Are QoE requirements for Multimedia Services different for men and women? Analysis of Gender Differences in Forming QoE in Virtual Acoustic Environments

Mansoor Hyder, Khalil ur Rehman Laghari, Noel Crespi, and Christian Hoene

[1] Interactive Communication Systems (ICS),University of Tuebingen
Sand 13, 72076, Germany
[2] Institut Telecom SudParis Evry France
{mansoor.hyder,hoene@uni-tuebingen.de,
khalil.laghari,noel.crespi@it-sudparis.eu}

Abstract. In recent years, the Quality of Experience (QoE) notion has become a major research theme within the telecommunication community. QoE provides an assessment around a human perception, feeling, performance and behavior. Normally, research studies on multimedia, Quality of Experience (QoE) and gender differences are carried out separately. To develop multimedia solutions covering QoE aspects, keeping the gender differences in mind, are the need of time. In current study, we are interested in knowing how QoE is shaped in virtual acoustic environment for male and female. We present experimental test results which provide interesting findings that both male and female have difference in their performance and perception in locating concurrent talkers in small and big sized virtual conferencing rooms. The middle-sized virtual room was suitable for teleconferencing because both male and female participants' performance and perception converge to similar trend and obtained better QoE values.

Keywords: Multimedia, 3D Audio, Quality of Experience, Virtual Acoustic Environment, Gender Differences

1 Introduction

Quality of Experience (QoE) is based on study of social psychology, cognitive science, economics and engineering science to assess overall human quality requirements, expectations, feelings, perceptions and congnitions with respect to a particular product, service and application [1, 2]. In short, QoE is the blueprint of all human quality needs and requirements. There is burgeoning trend to assess the quality of multimedia services and products on the basis of user centric Quality of Experience (QoE) benchmarks. Normally, QoE requirements are considered similar to every user of a particular multimedia service. But in this competitative market, it is important to provide more personalized and differentiated experience to each individual user/customer. Further, the previous research has also revealed that there are gender differences in perceptions and behaviours as well [3, 4]. Additionally, this is evident that male and female communicate differently in their role as a customer. For instance, females prefer accommodating communication which invloves listening and understanding the customer's needs during service interactions [5]. However, the males prefer specific and logical information because they are task oriented and process information selectively [6].

Traditionally, Quality of Experience (QoE) and gender difference research have been carried out separately or in an isolated manner. Their exact correlation consequently remains unclear. With this current study, we intend to investigate the relationship among multimedia service (for which we have selected 3D audio based on virtual acoustic environment), gender difference and QoE. The current experimental study is conducted to bridge the gap that exist among multimedia, QoE and gender differences. We evaluate 3D audio telephony and teleconferencing service on the basis of QoE. Virtual acoustic environment, which is part of 3D Telephony and teleconferencing service [7], helps participants of a conference call to spatially separate each other, locate concurrent talkers in space and understand speech with clarity. We define two QoE factors in this study, Localization Performance LP and Localization Easiness LE for the evalaution of virtual acoustic environment. Localization Performance is an objective QoE factor and it is related to human cognitive capabilities. LP is defined as, an assessment of how correctly listeners could locate the positions of the concurrent talkers in virtual teleconferencing room. Localization Easiness is a subjective QoE factor. It represents human perception and feelings of easiness in locating talkers. LE is defined as, how easy listeners feel it to locate concurrent talkers in virtual acoustic environment?. Specifically, this study investigates the effect of gender differences on QoE based on the study of localization performance and localization easiness in three different virtual acoustic rooms having dimension of ($10m^3$), ($15m^3$) and ($20m^3$). Furthermore, it is also important to understand and analyze the difference between human perception and performance capabilities with respect to their interaction with the virtual acoustic environment of 3D Telephony.

The overall goals are:

- To determine which virtual acoustic environment is most appealing to which gender in the context of 3D audio supported telephony and conferencing calls
- To design a virtual acoustic room which enhances quality of experience of the conference call participants (both male and female)
- To know whether human perception matches with human performance

This paper is organized as follows: In section 2, we present related work. In section 3, we present experimentation methodology and overview to data analysis techniques employed. In section 4, we discuss the results and in section 5, we conclude our work.

2 Related Work

We see it as an important task to understand the possible differences in forming QoE based on gender because gender differences may reflect difference in feelings, behaviour, performance and attentional task. It is reported in [8] that there are gender differences in attentional mechanism to collect spatial information in multiple sound sources environment. They found that, in the task of sound localization in multi-source environment, the superiority of males was apparent in the auditory domain. Further, in [9] sex differences in hearing are reported, the average girl hears the same sound with greater sensitivity than the average boy. In [10], superior performance by women on a task requiring object location memory has challenged the traditional view that men excel on all spatial tasks. A significant departure from the expected findings on sex differences has been the discovery that women excel also on a test of location memory for objects [11]. There is not much related work available concerning gender differences, QoE and VAE. Conclusively, literature study suggests to properly understand the role of gender in forming QoE in multimedia environments.

3 Experimental Study

3.1 Methodology

In current user study we opted three cubic shaped virtual acoustic rooms of ($10m^3$), ($15m^3$) and ($20m^3$) dimensions. The walls of the rooms represented the typical acoustic properties of concrete. Five participants were positioned at center of the room around the round table equidistantly as shown in Fig.1. One out of five participants was always performing a role of listener within the virtual acoustic environment. Remaining four participants always played the role of a talker. The layout of the postions of the listener and four talkers is depicted in Fig. 1.

At any given time, two out of four talkers were talking simultaneously. Listener's job was to locate and report the positions of the two simultaneous talkers and give score on the MOS scales on the level of effort he/she required to locate the two simultaneous talkers. Within all three different virtual acoustic rooms, the table always represented the radius of two meters (2 m) . Further, a rendered view of one of the virtual acoustic room with two simultaneous talkers and a listener is presented in Fig.2. Where, the green lines represent reflections of speech, yellow lines represent first order reflections and white lines represent direct speech from a talker to the listener. Red dots represent reflections point within the virtual acoustic environment. Small green triangles represent talkers while the small blue triangle represent a listener.

Furthermore, 26 paid subjects (13 male and 13 female) participated in this user study. All subjects had a university level education background and normal hearing threshold. Every subject participated in this experimental study as the listener and reported his performance and perception separately after judging six set of two simultaneous talkers. Therefore, every subject/listener had to

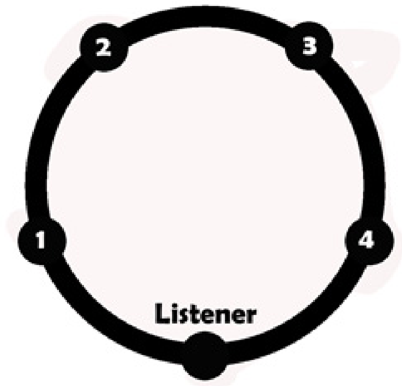

Fig. 1. *Layout of the listener and four talkers in the virtual acoustic room*

perform localization task and provide us with his/perception on 12 locations in one room. All in all, every subject contributed his/her responses for 36 locations of different combinations within all three virtual acoustic rooms where talkers were employed. Overall, 936 observations were obtained and analyzed in this experimental study.

Within the scope of this study, we were interested in knowing how listeners located multiple concurrent talkers in different virtual acoustic rooms. What was the successful vs unsuccessful performance of males and females in locating concurrent talkers in the virtual acoustic environment? To investigate the questions, we divided 26 participants (13 males and 13 females) into four groups by gender and number of talkers located. We considered a participant successful if he or she succussfully located at least 8 out of 12 talkers in one of virtual acoustic rooms and unsuccessful otherwise. We abstracted away information that distracted from our goal, such as specific positions located by the participants in a virtual acoustic room where talkers were employed. The virtual acoustic environment paramters, groups and number of participants are summarized in Table 1. Further, a view of the virtual acoustic environment under study can be seen in Fig.2.

3.2 Data Analysis Techniques

Localization Performance. To get data trends and first hand information, we use statistical approaches to measure user localization performance and corresponding confidence interval. The proportion probability p represents estimation of task performance rate. The most common way to measure a successful task completion rate and /or proportion of performance is to divide the number of participants who successfully completed the task (x) of localizing talkers in virtual environment by the number of participants who attempted the task (n) to

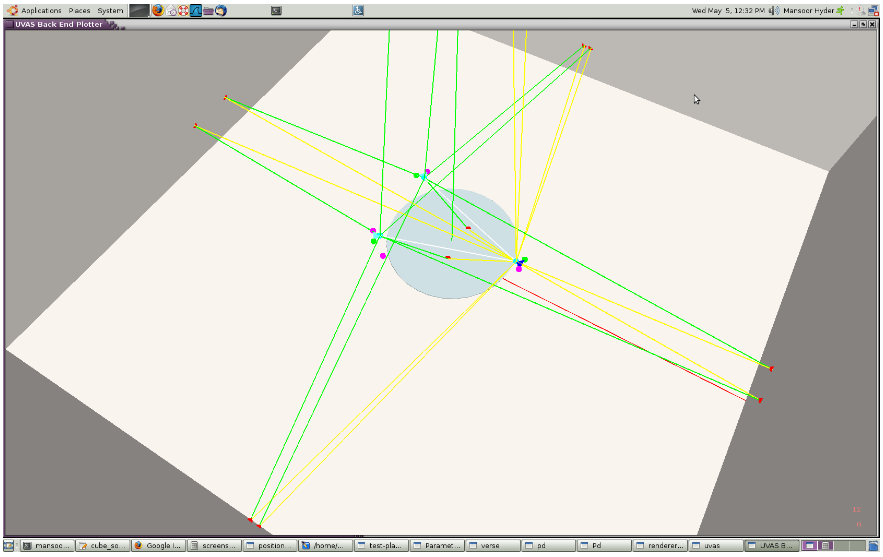

Fig. 2. *Virtual Acoustic Room with two talkers and one listener*

estimate p, the population probability of successful completion. The equation for general point estimator is given as below. p=((x+c^2/2))/$((n+c))(i)$

Other point estimators are special cases of general point estimator which are given below.

When c=0, it becomes Maximum Likelihood Estimator (MLE); p=x/n (ii)

When c=1, it becomes Jeffreys point Estimator; p=(x+0.5)/(n+1) (iii)

When c=2, it becomes LaPlace; p=(x+1)/(n+2) (iv)

When c=2, it becomes Wilsons point Estimator; p=(x+2)/(n+4) (v)

For the estimation of task performance, the Wilson point estimator is recommended, if proportion of success (x/n) is less than 0.5. Maximum Likelihood Estimation (MLE) is used, if the proportion of success(x/n) lies between 0.5 to 0.9 and LaPlace method may be used when proportion of success (x/n) is greater than 0.9 [13]. These estimation techniques produce statistically significant result about user localization performance. For calculating p, it is also important to calculate Confidence Interval (CI). CI is used to indicate the reliability of an observed data by a certain confidence level. The confidence interval is double of the margin of error and it tells us the likely range the population means and proportion will fall in. There are many techniques to calculate confidence interval. In [12], they present methods (Wald, Adjusted Wald, Clopper Pearson Exact, and Score) to compute CI. They found that the Adjusted Wald technique is suitable technique to calculate error margin and confidence interval. We calculate localization performance based on various point estimates, and CI will be computed using Adjusted Wald method.

Virtual Acoustic Environement			Groups	Number of Participants
Room Size	20 m³	⇒	Successful Males	6
		⇒	Unsuccessful Males	7
		⇒	Successful Females	8
		⇒	Unsuccessful Females	5
	15 m³	⇒	Successful Males	10
		⇒	Unsuccessful Males	3
		⇒	Successful Females	9
		⇒	Unsuccessful Females	4
	10 m³	⇒	Successful Males	11
		⇒	Unsuccessful Males	2
		⇒	Successful Females	8
		⇒	Unsuccessful Females	5

Virtual Acoustic Environment ⇒ **Successful and Unsuccessful Gender Groups of Test Participants**

Table 1. *Virtual acoustic enivronment and successful and unsuccessful gender groups of test participants*

Localization Easiness. Localization easiness is obtained using Mean Opinion Score (MOS). Within our experimental setup, subjects were asked to report the level of effort they required to localize two simultaneous talkers at any given time during the test. These MOS easiness values were measured on the discrete scale from 1 (bad), 2 (poor), 3 (fair), 4 (good) to 5 (excellent).

4 Results and Discussion

4.1 Localization Performance

Based on the discussion in (refer Section 3.2), we employed adjusted Wald method to calculate Confidence Interval (CI) and point estimators to calculate performance proportion p. Further, virtual acoustic enivronment and successful and unsuccessful gender groups of test participants are summarized in Table 1. We classify data into successful and unsuccessful gender groups of test participants with respect to each virtual room. From table, we get the values for x (successful participants) and n (total numbers of participants). We use point estimators to calculate localization performance proportion p and adjusted Wald to compute CI. The results of computation are presented in Table 2

Male. As per the results presented in Table in Table 2; localization performance increases for male participants as the size of virtual room decreases. The overall trend suggests indirect relationship between localization performance rate and virtual room size. It means male participants successfully localized more concurrent talkers in small-sized room ($10m^3$) than big-sized room ($20m^3$).

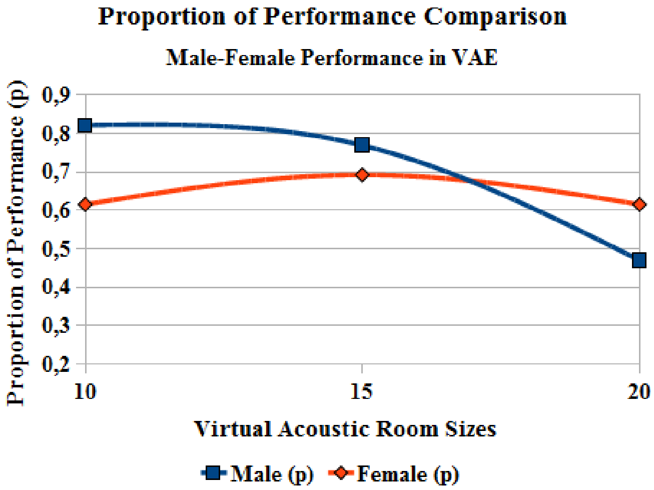

Fig. 3. *Comparison of proportion of performance (p) for male and female test participants*

Female. Unlike male participants data, female performance proportion rate is same (0.6154) in both large room ($20m^3$) and small room ($10m^3$). However the highest performance proportion rate is achieved in a middle size room ($15m^3$). It means female participants' performance is the highest in middle size room, and the small and big size room bring no big difference in their localization performance.

Comparison. It is quite clear from Table 2 and graph in Fig. 3 that both male and female differ in their localization performance capabilities.

4.2 Localization Easiness

Since, easiness measures a person's belief in his or her ability to perform a particular task [14], therefore, in this study, we were particularly interested to investigate whether quality scores for localization easiness of test participants play any clear role in performing localization of concurrent talkers in virtual acoustic environment. Further, distribution of quality scores on human localization easiness are reported in Fig. 4) along with the proportion of performance against different virtual acoustic environments.

The virtual acoustic environment paramters, proportion of performance of the participants and qualtiy scores for localization easiness are summarized in Table 2.

Male. Male participants perception of easiness is the highest in middle size room ($15m^3$) which is 3.85 MOS score and the lowest is in small size room

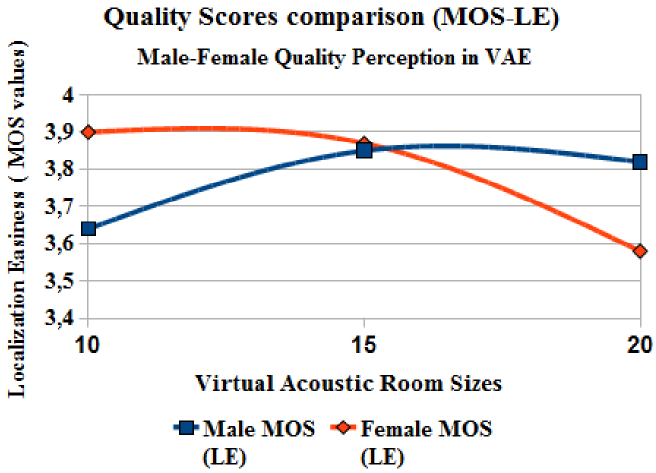

Fig. 4. *Quality score MOS-Localization Easiness comparison for male and female participants in virtual acoustic rooms*

($10m^3$) which is 3.64. It means male participants feel more easiness in localizing concurrent talkers in big room than in small room.

Female. For female participants, the MOS score data trend suggests that localization easiness and virtual acoustic room size are inversely proportional, i.e., as room size reduces, the perceived localization easiness scores increases. It means female participants feel easier in localizing concurrent talkers in small rooms than big room.

Comparison. The data trend in Table 2 and graph in Fig. 4 suggest that male participants feel more easiness in big room, conversely, female participants feel more easiness in localizing talkers in small room. While both male and female participants have similar scores in middle size room ($15m^3$). It means the male and female participants also keep different perceptual levels.

4.3 Localization Performance vs Localization Easiness

Male. In big-sized room ($20m^3$), male participants showed poor performance in successfully localizing talkers in various positions but they gave considerably good MOS score (3.82). While in small-sized room ($10m^3$), the overall MOS score i.e. 3.62, was lesser than big-sized room ($20m^3$). However, LP score was the highest (0.82). It means male participants perceive it easy to localize in ($20m^3$) room. But when male participants were asked to locate the talkers in ($20m^3$) room, their performance proportion rate was the lowest. At medium-sized room ($15m^3$), both LP and LE start to converge. It suffices to conclude

Virtual Acoustic Environement			Human QoE Factors			
		⇒	Proportion of Performance		Quality scores on Human Localization Easiness	
			Male (p) + CI	Female (p) + CI	Male	Female
Room Size	20 m³	⇒	0.4703 ± 0.2384	0.6154 ± 0.2350	3.82 ± 0.22	3.58 ± 0.27
	15 m³	⇒	0.7692 ± 0.2172	0.6923 ± 0.2280	3.85 ± 0.17	3.87 ± 0.15
	10 m³	⇒	0.8214 ± 0.2018	0.6154 ± 0.2350	3.64 ± 0.39	3.90 ± 0.31

Virtual Acoustic Environment ⇒ Analysis of Human QoE Factors in relation to Virtual Acoustic Environment

Table 2. *Analysis of Human QoE Factors in relation to Virtual Acoustic Environment*

that male perception and performance differ in both small and large room. But both LP and LE converge to similar trend in middle size room.

Female. LP values are similar in both big-sized room ($20m^3$) and small-sized room ($10m^3$), but LE-MOS scores differ in these rooms. In large room ($20m^3$), female participants perceive it harder to locate participants than small-sized room. In reality when they were asked to localize talkers in both large and small room. They performed equally well in both rooms. However in middle-sized room ($15m^3$), both LP and LE scores converge to similar trend.

Comparison. The results suggest that male and female participants have slightly different trends between performance rates and LE-MOS scores in small-sized ($10m^3$) and big-sized ($20m^3$) room but their perception and performance capabilities converge to similar trend in middle size room. It means the selection of proper virtual acoustic room size, and the gender of participants do matter in audio teleconferencing service. They should also be adjusted in suitable fashion to provide better QoE to both male and female participants/ end users.

4.4 Conclusion

The interest is increasing in both industry and academia to understand the QoE requirements of users and customers. In this paper, we investigated the impact of gender difference, and virtual acoustic enviornment characteristics over QoE. We learnt from this study that both localization performance and localization easiness are not only dependent on virtual room size but they also vary, when the gender of teleconferencing participants varies. Because when same subjects were tested in different virtual acoustic rooms, their performance and perception were different. Also this differnce is observed based on gender based groups. As per the findings from the user study, it is clear that male and female performance and perception varies in both big and small virtual acoustic rooms and we propose to consider the impact of gender differences over QoE. The middle-sized ($15m^3$)

room was found to be the most suitable room where human perceptual scores (LE MOS scores) and perfomance propotion rate (LP) match to similar trends. Hence, medium-sized room ($15m^3$) is suitable choice in terms of optimal QoE in virtual acoustic environment.

References

1. Laghari, K.U.R., Molina, B., Crespi, N., Palau, C.: "QoE aware Service Delivery in Distributed Environment", Advanced Information Networking and Applications Workshops, March 22 - 25,Biopolis, Singapore, (2011).
2. Kilkki, K.: "Quality of experience in communications ecosystem", Journal of universal computer science, pp. 615–624 (2008)
3. Iacobucci, D. and Ostrom, A., Gender differences in the impact of core and relational aspects of services on the evaluation of service encounters, Journal of Consumer Psychology, pp. 257–286 (1993)
4. McColl-Kennedy, J.R. and Daus, C.S. and Sparks, B.A., The role of gender in reactions to service failure and recovery, Journal of Service Research, pp. 66–82 (2003)
5. Sparks, BA and Callan, VJ, Communication in the service provider-customer relationship: the role of gender and communication strategy, Journal of Hospitality and Leisure Marketing, pp. 3–24 (1997)
6. Meyers-Levy, J., The influence of sex roles on judgment, Journal of consumer research, pp. 522–530 (1988)
7. Hyder, M., Haun, M., Hoene, C.: "Placing the participants of a spatial audio conferencecall", in IEEE Consumer Communications and Networking Conference - Multimedia Communication and Services (CCNC 2010), Las Vegas, USA, Jan. (2010).
8. Zuendorf, I.C. and Karnath, H.O. and Lewald, J., Male advantage in sound localization at cocktail parties, Cortex (2010)
9. Sax, L., Sex Differences in Hearing Implications for best practice in the classroom (2010)
10. Kimura, D., Sex, sexual orientation and sex hormones influence human cognitive function, Current Opinion in Neurobiology, pp. 259–263 (1996)
11. Silverman, I. and Eals, M., Sex differences in spatial abilities: Evolutionary theory and data, The Adapted Mind evolutionary psychology and the generation of culture, pp. 533–549 (1992)
12. Lewis, J.R. and Sauro, J., When 100
13. Sauro, J. and Lewis, J.R., Estimating completion rates from small samples using binomial confidence intervals: comparisons and recommendations, pp. 2100–2103 (2005)
14. Bandura, A., Social foundations of thought and action: A social cognitive theory, Prentice-Hall (Englewood Cliffs, N.J.) (1987)