https://doi.org/10.2352/ISSN.2470-1173.2018.12.IQSP-235

Assessing the Quality of Videoconferencing: from Quality of Service to Quality of Communication

Muhammad Shahid, Mekides Assefa Abebe, Jon Yngve Hardeberg ; The Norwegian Colour and Visual Computing Laboratory, Department of Computer Science, Norwegian University of Science and Technology, Gjøvik, Norway

Abstract

Various approaches for the assessment of quality of experience for video conferencing exist that try to quantify the level of satisfaction of the user. In order for videoconferencing to really succeed as a substitute of face-to-face meetings, it is constructive to explore the associated quality of communication. One of the most significant factors to ascertain in this regard would be the assessment of the level of user satisfaction for a videoconference instead of a face-to-face meeting. Various recommendations of ITU-T related to this field (such as ITU-T Recommendation P.920) deal with subjective experiments that involve performing interactive tasks in order to quantify the impact of terminal and communication link performance. In addition to looking at different quality paradigms, in this paper, we review a number of subjective studies on videoconferences and investigate the possibilities of discovering more of the user feedback in order to assess the perceived quality of communication experienced in a videoconference. Based on the review, we present a set of future work items which can be useful in finding a comprehensive definition for quality of communication.

Introduction

Videoconferencing is becoming more popular, developing to advanced forms (such as telepresence), and being offered by various service providers both as a free or a premium service. Cisco reported a growth of 25% in videoconferencing during 2014-2015 [1] and Skype reported 10 times increase in group video-calling in a period of less than two years [2]. Broadly speaking, the experience of videoconferencing can be divided into the ordinary or conventional videoconferencing and telepresence (or e-presence) [3]. International Telecommunication Union (ITU) recommendation (ITU-T Rec. F. 734) defines telepresence in terms of an interactive audio-visual communications experience between remote users with a strong sense of realism and presence [4]. The same recommendation enlists a set of requirements on a telepresence system including actual size high definition imaging, immersive presence, and provision of awareness of both direct gaze and averted gaze. These and other requirements greatly increase demands on the infrastructure such as hardware and software resources, thus making a telepresence system increasingly expensive. Such constraints motivate development in the traditional videoconferencing such that it can approach the experience offered by a telepresence system.

Videoconferencing is being used in a variety of fields ranging from private consumer applications to more advanced purposes, such as business, education, and health-care related applications. The level of audiovisual quality of the videoconferencing system that can be experienced by the users also varies with the applications. Systems deployed in higher level business sectors usually have dedicated high capacity devices and a reserved bandwidth for a reliably good audiovisual quality. However, individual consumers of private services usually prefer not to use expensive services. Their systems' audiovisual quality can vary depending on the capacity of their devices as well as the availability of the network. Despite the rapid advancements of communication technologies, different factors such as the transmission techniques, hardware, the information coding systems, and the users' viewing condition can have high impact on the perceived quality of the transmitted audiovisual content and the overall communication quality.

The ultimate goal of the videoconferencing systems is preferably to allow users communicate as efficiently as the face-to-face meetings. To be able to assess such capabilities of the systems, both the technical performance of the systems (commonly known as Quality of Service: QoS) and the users' satisfactions (Quality of Experience: QoE) have to be measured. It is known that the most accurate method for measuring perceived audiovisual quality is through subjective experiments. Also, for more real-time and efficient implementation purposes, the resulted data sets are used to develop objective quality assessment models which can reproduce the subjective scores, as closely as possible [5]. Consequently, over the years, a number of subjective experiments have been conducted to collect users' opinions and numerically represent perceptual audiovisual qualities.

The subjective experiments allow researchers and system designers to investigate different audiovisual processing and transmission chain related technical factors of the videoconferencing systems, which can directly or indirectly influence the systems' perceptual audiovisual qualities. The standard experimental guidelines introduced through several ITU technical committee recommendations were followed by most of the experiments. ITU-T recommendations related to audiovisual applications [6, 7, 8, 9, 10] include definitions of technical terms as well as guidelines for different subjective experiments. Particularly, recommendations like ITU-T Rec. P.920 present details on the experimental design, the methodology, contents as well as digital materials to be used during a particular audiovisual quality assessment experiments of multimedia services.

Most of the state-of-the-art videoconferencing quality assessments are intended for the evaluation of the performance of the system's design, the network architecture or the users' satisfaction in terms of the perceptual audiovisual qualities. Quality factors used for the measurements of QoS as well as QoE differ from one experiment to the other due to the experiments' intended purposes. Many of the subjective studies related to the evaluation of conventional videoconferencing systems, opt to consider factors affecting the users' satisfaction which are only related to the hardware, audiovisual signal processing and network transmission performance.

However, considering the exponential growth in the capacity of computing devices and the availability of higher Internet bandwidth, the usual system performance and perceptual quality measurements might not be enough. We argue that, even with a level of high perceptual quality, it is still pertinent to explore the videoconferencing services for user satisfaction in terms of quality of the overall communication. Additional factors (such as the naturalness, the accuracy, the sense of co-presence, as well as others) of the users' experience compared to the real face-to face meetings have to be analyzed. The level of satisfaction of the people engaged in a videoconference, which is usually assessed in terms of QoE, is of paramount importance in order for it to succeed as a full replacement of the face-to-face meetings. Thus, it is pertinent to explore the videoconference services for user satisfaction in terms of what we term as quality of communication (QoC).

In general, the objectives of this study are the following.

- Discussing the definitions and boundaries of QoS and QoE.
- Summarizing the state of the art subjective experiments, focusing on the measurement and evaluation of the conventional videoconferencing systems' audiovisual quality.
- Discussing the term QoC and recommending possible experimental design for subjective QoC measurements of videoconferencing systems.

Quality of service (QoS) and Quality of Experience (QoE)

ITU-T. Rec. G.1000 defines Quality as "the totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs" [7]. The type of needs that the system has to satisfy varies depending on the task at hand. Considering the multimedia quality of videoconferencing systems, the two frequently used quality terms are QoS and QoE. The terms have been used, very regularly, in different systems' quality assessment reports. Sometimes, it gets harder to make a distinction between them since they have been used interchangeably and the difference between them is not clearly defined [11]. Consequently, a few QoS subjective experiments reported their QoS results using the term QoE and vice versa. As it is mentioned in [11], the similarity and difference between the two terms have to be clearly defined and used accordingly.

In this regard, the ITU-T committee [6, 7, 8], European Telecommunications Standards Institute (ETSI) [12] and Internet Engineering Task Force (IETF) [13] have released a number of recommended definitions for the terms. However, we will be relying on the definitions provided by ITU-T committees for the rest of our discussion. For the related definitions of the terms related to telecomunications and Internet network performance, the readers can refer to the above mentioned standards and other related references [11].

ITU-T. Rec. E.800 defines QoS as "Totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service" [6]. The term is mainly defined from a system's perspective and it mostly shows the performance of the service offered to the user. Whereas, QoE (according to ITU-T. recommendation of P.10/G.100) is defined as "The overall acceptability of an application or service, as perceived subjectively by the end-user". The recommendation also note " that QoE includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.) and that the overall acceptability may be influenced by user expectations and context" [8]. Therefore, QoE is a more user oriented and broad concept.

It can also be seen that there is an indirect relationship between the two quality paradigms. Since the systems performance also affects QoE, a good QoS of a system can lead to a better QoE. This, however, does not mean that the system achieved a satisfactory QoE considering that QoE is additionally influenced by user intent and expectations. The users' expectations, on the other hand, are in turn affected by factors such as: the users' device characteristics, the viewing environment, the audiovisual content, and other related factors. The relationship between QoS and QoE for telepresence applications are more clearly presented in Wu et al.'s work [14]. QoS is presented as "a set of measures for tuning or quantifying the performance of applications, systems, and networks". Factors affecting QoS in several levels of the technological environment also affect the users' Quality of Experience. In addition, QoE is affected by users' influences that are triggered by their cognitive perceptions which in turn have an impact on their behavior.

State-of-the-art of Subjective Experiments for Quality Assessment of Videoconferencing

Perceptual audiovisual quality assessment is effectively performed based on subjective scores which are collected through several subjective experiments. In the experiment, a group of people participate and are presented with a set of selected audiovisual sequences or engage in a two way conversation. During and/or after the task, they are required to assess the quality. The appropriate contents of the stimuli, the experimental methodology and designs, the used materials and experimental environments, as well as the type of observers are not always the same for different applications. For example, the ITU-T provides recommendations on the general guidelines of subjective experiments of video quality assessments for multimedia applications [15, 9], Internet video and distribution quality television in any environment [16, 10], audio and audiovisual multiparty telemeetings [17], and many more. The focus of this paper, however, is mainly the subjective quality evaluation of videoconferencing services. In the following, state-of-the-art in this area.

Generally speaking, QoS is extensively used in the telecommunications field mainly for performance measurement of IP-based broadband, wireless, and multimedia services. For example, subjective experiments performed for determining transmission quality of the transmission equipment of different telecommunication networks are conducted based on the ITU-T. Rec. P.800 [18] guidelines. Evaluations like cellular network performance for speech quality [19], large scale management of multimedia flows [20], and other related studies use QoS as quality measures. These type of experiments usually lead to application-specific quality metrics and can not be generalized for all types of situations.

However, considering videoconferencing systems and current advancements in the performance of network architectures; application and user expectation related factors are the most dominating in subjectively perceived audiovisual quality [21]. Hence, the design of most of the current videoconferencing systems are guided by QoE. To measure the QoE of videoconferencing systems, researchers consider technical factors that are mainly related to audiovisual signal acquisition technologies, audio/video encoding/decoding algorithms and network transmission/reception protocols. Some examples of factors for the different parts of videoconferencing systems are summarized in Table 1.

There are a number of quantitative representations of QoE proposed over the years [22, 23, 10, 21].According to ITU-T Rec. P.920, the measure of QoE can be represented by Mean Opinion Score (MOS) of a Likert-scale subjective testing ranks which are collected from observers [10]. QoE based subjective experiments have been performed for applications such as: quality comparisons of existing videoconferencing applications [24], evaluation of QoE factors on telepresence meetings [25], user preference comparison of 2D versus 3D videoconferencing scenarios [26], study of the impacts of factors like delay, bitrate and packet-loss rate on QoE in a group based settings and different level of user interactions [27, 28], performance evaluations and optimization of mobile networks [29], quality assessment via crowdsourcing [30].

The results of these studies show that users with higher levels of interaction were quicker to perceive audiovisual quality degradation. They also show the people's great appreciation for the added value of 3D in videoconferencing. However, if the level of interaction does not require 3D features such as depth perception, people tend to prefer the 2D videoconferencing version. Many of the above studies are at least a couple of years old and some of them were using simulated video conversations rather than using video cameras for live conversational tasks. Multimedia technologies are, however, advancing enormously. Therefore, many of the experiments need to be revisited with current and up-to-date technologies as well as applications.

We also noticed that the individual experiments mainly assess a particular architecture, hardware or videoconferencing application. On the other hand, people are currently utilizing a range of different multimedia devices for their day to day personal or professional communications. For instance, one can prefer using a well designed and advanced videoconferencing system (such as Logitech and Google's Hangouts Meet material) and others might tend to just join the communication from their personal mobile phone. Up to the writing time of this work, we have not come across of any audiovisual quality study that takes this issue of diverse technology based communication into consideration.

In addition to the studies that are separately considering QoS and QoE measures, few studies combine both quality measures for a better design of multimedia solutions. Mellouk et al. [31], for example, combines the concepts of QoS and QoE for designing a better content distribution network architecture. They consider quality in perspective of the user, the application as well as the network components. Wu et al. [14], similarly, proposed a quality framework for telepresence systems by incorporating both quality measures. Their method was designed based on inputs from psychology, cognitive sciences, sociology, and information technology. The QoS and QoE measures and factors of telepresence were clearly described and metrics such as 'interactivity', 'visual context quality', and 'vividness breadth' were quantified based on the participants' scores gathered through several interactive experiments.

The studies of Wu et al. and Berndtsson et al. go beyond the conventional boundaries of QoE. They raised questions related to people's enjoyment, concentration, attitude towards the system and the system's ease of communication [14, 25]. However, the subjective experiment tasks used were not easy to adapt for conventional videoconferencing quality assessments. The tasks were also distracting since the users' were needed to separate their attentions from the screen many times.

To this end, one can plan to adapt and include certain tasks such as the two experiments [14, 25], ranging from participants involved in selling, doing negotiations, and playing interactive games. The same tasks could be performed beforehand in face-to-face meetings for the sake of comparison and knowledge-feedback which can also be used in the design of the experiments.

Experimental Procedures

As discussed in the previous sections, the audiovisual qualities of videoconferencing services are mostly measured and evaluated through subjective experiments. To reliably and accurately measure such services, the experimental situations and procedures have to be well designed and carefully planned. For this reason, organizations like ITU have been providing several recommendations and experimental guidelines. The guidelines are usually comprised of information such as the specification of experimental stimuli, experimental conditions and designs, types of observers/participants and also processing and presentations of experimental results.

In this regard, ITU-T Rec. P.920 and P.911 present interactive test methods for subjective audiovisual quality assessment of multimedia services [10, 15, 9]. We present here a short review of experimental procedures, with respect to the ITU-T P.920 guideline, which have been extensively used in various subjective studies.

Experimental Stimuli

As of now, different types of stimuli have been used for the evaluation of subjective audiovisual quality. Short stimuli, which are mainly 8 to 10 seconds long, were used for the assessment of momentary quality assuming that the quality of the system is stable over the duration of the stimulus. These type of stimuli are useful to measure the quality degradation magnitude which tends not to change over time. However, some quality impairments such as network packet loss have to be evaluated over time. In such cases, long stimuli which are, according to ITU-T Rec. P.880 [32], between 45 seconds to 3 minutes long should be used instead. Both the short and long types of stimuli require to be presented and rated by observers. The quality rating is mainly done while or after perceiving the stimuli quality, stimulus by stimulus. This type of assessment process is found to be passive, destruc-

Table 1: QoE quality factors

Related to:	**Audio**	**Video**
Acquisition technologies	sampling frequency pre-processing algorithms	spatial and temporal resolution camera setting, noise and motion
Multimedia Encoding/Decoding	Same as the video	coding latency type of the coding algorithms DCT quantization step and function synchronization adjustments post processing during decoding
Network transmission/reception	maximal transfer unit size of the network size of the jitter buffer the use of forward error correction packet loss concealment algorithms	same as the audio
Others	communication delay (interactivity) ...	...

tive and disturbing for the observers. Other possible assessment methods which can be applied for these types of stimuli are also not comfortable for more interactive evaluations, like videoconferencing [33, 21] .

To reduce the artificiality of such assessment environments and to stimulate natural interactive communications, ITU-T Rec. P.920 provides conversational type of stimuli for videoconferencing applications. The conversational stimuli are some type of tasks that should be performed by the observers, in a two-way conversation, and they should be at least 5 minutes long. To formulate these type of tasks, the recommendation additionally provides the following three guidelines together with example tasks such as name guessing, picture comparisons, story telling, and block building.

- ”The task should be designed such that, during their conversation, the subjects primarily maintain their attention on the audiovisual terminal.”
- ”The task must have sufficient face value, that is, it must resemble real-life audiovisual communication to a sufficient degree. In particular, it is preferable that the task be performed by two subjects and not by one subject and an experimental leader.”
- ”The task must yield reproducible quantitative results that represent adequate measures of communication efficiency.”

Additionally, some of the subjective studies rely on already captured videos from different video databases while others use simulation of real-time videoconferencing situation. However, many of the audiovisual quality evaluations are performed with a real videoconferencing conversations in a controlled room. Very few others use real videoconferencing in a real environmental conditions. However, there are also some studies on audio speech quality which show that the controlled environment results were actually equivalent to the ones found from real world environments [34]. However, there is still more to investigate on the videoconferencing perceptual quality in a real usage situation.

Experimental Methodologies

Almost all of the audiovisual quality related ITU-T recommendations suggested to use the absolute category ratings (ACR) method [35]. In ACR methodology, observers are presented with a stimuli or take part in a short task based conversation. After cognitively assessing the stimulus, the observers are asked to rate it. The frequently used and also the recommended rating scale associated to this methodology is mean opinion score (MOS) scale. ITU-T Rec. P.920 points out that the observers should rate the overall audiovisual quality, the audio, and then the video quality of the presented stimulus. This particular order of rating is proposed to inhibit the observers from giving the average rating for the overall audiovisual quality.

However, more recent studies show the need for looking a more exploratory methods beyond MOS. Alternative methods such as standard deviation, cumulative density functions (CDF), and quantiles as well as other rating methods are examined in [23]. Their results show that the proposed alternatives were able to better explore the uncertainty of the opinion scores.

Experimental Conditions and Design

It is also very essential to organize the experimental process so that the chosen methodology can be implemented without different distractions and observer fatigue that can happen during the assessment process. In this stage, the number of stimuli as well as their distribution and order of presentation are determined. The experiment conditions such as the room lighting and equipment characteristics as well as the experimental design are also be chosen. It should be noted that the appropriate conditions and designs which have to be chosen for the subjective assessments are different depending on the nature of the application under investigation. Since we are only interested in videoconferencing applications, we present some of the recommendations of ITU-T Rec. P.920.

As per the guidelines of P.920, the total number of stimuli should be constrained to avoid observers’ fatigue because the duration of a single stimulus for audiovisual quality evaluation is chosen to be at least 5 minutes long. The viewing and listening condition are also required to be fixed and equal for all laboratories taking part in the same experiment. Additionally, the size and type of monitors, speakers as well as cameras should be appropriate for the application under evaluation. A list of typical viewing and listening conditions are also provided in the recommendation [10].

The selection of the experimental design, however, is a very critical task due to a possible impairment of observers' judgment by the context effects that are likely to occur. These effects include the tendency of observers: to categorize the stimuli and use those internal category equally by continuously updating their standard of judgments, to always rate a good quality stimuli to the highest rank if it is preceded by bad stimuli, and to use internal reference that they keep comparing to. This type of situations are often avoided by making range of quality degradation ecologically valid, the level of degradation quantitatively balanced and by training observers with example pre-test. According to P.920 of ITU-T, Latin, Greco-Latin, or Latin-Latin square design are recommended experimental designs for subjective assessment of the audiovisual quality of videoconferencing systems.

For more detail information about conditions and experimental designs, we suggest our readers to refer to the ITU-T recommendations and related books [10, 15, 9, 21].

Observers

As stated in the above references, at least 16 observers have to participate in a particular subjective assessment of audiovisual quality. It is also mentioned that the maximum of 40 observers are sufficient since there are no statistical accuracy improvements expected beyond that point. All of this observers have to be non-experts in a sense that their normal work is not related to either audio or video technologies. However, a group 4 to 8 experts can provide indicative results.

As per P.920 recomendation of ITU-T, a wide range of observers, including elderly as well as viewing and hearing-impaired people, should also be allowed to participate in such types of experiments.

QoC: Discussion and Conclusion

Most of the experiments to measure QoE are designed or inspired from its definition which is anchored in the perceived quality of the multimedia content. Clearly, in order to evaluate a videoconferencing system to be worthy of substituent for a face-to-face meeting, the relevant subjective experiment should be guided by a definition which is more directive. That naturally brings up the need of a term which is aimed at direct evaluation of how much the user felt close to the real-world while conducting a videoconference-based conversation. An analysis of the review presented in the earlier sections broadly highlights the need of measuring the quality of videoconferencing by including factors that encompass the naturalness of the experience. While QoS mainly focuses on level of salification achieved by a service, the focus of QoE is more focused on the its level of acceptance. By examining the ITU-T definitions of QoS and QoE, and how they are applied in practice for performing quality measurements, Figure 1 provides their comparative view in relation to what we propose as QoC.

In literature, the term QoC has been used largely to denote affectivity of a communication happening between two stakeholders such as a doctor and patient [36]. Close to the field of videoconferencing, one of the earliest uses of this term is found in a paper related to the communication system of a power company [37]. Similarly, the use of an equivalent term 'Quality of Communicating' is found in [38], where it is employed in order to

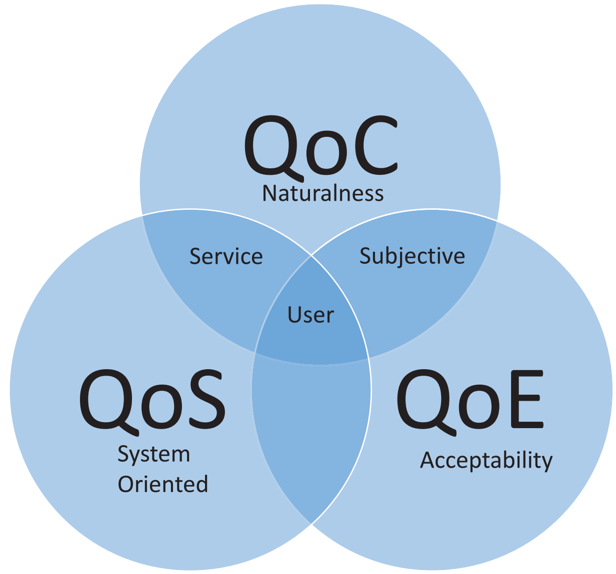

Figure 1: A comparative view of the three paradigms of quality measurements.

investigate 'Quality of Experience' of remote meetings. As such, we do not find any formal definition of QoC in the context of video communications. Nonetheless, Liu *et al.* define the Quality of Communication Experience with reference to conducting negotiations in intercultural and intracultural scenarios [39]. They identify three factors contributing to it–namely, Clarity, Responsiveness, and Comfort. Moreover, [40] estimate the importance of avatars in enhancing the QoC in remote meetings. They propose to quantify QoC by collecting subjective feedback in terms of: how natural the conversation felt, degree of involvement in the conversation, sense of co-presence and positive or negative evaluation of the conversation partner. Apart from the aforementioned scenarios, up to the time of writing, we were not able to find any formal definition of the term QoC which is adequate for accurately describing videoconferencing quality. We propose to adapt these factors in video communications by considering the clarity, responsiveness, and comfort resulting from the use of a videoconferecing solution. In line with the definition of Quality of Experience (QoE) in this context [14], we propose to conceive QoC as a quantity that is dependent on the service of a videoconferencing solution as a substitute to face-to-face meeting. We understand that a definition of QoC for videoconferencing can be evolved by keeping in consideration some factors that are to be determined through related subjective experiments.

In order for obtaining a conceptual framework and a functional definition of QoC, it would be constructive to conduct subjective experiments of videoconferencing. Moreover, in continuation of the aforementioned related work and to offer some insights into future research directions in order to propose a system of videoconferencing which can substitute the face-to-face meetings in a competitive manner, relevant subjective experiments have to performed for necessary knowledge discovery. To this end, the following list gives an overview of what should be done in order to explore this topic further:

- So far, the main focus of the subjective studies related to QoE of videoconferencing has been on evaluating the audio-visual quality. Though some studies [25] do obtain participants' feedback on how close is the experience of video conference to the real life meeting but a related detailed analysis is missing. Through designing subjective experiments by keeping the factors of naturalness in consideration, it would be possible to determine the factors that are significant for making the user of videoconference feel close to a real life face-to-face meeting experience.
- In relation to telepresence, ITU-T Rec. F. 734 [4] provides a list of requirements including the ones on user experiences and control functions; these requirements can be adapted to enhance the design factors of traditional videoconferencing.
- As the previous research has already focused on system factors (such as bandwidth, delay, depth (2D Vs 3D)), future focus can be more on the physical environments factors [41] that deal with perception and cognition about the environment, spatial cognition, social processes such as personal space and crowding (groups), and the analysis of human transactions with the built (room conditions).
- There has not been much focus on the intelligent information capturing (such as whiteboard contents, information radiators, etc.) in the context of videoconferencing. There has been some research on digital whiteboard [42] but it is known that digital whiteboards are not very common yet.
- Apart from adding to the body of knowledge in the area of videoconferencing, it is expected that after performing such experiments, a definition of QoC can be proposed such that it can enable building videoconference systems that are capable of providing the experience of naturalness or face-to-face meetings.

Conclusion

In this paper, a review on the state-of-the-art of subjective studies for videoconferencing has been presented. The main motivation behind the review has been to investigate the QoE for videoconferencing services in order to ascertain if knowledge of QoE is enough for successful replacement of face-to-face meetings with videoconference meetings. Starting with definitions of QoS and QoE, the presented review highlights the related research and the need of further studies. A number of research directions are identified to envisage QoC for videoconferencing. We expect that our findings can be helpful in performing further studies in this area, resulting into measurable improvements in the videoconferencing services.

References

[1] C. V. N. Index, "The zettabyte era: Trends and analysis," *Cisco white paper*, 2016.

[2] "Ten years of Skype video: yesterday, today and something new," https://blogs.skype.com/stories/2016/01/12/ten-years-of-skype-video-yesterday-today-and-something-new/, [Online; accessed 12-April-2017].

[3] "Telepresence vs. videoconferencing: Resolving the cost/benefit conundrum," https://www.ivci.com/telepresence-vs-videoconferncing-resolving-the-costbenefit-conundrum/, 1 2008, [Online; accessed 12-April-2017].

[4] "Definitions, requirements and use cases for telepresence systems," *ITU-T Recommendation*, vol. F.734, 2014.

[5] M. Shahid, "Methods for objective and subjective video quality assessment and for speech enhancement," Ph.D. dissertation, (Doctoral dissertation, Blekinge Institute of Technology, Karlskrona, Sweden, 2014.

[6] "Definitions of terms related to quality of service," *ITU-T Recommendation*, vol. E.800, 09 2008. [Online]. Available: https://www.itu.int/rec/T-REC-E.800-200809-I

[7] "Communications quality of service: A framework and definitions," *ITU-T Recommendation*, vol. G.1000, 11 2001. [Online]. Available: https://www.itu.int/rec/T-REC-G.1000-200111-I/en

[8] "Vocabulary for performance and quality of service," *ITU-T Recommendation*, vol. P.10, 07 2006. [Online]. Available: https://www.itu.int/rec/T-REC-E.800-200809-I

[9] "Subjective audiovisual quality assessment methods for multimedia applications," *ITU-T Recommendation*, vol. P.911, 12 1998. [Online]. Available: http://handle.itu.int/11.1002/1000/4538

[10] "Interactive test methods for audiovisual communications," *ITU-T Recommendation*, vol. P.920, 2000.

[11] M. Varela, L. Skorin-Kapov, and T. Ebrahimi, *Quality of Service Versus Quality of Experience*. Springer International Publishing, 2014, pp. 85–96. [Online]. Available: https://doi.org/10.1007/978-3-319-02681-7-6

[12] "Network aspects (NA),general aspects of quality of service (qos) and network performance (NP)," *ETSI Technical Reports*, vol. ETR.003, Aug. 1994. [Online]. Available: http://www.etsi.org/deliver/etsi_etr/001_099/003/01_60/etr_003e01p.pdf

[13] E. Crawley, H. Sandick, R. Nair, and B. Rajagopalan, "A framework for QoS-based routing in the internet," 1998.

[14] W. Wu, A. Arefin, R. Rivas, K. Nahrstedt, R. Sheppard, and Z. Yang, "Quality of experience in distributed interactive multimedia environments: Toward a theoretical framework," in *Proceedings of the 17th ACM International Conference on Multimedia*, ser. MM '09. ACM, 2009, pp. 481–490.

[15] "Subjective video quality assessment methods for multimedia applications," *ITU-T Recommendation*, vol. P.911, 09 1999. [Online]. Available: https://www.itu.int/rec/T-REC-P.910-199909-S/en

[16] "Methods for the subjective assessment of video quality, audio quality and audiovisual quality of internet video and distribution quality television in any environment," *ITU-T Recommendation*, vol. P.913, 03 2016. [Online]. Available: https://www.itu.int/rec/T-REC-P.913-201603-I/en

[17] "Subjective quality evaluation of audio and audiovisual multiparty telemeetings," *ITU-T Recommendation*, vol. P.1301, 07 2012. [Online]. Available: https://www.itu.int/rec/T-REC-P.1301-201207-S/en

[18] "Methods for subjective determination of transmission quality," *ITU-T Recommendation*, vol. P.800, 08 1996. [Online]. Available: https://www.itu.int/rec/T-REC-P.800-199608-I/en

[19] Ali, Y. Dalveren, R. Kadioğlu, and Kara, "Quality of service assessment: a case study on performance benchmarking of cellular network operators in turkey," *Turkish Journal of Electrical Engineering and Computer Science*, vol. 23, pp. 548 – 559, 2015.

[20] J. L. García-Dorado, J. Aracil, J. A. Hernández, S. Lopez-Buedo, J. E. L. de Vergara, P. Reviriego, G. Huecas, S. Pavón, and J. Quemada, *A Quality of Service Assessment Technique for Large-Scale Management of Multimedia Flows*. Springer Berlin Heidelberg, 2007, pp. 173–176. [Online]. Available:

https://doi.org/10.1007/978-3-540-75869-3_15

[21] B. Belmudez, *Audiovisual Quality Assessment and Prediction for Videotelephony*. Springer, 2015. [Online]. Available: http://www.ebook.de/de/product/23542787/benjamin_belmudez_audiovisual_quality_assessment_and_prediction_for_videotelephony.html

[22] R. Beauregard and P. Corriveau, *User Experience Quality: A Conceptual Framework for Goal Setting and Measurement*. Springer Berlin Heidelberg, 2007, pp. 325–332. [Online]. Available: https://doi.org/10.1007/978-3-540-73321-8_38

[23] T. Hoßfeld, P. E. Heegaard, M. Varela, and S. Möller, "QoE beyond the MOS: an in-depth look at QoE via better metrics and their relation to MOS," *Quality and User Experience*, vol. 1, no. 1, p. 2, Sep 2016. [Online]. Available: https://doi.org/10.1007/s41233-016-0002-1

[24] Y. Lu, Y. Zhao, F. Kuipers, and P. Van Mieghem, "Measurement study of multi-party video conferencing," in *Proceedings of the 9th IFIP International Conference on Networking*, ser. NETWORKING'10. Springer-Verlag, 2010, pp. 96–108.

[25] G. Berndtsson, M. Folkesson, and V. Kulyk, "Subjective quality assessment of video conferences and telemeetings," in *19th International Packet Video Workshop (PV)*, May 2012, pp. 25–30.

[26] H. Rizek, K. Brunnström, K. Wang, B. Andren, and M. Johanson, "Subjective evaluation of a 3D videoconferencing system," pp. 90 110A–90 110A–10, 2014.

[27] M. Schmitt, S. Gunkel, P. Cesar, and D. Bulterman, "The influence of interactivity patterns on the quality of experience in multi-party video-mediated conversations under symmetric delay conditions," in *Proceedings of the 3rd International Workshop on Socially-Aware Multimedia*. ACM, 2014, pp. 13–16.

[28] M. Schmitt, J. Redi, P. Cesar, and D. Bulterman, "1Mbps is enough: Video quality and individual idiosyncrasies in multiparty hd videoconferencing," in *Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, June 2016, pp. 1–6.

[29] C. N. Pitas, A. D. Panagopoulos, and P. Constantinou, "Speech and video telephony quality characterization and prediction of live contemporary mobile communication networks," *Wireless Personal Communications*, vol. 69, no. 1, pp. 153–174, Mar 2013.

[30] M. Shahid, J. Søgaard, J. Pokhrel, K. Brunnström, K. Wang, S. Tavakoli, and N. Gracia, "Crowdsourcing based subjective quality assessment of adaptive video streaming," in *Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*, Sept 2014, pp. 53–54.

[31] A. Mellouk, S. Hoceini, and H. A. Tran, "Quality of experience vs. quality of service : Application for a CDN architecture," in *2013 21st International Conference on Software, Telecommunications and Computer Networks - (SoftCOM 2013)*, Sept 2013, pp. 1–8.

[32] "Continuous evaluation of time-varying speech quality," *ITU-T Recommendation*, vol. P.880, 05 2004. [Online]. Available: https://www.itu.int/rec/T-REC-P.880-200405-I/en

[33] B. Weiss, S. Möller, A. Raake, J. Berger, and R. Ullmann, "Modeling call quality for time-varying transmission characteristics using simulated conversational structures," *Acta Acustica united with Acustica*, vol. 95, no. 6, pp. 1140–1151, 2009.

[34] L. Gros, N. Chateau, and S. Busson, "A comparison of speech quality judgments in laboratory and in real environment," in *Audio Engineering Society Convention 114*, Mar 2003. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=12590

[35] M. Marshall, *OCR Psychology Student Guide 3: Component 3 Applied psychology*. Hodder Education Group, 07 2016.

[36] P. Bolton, M. Mira, P. Kennedy, and L. MM., "The quality of communication between hospitals and general practitioners: An assessment," *Journal of Quality in Clinical Practic*, vol. 18, no. 4, pp. 241–247, December 1998.

[37] R. B. Ashbrook and F. B. Doolittle, "The communication system of the southern california edison company, ltd." *Transactions of the American Institute of Electrical Engineers*, vol. 50, no. 1, pp. 89–99, March 1931.

[38] S. Möller and A. Raake, *Quality of Experience Advanced Concepts, Applications and Methods*. Springer, 2014.

[39] S. G. Liu LA, Chua CH, "Quality of communication experience: definition, measurement, and implications for intercultural negotiations," *Journal of Applied Psychology*, vol. 95, no. 3, pp. 469–487, May 1999.

[40] M. Garau, M. Slater, S. Bee, and M. A. Sasse, "The impact of eye gaze on communication using humanoid avatars," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '01. ACM, 2001, pp. 309–316.

[41] R. Gifford, *Environmental psychology. In Encyclopedia of human behavior (2nd ed.)*. Elsevier, 2012.

[42] K. Higuchi, Y. Chen, P. A. Chou, Z. Zhang, and Z. Liu, "Immerseboard: Immersive telepresence experience using a digital whiteboard," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 2383–2392.