users' Subjective QoE Evaluation to Signal Strength in an IEEE 802.11b/g Wireless LAN Environment", EURASIP Journal on Wireless Communications And Networking, vol. 2010, pp. 1-12, Article ID 541568, doi: 10.1155/2012/541568.

[5] ITU-T, Recommendation G.1080, "Quality of experience requirements for IPTV services", International Telecommunication Union, 2008.

[6] P. Huang, Y. Ishibashi, N. Fukushima, and S. Sugawara, "QoE Assessment of Group Synchronization Control Scheme with Prediction in Work Using Haptic Media", International Journal of Communications, Network and System Sciences, vol. 5, 2012, pp. 321-331, doi:10.4236/ijcns.2012.56042.

[7] ITU-R BT.500-13, "Methodology for the Subjective Assessment of the Quality of Televiosion Pictures", International Telecommunication Union, 2012.

[8] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, " Study of Subjective and Objective Quality Assessment of Video", IEEE Transactions on Image Processing, vol. 19, Issue: 6, June 2010, pp. 1427-1441.

[9] H. R. Sheikh and A. C. Bovik, "An Evaluation of Recent Full Reference Image Quality Assessment Algorithms", IEEE Transactions on Image Processing, vol. 15, Issue: 11, Nov 2006, pp. 3440-3451.

[10] L. S. G. Piccolo, G. C. Bonadia, and C. A. Tambascia, "Technical application and HCI methods to evaluate interactive TV service QoE" (Aplicação de técnicas e métodos de IHC para avaliar a QoE de serviços de TV interativa), IHC2010, vol. I, 2010, pp. 225-230, Belo Horizonte, Brazil.

[11] C. A. Tambascia, E. M. Menezes, and R. E. Duarte, "Usability evaluation using eye tracking for iconographic authentication on mobile devices", Mobility 2011 – The First International Conference on Mobile Services, Resources, and Users, October 23-29, 2011, Barcelona, Spain.

[12] T. L. Saaty, "The Analytic Hierarchy Process – Planning, priority setting, resource allocation", McGraw-Hill, Inc. , 1980.

[13] D. Kim, "An Integrated Framework of HoQ and AHP for QoE Improvement of Network-Based ASP Services", Engineering Annals of Telecommunications, vol. 65, no. 1-2, 2010, pp. 19-29, doi: 10.1007/s12243-009-0143-9.

[14] D. C. Montgomery, Design and analysis of experiments, 8th ed., New York: John Wiley & Sons INC., 2012.

[15] D. W. Hosmer and S. Lemeshow, Applied Logistic Regression, 2nd ed., New York: John Wiley & Sons INC., 2000.

[16] D. C. Montgomery, E. A. Peck, and G. G. Vining, Introduction to Linear Regression Analysis, 5th ed., New Jersey: John Wiley & Sons INC., 2012.