A QoE Model for Digital Twin Systems in the Era of the Tactile Internet

by

Mohammad Alja'Afreh

A Thesis submitted to the

In partial fulfillment of the requirements

For the Ph.D degree in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

School of Electrical Engineering and Computer Science

Faculty of Engineering

University of Ottawa

uOttawa

Abstract

The idiom by Thomas Fuller fantasizes the fact that seeing is believing, but the feeling is the truth. This ideology has fired the vision and innovation of the Mulsemedia, multiple-sensorial media, and Internet of Skills (IoS) which enable the exchange of control, skills, and expertise anytime/everywhere across the Internet. With the emergence of the new generation of mobile network (5G), Tactile Internet, as well as the deployment of Industry 4.0 and Health 4.0, multimedia systems are moving towards immersed haptic enabled human-machine interaction systems such as the Digital Twin (DT). Specifically, Industry 4.0 will be using DT and robots on a large scale. This will increase human-machine and interaction to a great extent. There will be multimodal communications used to interact with digital twins and robots, specially haptics. Hence, tactile internet will replace the conventional internet today. In fact, a DT system can also be extended in Health 4.0 domain to act as a COVID-19 early warning system. Tracking a person's temperature and other symptom data in real-time can signal if as well as when it's time to see a doctor or take a COVID test. Link to a COVID tracing app, the digital twin might help get more information about the virus relative to the person itself. Since there are currently no well-recognized models to evaluate the performance of these systems, to address this research lacuna, we proposed a Quality-of-Experience (QoE) model for DT systems containing multi-levels of subjective, objective, and physiopsychological influencing factors. The model is itemized through a fully detailed taxonomy that deduces the perceived user's emotional and physical states during and after consuming spatial, temporal, proximal, and abstracted multi-modality media between humans and machines. Further, the taxonomy was modelled using the best practice of machine learning methods to show how QoE for digital twin applications can be inferred and predicted from interactions and biosignals in this class of applications. Furthermore, the taxonomy was applied to two use cases. The first one addresses the objective quality optimization for transmission in a large scale immersed haptic virtual reality over the Internet while the second one aims to objectively infer an important DT QoE physiological aspect i.e, fatigue.

Acknowledgements

I would like to sincerely show my deepest appreciation and respect to my supervisor and mentor Prof. Abdulmotaleb El Saddik for his continuous and invaluable guidance and support, not only in academia but also in my real life.

Many and sincere thanks to Dr. Mohammad Huda for his great assistance, guidance, and step-by-step support from day one of my PhD journey.

I also would like to extend my sincere thanks to all my research mates in the uOttawa Multimedia Communications Research Laboratory, Engineering Faculties at the Technical University of Munich (TUM), and University of Qatar, for their participation in the studies throughout this PhD thesis and journey.

Special thanks go Dr. Julia Kucharski and Dr. Emad Abdulkarim for their guidance, help, and never hesitating in answering all medical consultations from the first day until the last. In addition, particular credits go to my close friend John Baroud for his endless moral support. It is veritably a pleasure to deal with such respectful individuals.

Last but not least, I am grateful to have a wonderful family especially my dear mothers (Khaleda and Joan), Tara my lovely sister, and siblings in Canada and Jordan who always encourage me to excel in my life as well as stand at my side to let me tackle all hardships and difficulties during this journey.

THANK YOU ALL.

Dedication

To honor my beloved deceased wife "Krystal Schippers"

May God rest her soul peacefully in heaven.

Contents

List of Tables

List of Figures

List of Abbreviations

3D : Three-Dimensional

4G : The fourth generation of broadband cellular network technology

5G : The fifth generation technology standard for broadband cellular networks

ACR : Absolute Category Rating

AI : Artificial Intelligence

ANS : Autonomic Nervous System

AoI : Area of Interest

AR : Augmented Reality

AV : Audiovisual

CAHVE : Collaborated Haptic Virtual Environment

CNS : Central Nervous System

CPS :Cyber Physical System

DT : Digital Twin

DoF : Degree of Freedom

EDA : Electrodermal activity

EEG : Electroencephalogram

ECG : Electrocardiogram

EMG : Electromyography

IoT : Internet of things

IoS : Internet of Skills

LTE : Long-Term Evolution

HEVC :High Efficiency Video Coding

HD : High Definition

HMD : Head-Mounted Display

VR : Virtual-Reality

VE : Virtual Environments

HVR : Haptic Virtual Reality

HRV : Heart Rate Variability

FOV : Field-of-View

QoE : Quality-of-Experience

QoS : Quality-of-Service

HCI : Human-Computer Interaction

HAVE : Haptic-Audio-Visual Environments

ML : Machine Learning

MDF : Median Frequency

MNF : Mean Frequency

MEC : Mobile Edge Computing

MoS : Mean Opinion score

MVS : Multi-View Stereo

MR : Mixed Reality

MMT : Model-Mediated Teleoperation

NFV :Network Functions Virtualization

PNS : Peripheral Nervous System

PD : Perceptual deadband

RD : Research and Development

SDN : Software-defined Networking

TI : Tactile Internet

TDPA : Time Domain Passivity Approach

UX : User Experience

Chapter 1

Introduction

Currently, the Internet-of-Things (IoT) has been introduced as an umbrella to cover the extensions of the Internet and into the physical world, by means of the widespread deployment of spatially distributed devices with embedded identification and sensing capabilities. As stipulated in the ITU-T Y.2060 [9], "the IoT is an infrastructure that enables advanced services by interconnecting physical and virtual things, based on existing and evolving interoperable information and communication technologies". The fourth industrial revolution originates from the advent of IoTs. Through the advancement of information capture and sharing, processing and communication, systems may interact to deliver services not formerly attainable. Tangible objects, the things, are in the physical world and are efficient of being sensed and attached to others in its network. These things every so often have a virtual controller that enables the interaction and data exchange with sensory things in the environment. This infrastructure manifests itself in the domain of Industry 4.0 as a technology known as Cyber Physical Systems (CPS). CPS are defined in [10] as: "Systems where physical and software components are deeply intertwined, each operating on different special and temporal scales, exhibiting multiple and distinct behaviours, and interacting with each other in a myriad of ways that change with context." The essential attributes of CPS elements comprise 1) cybernetic capability in every component, 2) a high degree of automation, 3) networking at multiple

scales, 4) integration at multiple temporal and spatial scales, and 5) a modular dynamic capable of reorganization and reconfiguration. Consequently, the next wave of innovation demands enabling haptic interaction with audiovisual feedback, as well as technical systems supporting not just visual interaction, but also that involving robotic systems to be steered and controlled with an imperceptible time-lag. This next wave of innovation will create what is referred to as the the Tactile Internet (TI). The vision of digital twin is a digital replication of a living or non-living physical entity, proposed by El Saddik [1]. By bridging the physical and the virtual worlds, data are transmitted seamlessly, allowing the virtual entity to exist simultaneously with the physical entity. A digital twin is an immersed multimedia that enriches the segways to monitor, understand, and optimize the functions of the physical entity and provides continuous feedback to improve quality of life and well-being. A digital twin is hence Figure 1.1 the convergence of several technologies such as artificial intelligence (AI), Mixed reality (MR) and Haptics, IoT, Cybersecurity, Tactile Internet, Industry and health 4.0 and in particular Quality of experience .

A key component of the digital twin vision is the optimization of multi modal communication/interaction from the user perspective, which did not get any attention from industrial and academic stake holders until so far. As a result, the demands to conduct extensive researches and developments in this domain are still growing and infancy stages.

1.1 Motivations

While auditory and visual multimedia have reached an advanced quality level which is characteristically referred to as high definition (HD) and beyond. For example, high-end video cameras capture ultra-high-definition content, highly efficient video codecs such as H.265/HEVC achieve remarkable compression factors, and high-resolution monitors and virtual reality (VR) head-mounted displays (HMDs) enable high-end virtual experiences.

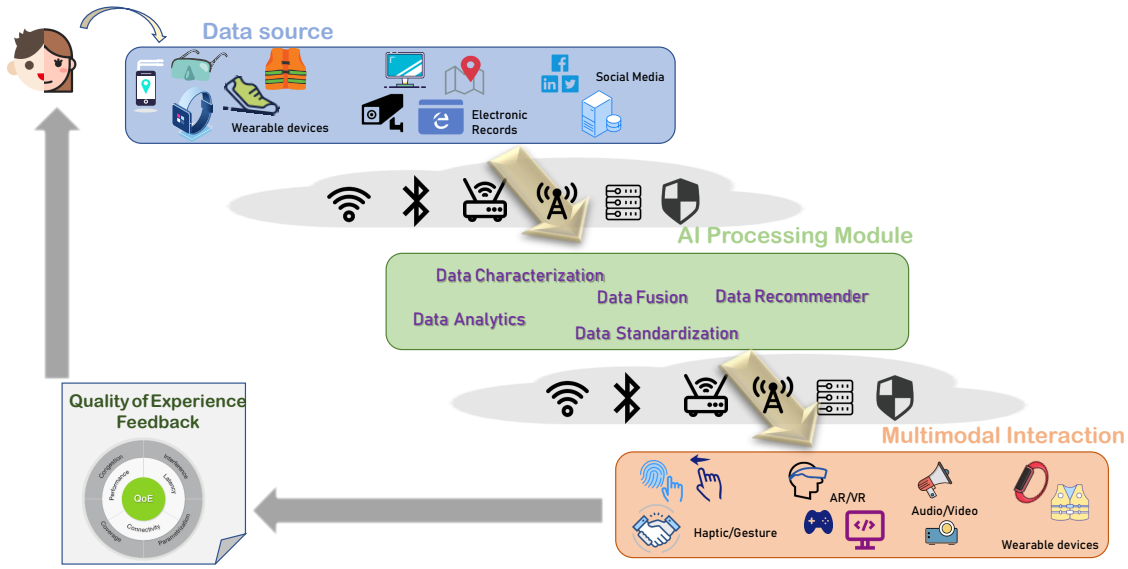


Figure 1.1: Digital Twin Technology adopted from [1]

Similar HD technology for audio is also available such as Hi-FI and Dolby Atmos. On the contrary, technical solutions addressing the sense of touch, which is typically referred to as haptic technology, have not yet got the same level of attention and evolution. With the development of the new mobile communication infrastructure (5G), the Internet will be revolutionized to what is referred to as The Tactile Internet [11]. The Tactile Internet is the next wave of inter-networking innovation that aims to provide ultra-low delay, ultra-high bandwidth, and ultra-high reliability communications. This certainly has enabled a paradigm shift from conventional content-oriented communication to control-oriented communication. Consequently, new technological solutions to address the new form of multimedia will significantly gain relevance. Enabling remote physical interaction with convincing touch, smell, and even taste experiences is one of the key technologies of the digital twin which be realized over the TI. Haptic, in particular, allows motor skills to be available across distances and enables fully immersive multi-sensory remote exploration of real or virtual environments. Therefore and with the incorporating of olfaction and

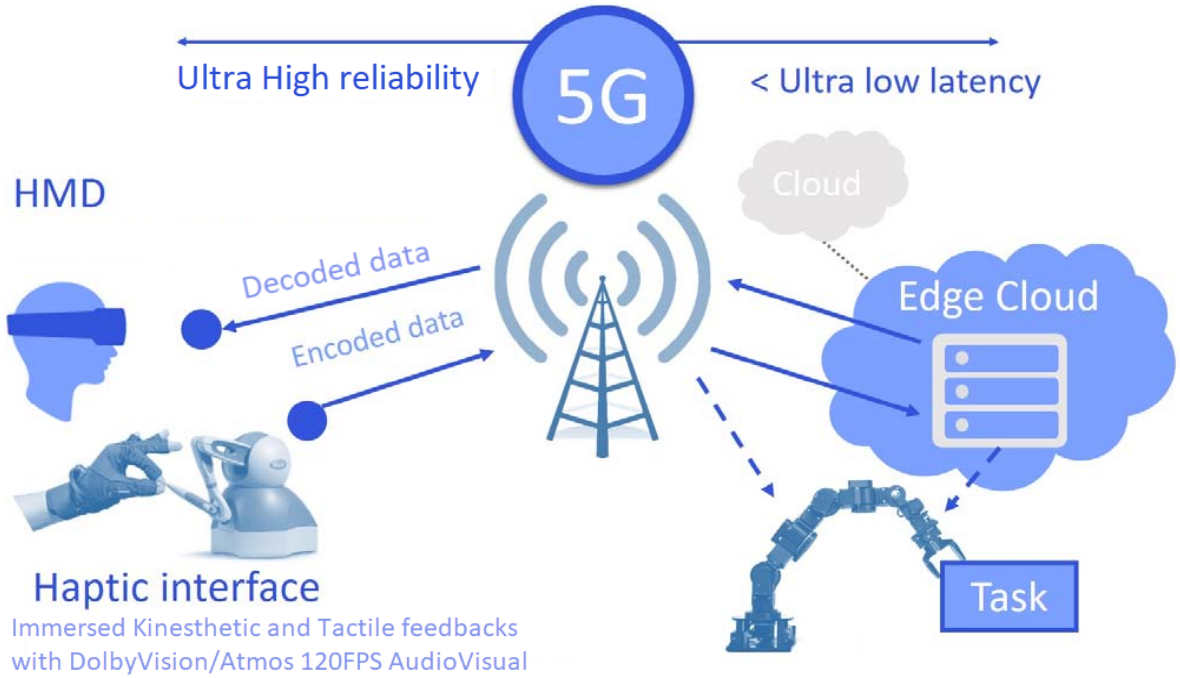


Figure 1.2: Industry 4.0 using Digital Twin (DT) Architecture over the Tactile internet

gustatory, users will not only see and hear remote Virtual/Augmented Reality (VR/AR) objects but they will be capable to smell, taste, feel and manipulate those objects. For the latter, haptic, olfaction and gustatory information need to be captured, compressed, transmitted and displayed with minimum latency.

For example, in the context of Industry 4.0, haptic more specifically teleportation is the main technology used to enable this field as depicted in Figure1.2. Teleoperation denotes the bidirectional ability to interact with a mechanical system remotely. Teleoperation is taken place by a machine or a robot at the slave site and is controlled by the worker at the master site placed at a remote location. The TI communication channel between the master and the slave is taken place via haptic feedback e.g. positions force, pressure, torque, vibrations, velocity, and touch and the delivery of ultra high-audiovisual resolution to the master site for proper interaction. Multiple merits will be gained from these virtual telehaptic interactions for instance the risk of injurers will be avoided as well as the time and cost to complete the industrial operations will be minimized. In

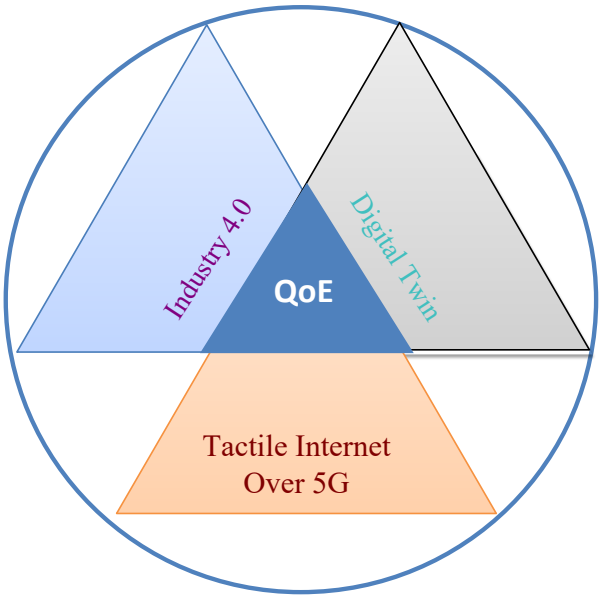


Figure 1.3: The technological intersection that is discussed in this dissertation.

this context, the Digital Twin (DT) architecture is a promising technology to enable teleoperation for an Industry 4.0 based system. The digital twin architecture requires the formation of a digital replica of some tangible person/object in a virtual environment, then bridging the gap between the two by a reliable network connection, enabling a real-time exchange of information between the twins. As an industry-based immersed multimedia system, a digital replica of a real-world industrial unit can be deployed in a virtual environment, then connected to the real unit over a network. Consequently, humans could interact with the virtual twin and any operations carried out within the virtual environment can be mirrored in the real-world twin.

As illustrated in Figure 1.3 optimizing on a large scale the transmission of the immersed multidimensional media as well as implementing a subjective and objective quality evaluation is the foremost focus of this thesis.

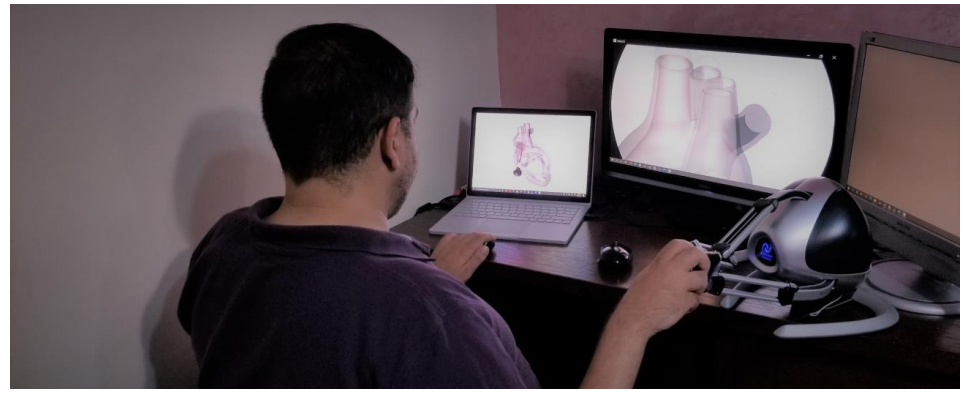
1.2 Problem Statement

Due to COVID 19 outbreak, most of the daily activities such as work, research, education are taken place online rather than in an inline style. According to [12] applications for teleworking and online education, including VPN and video conferencing, see traffic growth beyond 200%, since March 2020. Consequently, internet traffic has been increased drastically. As a matter of fact and according to a recent CNN news article [1], popular content providers such as Netflix and YouTube are slowing down in North America and Europe to keep the internet from breaking. The Covid 19 pandemic has altered how we live, breathe, and do business. It has brought the global economy to a halt, posing insurmountable continuity problems to any industry. With that being addressed, there is an increasing demand to enable digital twin technology and ultimately over the Tactile Internet.

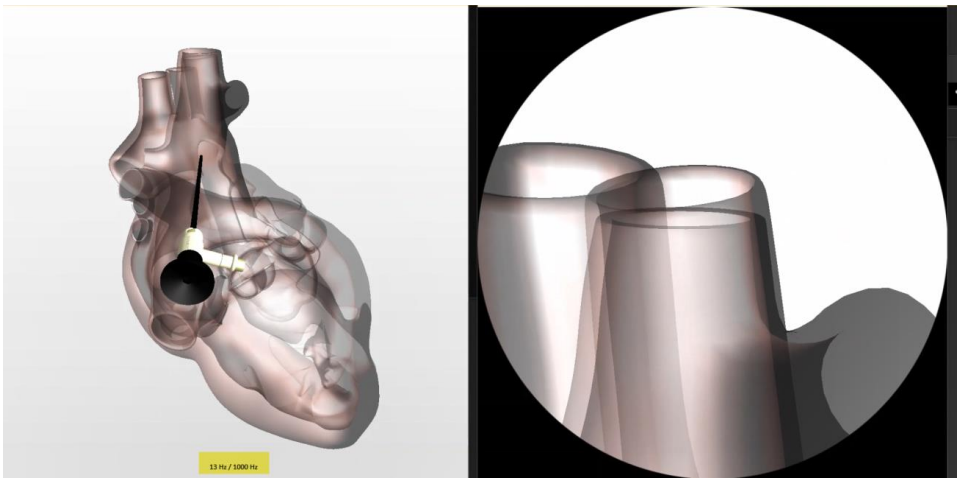
When stepping into the era of 5G, the emergence of the Tactile Internet (TI) can help haptic communication enter the new world for digital twins' applications. The most directly related is people's health and well-being in life. Up to now, people who live in some suburban areas still suffer from not getting high-quality and timely treatment. With the help of superb network conditions and the support from the haptic technical system, the patients whose physical conditions are not allowed to travel in long-distance to seek medical treatment can be diagnosed and treated at home or local hospital by some experienced doctors at remote places all over the world. One of its well-known applications is telesurgery, as illustrated in Figure 1.4.

After collecting and cloud computing real twins' information, its IoT data can help to reflect the physiological information on the digital twin. If doctors have different opinions about treatment, they can predict the response of the real twin's body by conducting experiments on digital twins without pain and damage on the real entities. A similar example is that we can use digital twins to clinically test drugs or surgical

[1] https://www.cnn.com/2020/03/19/tech/netflix-internet-overload-eu/index.html



(a)



(b)

Figure 1.4: (a) An example of Health 4.0 remote surgery using the DT model, (b) A proof of concept of Tele-heart-endoscope operation with 1KHZ haptic closed loop sampling rate.

treatment during the research and development testing phase, which also needs us to gather and deal with "big data" as well as use intelligence algorithms, deep learning, and machine learning (AI) to achieve physiological responses almost similar to the real twin. Obviously, minimizing the physical testing of drugs on the real body makes an effort to protect animals that will be used as experimental subjects. More humanitarian and less potential harm. What's more, the communication between haptic applications in the digital twins can not only resolve the health issues but also prevent people from physical damage. For instance, when executing tasks under some dangerous situations or in tough surroundings, haptic devices can display as humanoid robots to help experienced experts or technical engineers to handle hard issues in high-risk areas. At the execution end, digital twins perform image observation, data acquisition and implement real-time operation based on the performance of the remote real twins' control end. At the same time, the actual control unit can feel the feedback from the scene in real-time, including the tactile pressure, temperature, and pain in skin sensing, or feedback in audition, olfaction, and gustation. What's more, visual cues and sound must be synchronized with haptic interactions and must respond to continuous input data to increase the quality of perception. At the remote real twins' control end, we can apply the wearable devices to gather physiological data from real twins, while the devices can also bring an immersive experience to entities in terms of sight, touch, hearing, smell, taste, sense of direction, etc. In order to let haptic machines meet the speed of human's natural reaction time, 1ms end-to-end latency is necessary for Tactile Internet and it will need 5G as the underlying network infrastructure [13]. Tactile Internet along with IoT intelligence various aspects of human's life, particularly the haptic communication brings the immersive experience to achieve the human-in-loop meanwhile. The tactile internet will perform an innovative step by combing haptic interaction with visual feedback and other multimedia under the guarantee of its high availability, reliability, and low latency. In many cases, low tactile latency keeps the application safe, yet the current wireless and mobile communication networks have not achieved a round-trip latency in 1ms [11].

With the rapid development of fifth-generation (5G) mobile communications and infrastructure throughout the world, DT applications will become more popular. The DT is based on massive, real-time, real-twin data measurements across multiple dimensions. These measurements create an ever-evolving profile of the real twin (person, object, or process) in the digital world, that provides important insights on behaviour and performance which makes the interactions between the physical world and the virtual world seamless. as we mentioned, benefits of DT range from monitoring complex manufacturing processes, to analyzing individual people's behaviour using artificial intelligence and providing feedback, to predicting Quality of life features (QoL) such as possible health issues such as a heart attack or possible pandemic before it happens.

On one hand, evaluating these classes of applications using the traditional Quality-of-Service (QoS) models will fail to address the user's satisfaction with these applications due to the involvement of a large amount of human-computer interactions and the subjectivity of the end-user. On the other hand, existing Quality-of-Experience (QoE) models are proposed to meet specific targets related to conventional audiovisual multimedia i.e., QoE metrics differ from one system to another. In other words, many researchers will look at a semantic subjective mapping while others used mathematical modelling of QoE in terms of QoS. Nonetheless, Existing Quality-of-Experience (QoE) models of existing multimedia areas. Those models are tailored to fulfill specific purposes no dedicated multimedia devices, hence cannot be applied in the DT domain. For instance, it has been shown [14] that due to the bidirectional nature of positions (X, Y, Z) and its correspondent forces, torques, and velocity haptic signals and varying communication conditions, researching the human perception in haptic teleoperation systems is challenging compared to audio-visual perception [2]. Consequently, existing QoE models and approaches cannot be applied to DT applications where haptic and other emerging forms of multimedia (e.g, olfaction, gustatory, and other kinds of hard and soft social media) are heavily incorporated with the said system.

Consequently, the research lacuna in this work can be articulated as : *how can*

we quantitatively measure and predict users' Quality of Experience of DT system given the fuzzification of the human nature and the involvement of large amount of human-computer interactions ? As there are presently no well-established models for evaluating the QoE of a digital twin system, we propose a QoE taxonomy and model to best capture the amount of overall satisfaction and benefits earned from this system in addition to setting the foundation for the future ideal user-centric design of the aforementioned system.

1.3 Contributions

The main contribution in this thesis can be enunciated on how can we can help industrial, academical, and communication stakeholders building the digital twin systems, such as Industry 4.0 and Health 4.0 by putting in mind and proactively focusing on those QoE relevant metrics and parameters that guarantee an enhanced end-user perception while modeling, implementing, and fabricating the hardware, software, and communication complements of the DT systems. Overall the research contributions of this PhD work can be listed as:

- Conceptualization and modeling of the integration, communication and quality control properties of a generic Digital Twin System.
- Design and development of a QoE taxonomy containing multi-levels of QoE parameters for DT system. the taxonomy was built through analyzing existing QoE models of different multimodal systems in the literature and integrating technical metrics for new emerging media, observations, opinions, and more importantly bio feed-backs from our human subjects.
- Design and construction of a dataset that encloses the proposed QoE taxonomy
- We developed an optimized and automated QoE model based on effective machine learning algorithms to best capture the heterogeneous nature of multimodal sys-

tems such as the DT in a smart and cost effective way (incorperating subjective, objective, and physiopsychology mertics) that satisfies and predict the degree of perception of the most important entity of the multimedia system (i.e. the end-user).

- Design and development of an architecture that combines the three aspects: subjective, objective, and biomedical metrics.Hence, satisfying the pseudo-digital twin analogy.

- Benchmarking the interdependencies of 'QoE v QoS' over the traditional internet architecture such as LTE-A using different stabilizing, control, and QoS schems.

- Design and conduct experiment and user study based on the proposed taxonomy to objectively and subjectively estimate QoE for DT system taking fatigue as case study.

1.4 Scholarly Articles

The research works compiled in this dissertation have been reported or will be reported in the following publications.

- **Refereed journal Papers**:

    1. **Mohammed Al Ja'afreh**, Hikmat Adhami, Alaa Eddin Alchalabi, Mohamed Hoda and Abdulmotaleb El Saddik, "Toward Integrating Software Defined Networks with the Internet of Things: A review", *Springer Cluster Computing*, 2021.

    2. Abdulmotaleb El Saddik, Fedwa Laamarti, and **Mohammad alja'afreh.**, 2021. "The Potential of Digital Twins" *IEEE Instrumentation and Measurement Magazine*

3. Alowaidi, **M., Al-Ja'afreh, M.**, Karime, A. and El Saddik, A., 2019. "A Fuzzy Markup Language-Based Approach for a Quality of Location Inference as An Environmental Health Awareness" *International Journal of Extreme Automation and Connectivity in Healthcare (IJEACH)*, 1(2), pp.1-21.
4. **Al ja'Afreh. M**, A, Eid. M, El Saddik. A, Mahmoodi. T, Liu. Q, Liu. X, Steinbach. E , Strese. M, "Haptic Codecs for the Tactile Internet", Authors listed in alphabetical order, *In Proceedings of the IEEE (IF:9.237) special issue on Tactile Internet*, 2018.

- **Book Chapters and IEEE Standard:**

1. **Al Jaafreh M**, Alowaidi M, Al Osman H, El Saddik A. "Multimodal Systems, Experiences, and Communications: A Review Toward the Tactile Internet Vision". *Recent Trends in Computer Applications*, 2018 (pp. 191-220). Springer, Cham.
2. Adhami, Hikmat , **Al Ja'afreh, Mohammad**, and A. El Saddik, Abdulmotaleb, "Ontology based framework for Tactile Internet Applications" *"Smart Multimedia*, 2019 (pp. 81-86). Springer, Cham.
3. **Aljaafreh. M**, A. Hamam, A. El Saddik, "Objective and Subjective Quality Evaluation Metrics for the Tactile Internet Applications and Services", *case study contribution In IEEE P1918.1 Standard*, 2017.

- **Refereed Conference Papers**:

1. **M.Aljaafreh**, Somaya Al Maadeed, Jihad M. AlJA'AM3, and Abdulmotaleb El Saddik "Towards a Comprehensive Study of Fatigue Deducing Techniques for Evaluating the Quality of Experience of Haptic-Visual" in IEEE International Conference on Informatics Conference (Doha, Qatar), 2020.

2. H. Adhami , **M. Aljaafreh**, and A. El Saddik, "Ontology based framework for Tactile Internet Applications" in 2019 International Conference on SMART MULTIMEDIA, San-Diego USA.
3. H. Adhami , **M. Aljaafreh**, and A. El Saddik, " Can We Deploy Tactile Internet Applications overWi-Fi, 3G and WiMAX: a Comparative Study based on Riverbed Modeler" in 2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) (HAVE 2019), (Kuala Lampur, Malaysia), Oct. 2019.
4. **M. Aljaafreh**, H. Adhami, and A. El Saddik, "Experimental Qos Optimization For Haptic Communication Over The Tactile Internet," in 2018 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) (HAVE 2018), (Dalian, P.R. China), Sept. 2018.
5. Majed Abdullah Alowaidi , Ali Karime , **Mohammad Al Jaafreh**, Abdulmotaleb El Saddik, "Empirical Study of Noise and Air Quality Correlation Based on IoT Sensory Platform Approach" IEEE International Instrumentation and Measurement Technology Conference (I2MTC) 2018.
6. **M. A. Ja'afreh**, M. Aloqaily I. A. Ridhawi and N. Mostafa, "A hybrid-based 3D streaming framework for mobile devices over IoT environments," Third International Conference on Fog and Mobile Edge Computing (FMEC), Barcelona, Spain, pp. 211-216, 2018.
7. **M.Aljaafreh**, A. Hamam, A. El Saddik "A Framework to Analyze Fatigue for Haptic-based Tactile Internet Applications", IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE) 2017.

1.5 Thesis Organization

The rest of the dissertation is organized as follows:

- Chapter 2 presents the background information regarding the closely related state-of-the-art works fit the domain.
- Chapter 3 presents our proposed QoE taxonomy and the methodologies to deploy it.
- Chapter 4 presents a full projection and of the taxonomy into machine learning models to predict QoE for the developed DT applications.
- Chapter 5 presents a use-case on optimizing quality in the communication of immersive DT Haptic Virtual Reality over the existing internet using various modern networking technologies.
- Chapter 6 discusses a QoE Case Study about mathematically predicting and quantifying Fatigue for a haptic-audiovisual system.
- Chapter 7 presents the concluding remarks about the dissertation and discuss about the possible future works on the proposed topic.
- Appendix 8: shows the post-test questionnaire used in this thesis.

Chapter 2

Background and Related Works

In the past decade, multimedia has been widely based on vision and hearing. However, many forms of content especially those which stimulate other senses including olfaction, tactile, kinesthetic and gustatory have been recently proposed [15, 16, 4]. In addition to that, advances in IoT and Internet of Skills (IoS) technologies facilitate continuous data collection and interaction by means of connected hard sensors and multi-modal interaction. Hard sensors are sensors that are physical, hardware-based, such as pressure sensors, ambient light sensors, cameras. Hard sensors are part of what are referred to as today the Internet of Things and the Internet of Skills. Hard sensors can either be attached to the real twin. These include wearables and personal sensory devices, collecting mainly biometric data, and providing valuable insights on health and well-being. Hard sensors can also be ambient sensors, providing data on the real twin's environment at any given time. Additionally, continuous tracking involves data collected by soft sensors. Soft sensors collect data mainly from social networks such as Instagram, Twitter, Facebook, etc. We call them soft sensors because they are software-based, where information is entered into the platform by humans. Both IoT and Social Networks provide high amounts of sensory data and varied views on the real twin's state of health and well-being. Therefore, we are witnessing a rapid evolution of new technologies such as the digital twin that extends the audiovisual services into multimodal (i.e, multi-sensory) ap-

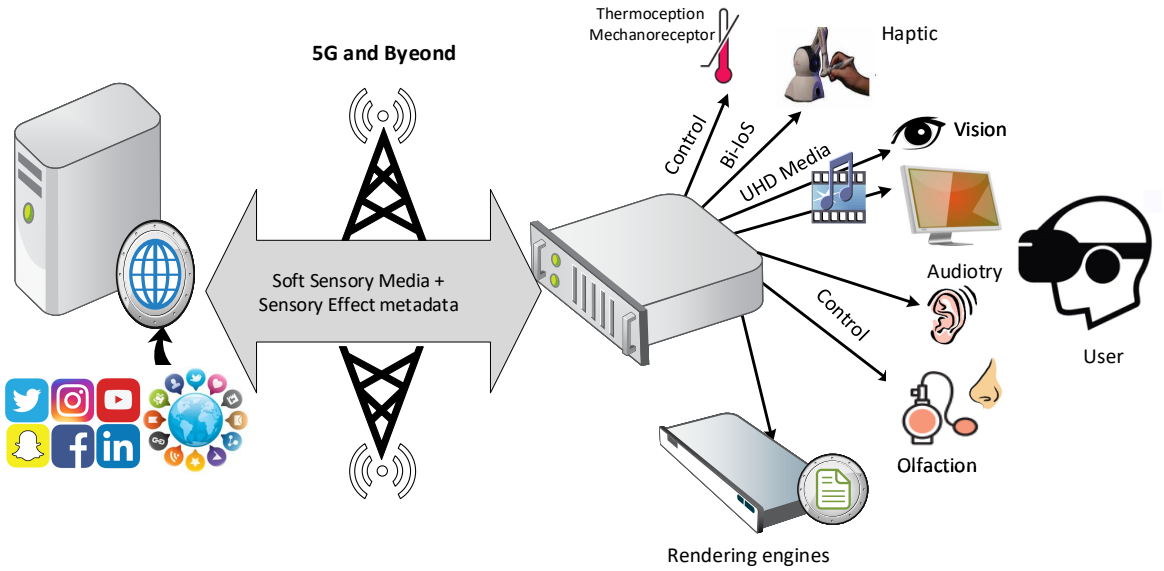


Figure 2.1: Hard and Soft Sensory Experience is Beyond AudioVisual Perception.

plications, as depicted in Figure 2.1.It should be noted that multimodal communications is the field referring to the: Representation, Storage, Retrieval, and Dissemination of temporally, spatially, and contextually correlated multisensory information expressed in multiple media such as: Text, Voice, Graphics, Images, Animations, Audio, Video, smell, touch, force feed-backs, and soft sensory data. These new modalities have been proven to potentially enrich the overall user's satisfaction in many applications ranging from interactive gaming to medical and military training [17, 18, 19]. In the following paragraphs, we provide a brief description of the haptic, gustatory, and olfaction modalities.

Haptic is referred to as the science of touch. The term is derived from the Greek word hapt esthai that refers to the sense of touch [16, 20]. The haptic perception was examined by many studies [16, 21, 22]. A haptic process can be considered as force, tactile vibration or heat stimuli that represents mechanical, thermal or pain receptor. Furthermore, haptic manipulation can be achieved as one of the following forms: tactile feedback, kinesthetic feedback, or motor action. Tactile signals are signals detected by the skin and internal organs such as touch, pain, vibration, heat etc. They are

coded by mechanoreceptors which respond to mechanical distortion of the sensory nerve membrane. Kinesthetic signals comprise information relevant for the human kinesthetic perceptual mechanisms, which allow for the perception of the position and orientation of our body parts and joints and external forces and torques applied to them. Hence, position, velocity, angular velocity, force, and torque all fall into the category of kinesthetic signals. In general, a haptic signal has a unique property of being both informatic and energetic [23, 24]. It is characterized by a bidirectional flow of information. Tele-haptics extends the haptic capability to include remote distance interactions [16]. Kinesthetic haptic interfaces (devices) are like small robots which enable a user to send and perceive kinesthetic signals. These interfaces are able to interact (modify) with a real or virtual environment and provide force or torque feedback to the user. A Touch X and Force dimension Omega.X are examples of a Kinesthetic haptic interface. The haptic interfacing device has to provide all the possible degrees of freedom (DoF) inherited from the biomechanical structure of a human's arms, joints, hands, etc. Nowadays, Sensible Touch X [25] device provides six DoFs positional sensing and three DoFs of force feedback which allow the human nervous system to interact haptically with the object in a very effective manner. As such, augmenting virtual environment and/or telepresence with haptic features implies that, *togetherness, tele-manupulation, and tele-touching* can be achieved among users in a realistic nature.

Olfaction is referred to as the act of smelling [26, 27, 28]. More specifically, it is the nose's ability to perceive a scent of material or substance in an environment. In spite of the fact that the olfaction modality can enrich the multimedia system, this field has been nearly neglected due to the limitations of hardware manufacturing required to remotely produce the desired scents [29]. However, many successful endeavours have lately succeeded in enhancing the quality of scent displays [30]. Furthermore, [26, 31, 32] conducted studies on olfactory displays and smell sensors to show the potential of transmitting scent information such as the aroma of fruits, cooked food, and flowers to users over networked virtual environments. In addition, the authors of [30] investigated

the current use of olfaction in many multimedia fields, from the filming industry to VR games [31].

The gustatory sense represents the taste sensation that can be perceived from a mixture of chemical signals [27, 4]. Gustation sensation involves a flavour of multi-faceted or a mixture of taste sensations. Flavour is defined by ISO [33] as different sensory integration modals of olfaction, gustatory, and trigeminal that are recognized during the tasting process. Generally, the taste is categorized in five basic qualities: sweet, sour, salt, bitter, and umami. Due to the complexity of chemical combinations required to produce distinct tastes and the fact that the gustatory sense is affected by other modalities like olfaction, vision or thermal, the research conducted on the perceived gustatory sense is still limited. Recently, [34] analyzed the diachronic and synchronic experiences of the five taste groups to establish a generic framework to design a taste experience human-computer interaction.

2.1 QoS requirements for each Modality

Traditionally, before the rapid emergence of the vast amount of multimedia content on the Internet, computer networks were designed to convey data traffic using a scheme known as best-effort delivery. This scheme was somewhat acceptable in the past, where data traffic was elastic (i.e., it can stretch under delay and bandwidth impairments). This best-effort service is unable to provide a predictable and reliable end-to-end packet delivery for real-time services and business mission-critical applications. The recent emergence of newly perceived multimedia classes of applications has highlighted the need for the development of a new standard that can accommodate their characteristics. This has been recently realized, in the form of a delivery scheme known as Quality of Service (QoS). QoS refers to a set of quality metrics used to tune and quantify the performance of applications, systems and networks. From a network point of view, QoS is defined as the set of service requirements that have to be met to enhance the overall utility of the network [35, 36].

This can be done by granting priority to a higher value or more performance-sensitive flows. Generally, QoS metrics are grouped into application factors, system factors, and network factors. Metrics in each group are bidirectionally affected by metrics in the other two groups. According to several research studies [37, 38, 39, 40, 41, 42], the most significant factors of the QoS metrics are considered: packet size, delay, jitter, throughput, data loss rate, rendering (update) rate, and arrival model.

The requirements for these five QoS factors in relation to the aforementioned media types are discussed in this section, and summarized in Table 2.1.

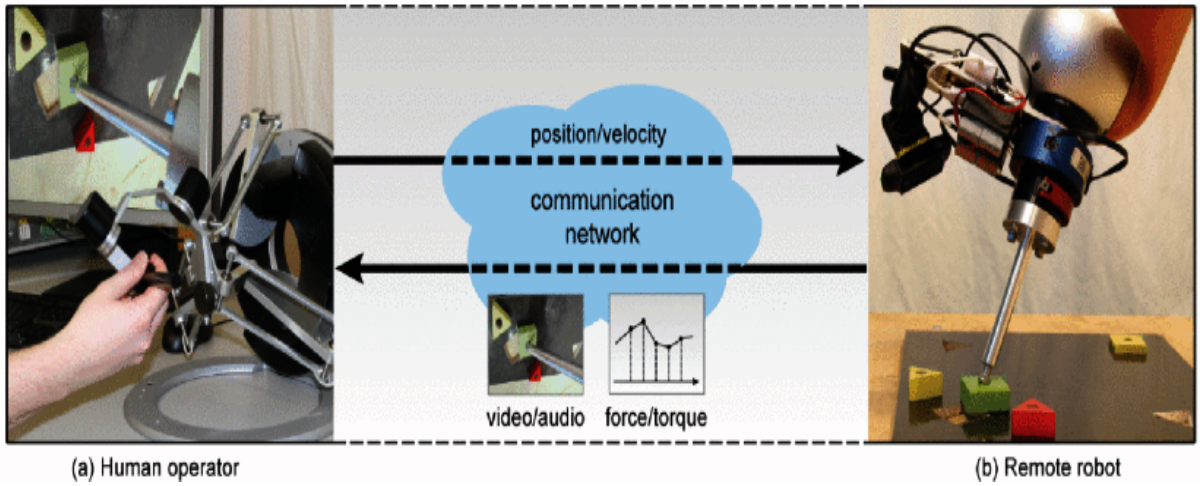


Figure 2.2: Bilateral teleoperation.The operator controls the position of the remote robot (teleoperator). Interaction forces are measured during contact and sent back to the operator. Additionally, visual and auditory information is streamed back to the operator.[2]

The internet-based haptic application enables the interaction between a human operator and a remote actuator. Such a system is referred to as a Master-Slave Teleoperation (MST) as illustrated in Figure 2.2. MSTs are measured by the number of Degree of Freedoms (DoFs) they provide [43] and can be used in several fields, such as military space robotics, underwater operations, and remote medical surgeries. DoF is one of the parameters used for characterizing haptic interfaces (both input and output). It defines the available number of degrees of freedom for position/orientation information (as in-

put), and force/torque feedback (as output). For example, 3 DoF position capability means that the device provides position information along x, y, and z axes. The haptics' bi-directional data flow property implies a distinguished requirement for teleoperation systems that maintains local control loop stability [44]. Unstable teleoperation interaction is a result of the unwanted delay, jitter, and packet loss. Haptic data is more sensitive to delay since the operation performance of kinesthetic or tactile feedback is typically positioned between 1 and 50 milliseconds. When The MST involves Machine to Mchine (M2M), the latency requirement is extremely critical with a range of 1-10 ms, in order to interact with fast-moving objects [45, 46, 22]. In MST systems, we might witness 2 scenarios, the master-slave, the sent data can be among others force, torque, or velocity (in 3 dimensions) sent in a 1kHZ rate, and the feedback from the slave which can be in the form of vibrotactile information or a kinesthetic with 1000 packet/per second. The 100 DoF tactile is denoted to include most of the body's biomechanical structure. Unstable latency (jitter) has the most adverse effect on the teleoperation process in terms of stability of performance. To ensure stability and transparency for MST, jitter has to be less than 2 milliseconds, and data loss has to be between 0.01% and 10% [47] [44]. To achieve a high fidelity output, the update rate should be greater than 1 kHz. This means that the haptic device should receive 1000 frame per second to render the output in a good resolution [48]. The bandwidth is another requirement of teleoperation communication, however, it has less of an impact since kinesthetic, tactile, and actuators have a small volume of data per frame (Usually 512 kbps) [47]. The arrival model of the haptic modality is heterogeneous by nature and can be described using the Gilbert-Elliot model with two-state Markov process [49]. In Summery, haptic systems are very critical in terms of delay, jitter, and update rate and much robust to data loss and bandwidth.

The video represents a sequence of continuous frames of visual data synchronized in a timely fashion [50, 51]. Video-conferencing and streaming are other examples of an application that cannot survive on the best effort standard for routing through an IP network [52]. One point worth noting is that in terms of video communication and com-

Table 2.1: Communication Requirements for DT modalities over Tactile Internet

QoS Metric	Human-to-machine					Machine -to- Human	
	Haptic	Video	Audio	3D	Olfaction	Haptic Feedbacks	
Packet Type	(Positions, Velocity, Force, Torque)	H.265/MPEG-4	Dolby Surround	Mesh, Texture	Scent	Kinesthetic Signals	Tactile Signals
Packet Size(B)	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48	1.5K	>50	1.5	1	1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48	1 DoF: 2-8 10 DoFs: 20-80 100 DoFs: 200-800
Jitter (ms)	1-2	30	30	30	≤ 23	2	1
Delay (ms)	1-50	≤ 400	≤ 150	100-300	0-1500	10	1
Throughput (kb/s)	≥ 512	≥ 2500	≥ 128	≥ 1200	0.008	≥ 512	≥ 1000
Data Loss Rate(%)	0.01-10	1	1	1-10	1	10	0.01
Update Rate(Hz)	≥ 1000	30	20	30	0.1-10	≥ 500	≥ 1000
Arrival Model	Heterogeneous (Periodic or Gilbert-Elliot)	Periodic	Periodic	Periodic	Periodic	Heterogeneous (Periodic or Gilbert-Elliot)	Heterogeneous (Periodic or Gilbert-Elliot)

pression of the MPEG-V standard, only three types of frames are taken into consideration [50, 51]: Intra-coded frames (I-frames) are used for coding real-time images frames that do not have references of previous frames, Predictive coded frames (P-frames) are used to reference the changes in previous I-frames or P-frames, and Bi-directional predictive frames (B-frames) are used as a data reference for previous and successive frames, so B-frames are aimed to enhance the quality of compression. To prevent video flickering, frames have to be refreshed momentarily. [53]. The refresh rate of video varies from one application to another, however, 30 frames per second is the common rate. In video applications, QoS communication requirements should be preserved within rigged time-based parameters in order to maintain temporal relation between information entities [47, 36]. A delay of 400 ms or less should be preserved to provide an acceptable human perception of real-time scenes. The jitter should be kept below 30 ms. Real-time video applications require a large volume of data per frame, therefore a high bandwidth provi-

sion may vary from 2.5 to 5 Mbps. In order to render a rich video application, the data loss should be constrained to around 1% and the update rate has to be around 30 frames per second [19]. Due to the manifested usage demands of high-resolution video content, higher dynamic range and higher frame refresh rate (60 fps) are significant to optimize video delivery [54].

Audio is defined as a continuous waveform of perceived voice/sound signal that propagates in a medium [50]. Voice Over IP (VoIP) applications such as Skype, Viper, and Hangouts are getting more popular for social multimedia users. VoIP applications transfer voice packets over an IP network while requiring guaranteed bandwidths, and very low delay and jitter. Audio delays expectations are based on the interaction level of the application i.e., one-way, two-way, or asymmetric two-way [55]. Generally speaking, the preferred range of the delay is less than 150 ms, whereas the jitter should stay around the 30 ms boundary. The audible frequency threshold that can be perceived by the human ear is roughly 20 HZ [56], hence the update rate has to be at least 20 frames per second. It has been shown [57] that the tolerated data loss of audio applications should be preserved below 1% and the throughput has to be more than 200 Kbps.

3D graphics is the representation of three dimensions of data in a geometric space. This space can be a composite of dimensions selected from the length, width, and depth. 3D plays a significant role in the creation and development of virtual reality (VR) and augmented reality (AR) applications [58]. Mixed reality (MR) is described as computer software that simulates a real or virtual physical existence where people can interact with it instantaneously and change it as if it is real. Augmented reality is the technique of enhancing the real world by getting information virtually to enhance a user's experience and senses. Multiview and stereoscopic scenes can enhance MR graphics which in turn will improve human perception [54]. MR streaming has been introduced to overcome the limitations of pre-installing and/or downloading the vast virtual environment's (VE) content. MR streaming is defined as continuous and real-time delivery of 3D contents, such as meshes, texture, and animations, over network connections to allow user inter-

actions with its virtual and physical world without a full download or a pre-installation [59, 60, 61]. Users can immediately render the 3D content when it is only partially received, and thus the interaction with their VE/AR occurs without having to wait for the entire download to be completed. MR streaming is similar to audiovisual streaming, where users can immediately interact with the displayed scene when the data are gradually downloaded. However, there exist some key differences between both approaches as illustrated in table 2.2. The first difference relates to the content itself, which is 3D mesh models, texture, and animation, real objects for the MR streaming and 2D images for audiovisual streaming. This means that the user in MR streaming can navigate the same scene at its all-possible resolutions and from any viewing angle without the need to download any extra data. However, in audiovisual streaming when a user wants to change his/her viewing angle, extra images for the same objects in the scene have to be streamed. In audiovisual streaming, the video is fragmented into sequentially ordered frames. These frames are transmitted according to the time sequence of the video. Hence, the data stream would be the same for everyone. The audiovisual streaming access pattern is considered linear and frame prediction can be applied; whereas in MR streaming, the 3D content cannot be ordered and has to be fragmented based on the user's behaviour, viewing angle and distance. Different viewing angles, areas of interest (AOI), or distances thus would produce unique transmission sequences. Obviously, in 3D streaming, the content access pattern is hard to be anticipated and lacks linearity[60]. The QoS factors for networked virtual environments have been evaluated in [62]. The findings are as follows: the end-to-end delay fluctuates between 100 ms and 300 ms, the jitter effect should be at most 30 ms, the 3D contents have to be periodically updated in a 30 frame per second rate, the data loss is equal or below 10%, and the bandwidth should be margined between 1.2 Mbps and 600 Mbps.

Tele-olfaction refers to the transmission of scents over a network [27, 63]. Users can sniff a mixture of odorants located at a remote place using an aromatic sensing system combined with olfactory display, as shown in Figure 2.3. Unlike other media types,

Table 2.2: MR Streaming Vs. Audiovisual Streaming

Steaming Characteristic	**MR**	**AV**
Content	3D meshes and textures	Image-based
Access Pattren	None-Linear	Linear
Transmission Order	Unique-based on user's (AOI)	Same frames for each user e.g., video streaming
Predictability	non-predictable	predictable

there is a fundamental challenge before realizing scents on demand. Given the fact that thousands of kinds of olfactory receptors are found in the human's nose [15], researchers find difficulties in creating a systematic, solid, and standard scheme for olfaction.

Although this topic is still relatively understudied, few works have been carried out to determine the tele-olfaction QoS metrics. In terms of virtual olfactory system, authors in [64] itemized the latency of any olfaction display in terms of the purging(cleaning) time for previous odour, the odorant forming time, and the time of delivery. An olfactory study [31] was conducted to assess the delay influence on a game when aromatic information is delivered to a player in a networked fruit harvesting game. The study found that to get a correct smell judgment from the player, the maximum allowable jitter has to be less than 23 ms. Even though a human can indicate the smell source (in respect to a median or sagittal plane) is about 0.1 ms [64], the perceived delay ranges between 0 and 1500 ms. Regarding the bandwidth needed to convey scents, a tele-olfactory system needs to rely on qualities of scents rather than their quantities as indicated in [30] where some scents have a different threshold than others. Also in [64], after carrying out a test on 60 odour classes, each one in a four concentration scale claimed that at most 8 bits of the capacity channel are needed to carry olfactory information. Besides the

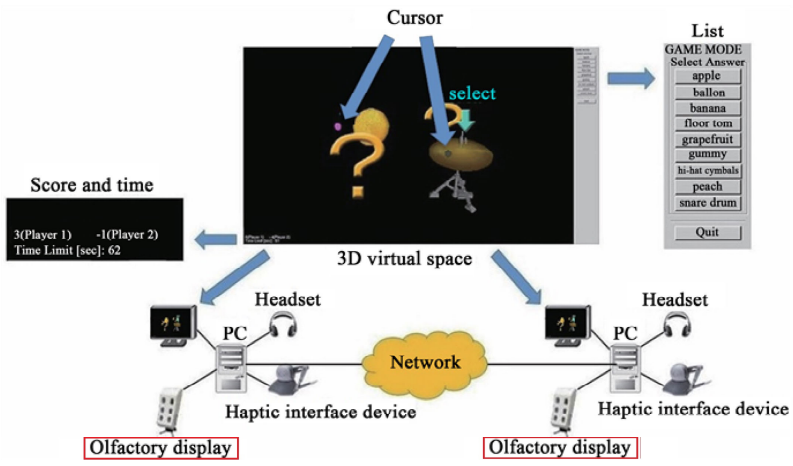


Figure 2.3: Experiment on Teleolfaction Using Odor Sensing System and Olfactory Display Synchronous with Visual Information.[3]

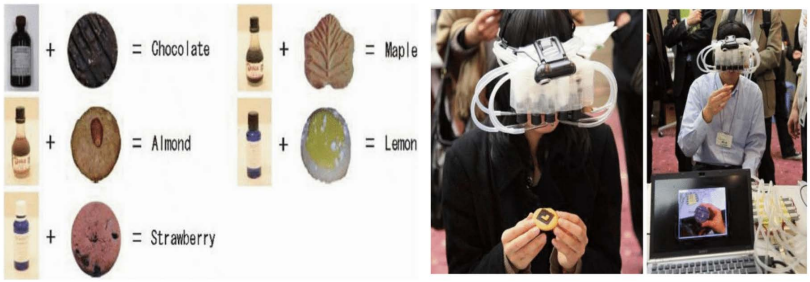


Figure 2.4: Flavor display based on cross-modal interaction among vision, olfaction, and gustation.[4]

aforementioned QoS metrics for olfaction, [64] performed a study that investigates the required refresh rate for a virtual olfactory application needed over the network. They found that the requirement for the update rate is 0.1-10 Hz and the shifting time between aromas between is 1 and 10 seconds.

As shown in figure 2.4 the recent implementation of the gustation is based on the cross modal interaction between VR and olfaction. Hence, gustatory QoS parameters are dependent on visual and olfaction communications factors [4, 27].

2.2 QoS Approaches for Immersed Media Traffic

A number of solutions have been proposed to deploy QoS on standard IP networks. These solutions include relative priority marking, service marking, label switching e.g., (Multiprotocol Label Switching MPLS), static per-hop classification, Integrated Services (IntServ) and Differentiated Services (DiffServ). The IntServ maintains a soft-state about the traffic that traverses the routers of the IP network. Resources allocation Protocol such as the RSVP which is a network layer signalling protocol used to guarantee QoS per individual application. The IntServ lacks scalability to large networks where multiple flows exist. On the contrary, DiffServ provides QoS to multiple streams by classifying, labelling, and shaping the data flows. Hence, it supports a multimodal application by aggregating different flows into a single class. DiffServ is implemented via the Type of Service (ToS in IPv4) and Traffic Class (TC in IPv6) fields in the IP header which specifies how the routers should treat each packet in a per-hop behaviour (PHB). The TOS/TC has 6 bits of the octet which supports up to 2^6 i.e., 64 different flows.

IETF DiffServ is designed to support multiple queuing and scheduling algorithms. The most widely used ones are First In First Out (FIFO), Priority Queuing (PQ), Custom Queuing (CQ), and Rate Queuing such as Weighted Fair Queuing (WFQ) Class-Based Weighted Fair Queuing (CBWFQ). FIFO is just what it sounds like, there is no priority traffic, no traffic classes, no queuing decision for the router to conduct. PQ has four types of queues: High, Medium, Normal, and Low. Packets in the high queue are sent before any packets in lower queues. PQ has to be configured carefully as it might not choke out traffic in the lower queues. Custom Queuing (CQ) takes PQ one step ahead, as it is based on a round-robin transmission approach and permits the allocation of the desired bandwidth for any and all traffic types. CQ has seventeen queues, however, only 16 queues can be configured. The queue indexed zero is reserved for storing system messages and carrying network control traffic such as keep-alive messages. Time consumption is the main drawback of CQ, as it has to be configured manually. To prevent one flow from

taking the whole bandwidth, weighted fair queuing was introduced. By assigning the proper weight, WFQ schedules delay-sensitive traffic to the front of the queue in order to minimize response time, and at the same time fairly shares the remaining bandwidth between high demanded bandwidth flows. CBWFQ is an advanced version of WFQ as it supports user-defined traffic classes. LLQ is an add-on providing a strict priority queueing feature to the regular CQ or CBWFQ.

The authors in [65] proposed configuration guidelines for IP DiffServ that includes multiple types of classes. The work induced a new real-time service class intended for interactive and variable rate inelastic applications that require low jitter and loss and very low delay. The authors recommended that this class of application has to be configured with a class selector CS4 while setting the Differentiated Services Code Point (DSCP) marking scheme with a PHB by Assured Forwarding AF21-23. The study in [66] investigated the effect of QoS required to transmit haptic stream on top of UDP in a Distributed Haptic Virtual Environment (DHVEs). The outcome of the study depicted that applying Weighted fair queuing and Class-based weighted fair queueing improved the transmission of a haptic stream in IP QoS-enabled architecture. The study was extended in [47] to study different levels of service forwarding applied to haptics and voice traffic. Comparing both experimental and simulation results, The author concluded that in order to have an acceptable haptic and audio experience, the haptic and audio CBWFQ classes have to be DSCP marked with at least AF22 or expedited forwarding (EF).

2.2.1 Software-defined Networking (SDN)

Deploying the traditional QoS protocols i.e., Interserv and Diffserv, is not an easy mission due to the fact that supervising and dynamic managing of resources necessitates the configuration (and re-configuration) of all network nodes is challenging. Consequently, flexible traffic management while concurrently establishing precise and dynamic QoS requirements is still crucial and challenging on the Internet. Alternatively, the effective and

rapid tuning of resources to the actual traffic necessity is one of the key features required to be effectively delivered by next-generation communication, which will additionally be a vital enabler for the Tactile Internet. A step forward in introducing flexibility and adaptability in network administration is achieved by the Software-defined Networking (SDN) [67]. Software-defined Networks (SDN) is characterized by "the decoupling of control and packet-forwarding planes in the network"[68]. It empowers the network to legitimately associate with applications through applying programming interfaces, supporting application execution and security, and creating a uniquely adaptable network design that can be changed as required. Apparently, evolving to be the most regularly utilized method for application deployment, SDN is now utilized by enterprises to send their applications more quickly, in the meantime cutting down deployment and operating expenses. IT heads utilizing SDN can oversee and provision their network services from an incorporated point. A network model that yields programmatic management, control, and network asset optimization, SDN applies open APIs to help keep up network control. This network control is established when SDN decouples the network design and traffic engineering, isolating them from their central network infrastructure. This splitting permits the utilization of OpenFlow and other open protocols. These open protocols can access network switches and routers that regularly utilize exclusive and generally closed firmware by applying globally aware software control at the network's edge. SDN helps clients virtualize their hardware and attempts to make a computer network by separating the system into the accompanying separate planes: The control plane offers the performance and fault management of NetFlow according to the type of the deployed protocols, hence it is often utilized for managing devices designs that are remotely associated with the software-defined network. The data plane advances traffic to its ideal destination. Before traffic arrives at the data plane, the control plane directs what path streams it will take by utilizing the flow control. This enables network administrators to effectively work with the software-defined network and manage the network. At the point when it was first deployed by huge enterprises, like Google and Amazon,

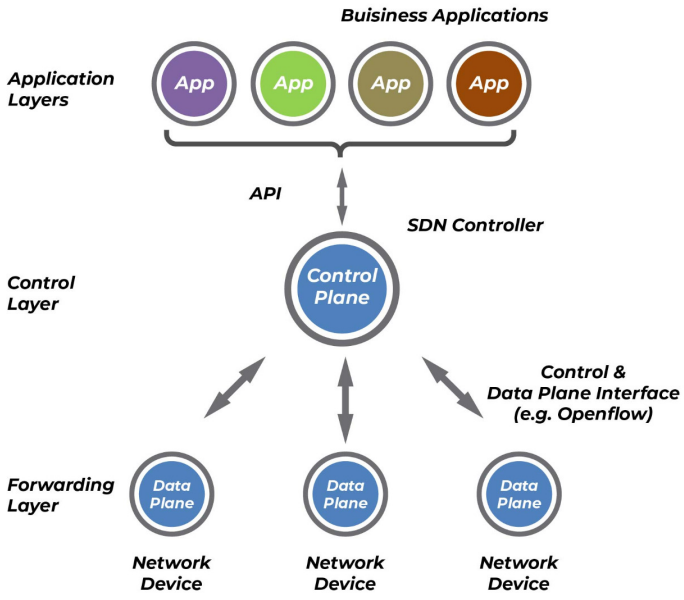


Figure 2.5: SDN Architecture

SDN helped them make adaptable server farms, encourage network resources and new cloud-based server development, as well as decrease the workload for IT directors. SDN streamlined the efficiency of the up-scaling procedure for these huge organizations and immediately drew the consideration of other huge organizations that quickly embraced SDN to improve their upscaling effectiveness. The framework of SDN is summarized and depicted in Figure 2.5.

2.2.2 Other QoS Schemes

In terms of QoS at the data link layer, Multiprotocol Label Switching (MPLS) is a routing technique in telecommunications networks that conveys data from one node to the next according to short path labels instead of long network addresses, hence dodging complex lookups in a routing table and speeding traffic flows [69]. As a result of that, MPLS can provide improved end-end QoS metrics for multimedia communications such as low

latency and jitter. In terms of QoS schemes at the application layer, the author in [70] proposed an Adaptive Multiplexing Framework (ADMUX) for haptic, visual, auditory, and scent data dissemination. To ensure QoS, Admux uses a statistical multiplexing scheme that fairly allocates the network resources. Adaptive Mulsemedia (multiple multimedia modalities) Delivery Solution (ADAMS) [71] is another client-server framework equipped with multiple modules that take into consideration information submitted by the client in order to refine the mulsemedia data streams to adapt according to the network condition. A visual-haptic multiplexing scheme is proposed by Cizmeci et al.[72] for teleoperation over constant bitrate (CBR) communication links. The proposed approach divides the shared channel into 1-ms resource buckets and controls the size of the transmitted video packets as a function of irregular haptic transmission events that are generated by a kinesthetic codec.

2.3 Quality of Experience

Although some applications may have optimized QoS parameters, users might still not be completely satisfied. For instance, users may still not be happy with an application as a result of rendering difficulties, entertainment limitations, subjectivity, or cybersickness (simulation dizziness). In fact, experimental results [73, 74] have shown that systems excelling in QoS do not necessarily translate to an enhancement of the perceived quality due to the gap between system and human-centric evaluations. According to [75], that gap is because the perception was traditionally treated as multiple modules acting independently. Therefore, the demand to find new methods for a collaborative perceptual quality assessment is increased accordingly [76]. A new model referred to as, Quality of Experience (QoE), has been proposed to tackle the issues not revealed by QoS measurements. Hence, Figure 2.6 depicts the bidirectional correlation between QoE and QoS. As a matter of fact, QoE is influenced by several factors. TThese factors can additionally be more generally categorized into three groups: Technical and system factors

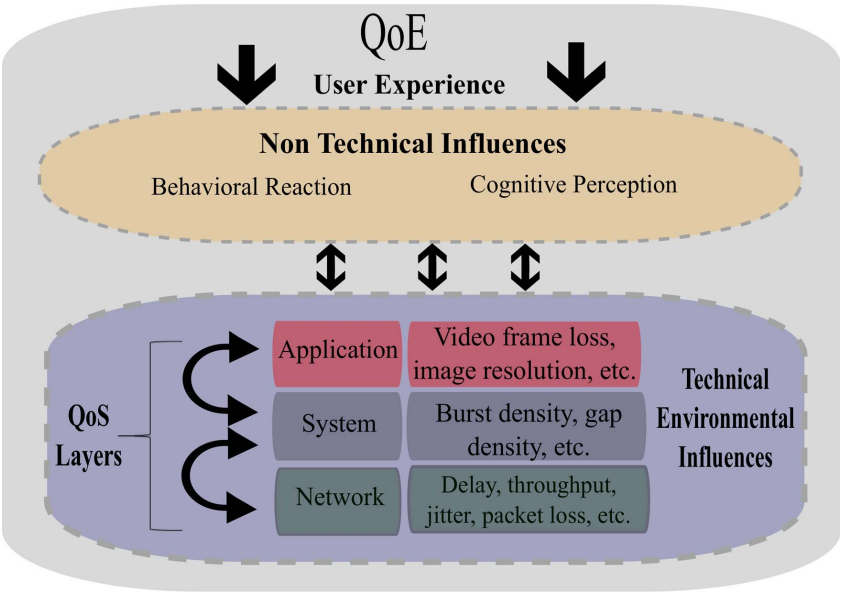


Figure 2.6: QoE and QoS dependencies

e.g., QoS layers metrics, human factors e.g., user experience (UX), and context factors. Measurements of these factors are differed between quantitative and qualitative performance indicators, yielding QoE a mixed measurement area of study. Unlike QoS, the QoE assessment system is still immature and little work has been done in this domain [77]. Throughout the literature, QoE has been defined in multiple ways and different contexts. In this thesis, we will list the most popular definitions of QoE.

"Quality of experience is the degree of delight or annoyance of a person during the process of experiencing. It results from the person's evaluation of the fulfillment of his or her expectations and needs concerning the utility (pragmatic and hedonic) in the light of the person's context, personality and current state"[78]. "QoE is not a metric but a concept comprising all the elements of a subscriber's perception of the network and its performance and how they meet expectations. QoS is the ability of the network to provide a service with an assured service level. QoE is how a user perceives the usability of service when in use – how satisfied he or she is with a service. The term QoE refers to the perception on the user about the quality of a particular service or network." [79] "QoE

can be defined as the qualitative measure of the daily experience the customer gets when he uses the services he is subscribed to including experiences such as outages, quality of the picture, speed of the high-speed Internet service, latency and delay, customer service, etc. The better the consumer's experience, the higher his QoE. And that affects customer loyalty."[80] All the aforementioned definitions share the following observations:

1. *QoE addresses the overall satisfaction of a given service, application, and network and can be affected by the context and the user expectations.*

2. *QoE assessment has to be carried out at the end-end level.*

3. *Subjectivity correlates strongly with the end-user judgment and appreciation.*

2.4 Requirements for the Design of QoE Evaluation Metric

QoE evaluation framework mainly falls into two categories. The first category, the user-based questionnaire, defines a subjective QoE evaluation metric at which the stakeholder will directly ask the users about their QoE evaluation. The questionnaire would determine how the users perceive the quality of the application and a subjective QoE metric could be derived. The second category describes an objective QoE metric, namely: Mathematical Modeling and the use of machine learning algorithms such as Decision Trees (DT), Support Vector Machines (SVM), and Fuzzy logic interface (FIS). In the following section, we will review the research works addressing these domains.

2.4.1 Subjective QoE Metric

Subjective assessments are taken place by a test panel of actual individuals and usually are formulated according to the International Telecommunication Union (ITU) standardized protocols and recommendations. To the best of our knowledge, there is no standard-

ized subjective framework for evaluating haptic(kinesthetic, vibrotactile) VR quality to date, but several protocols exist for the evaluation of audiovisual quality. Examples of the aforementioned protocols are ITU-T P.800; Rec. ITU-R BS.1116-3, Rec. ITU-R BS.1534-3 [81, 82, 83] used to formalize subjective testing for acoustic telephone transmission while the recommendations provided in ITU-T P.910 and ITU-T P.913 are used to propose the evaluation of the subjective metrics for video quality respectively[84, 85]. Mostly, three pilot tests were performed with three approaches as presented below.

- **Absolute Category Rating (ACR)**: As described in the ITU recommendations, ACR is an approach used to collected subjective experience for audiovisual stimulus. In this evaluation method, the end-users are presented with the original stimulus and after that a group of distorted stimuli of that pristine signal in an arbitrary order. The users perceive only one stimulus at a time and are then asked to share their experience about the perceived stimulus on a 5-point Likert scale as in Table 2.3 Questions are reported via group evaluation and the user might be

Table 2.3: Absolute Category Rating based on 5-point Likert Scale

1	**2**	**3**	**4**	**5**
Bad	*Poor*	*Fair*	*Good*	*Excellent*

uninformed of the numerical weighting of each group. Naturally, when the tests reach the third or fourth trial of impaired stimulus, users might have difficulties reporting scores since their remembrance of the original stimulus might fade up rapidly and previous test stimuli further disrupted the memory of original stimuli.

- **Degradation Category Rating (DCR)**: In this method, the users will perceive two stimuli of interest in pairs at the same time. The first stimulus of interest will be always the original i.e., reference, pristine and undistorted signal, while the following second one also referred to as the disrupted outcome will be randomly

selected from a group of distorted stimuli. Based on the preference of the participants, each pair can be perceived multiple times with a short period of time is used to distinguish between the outcome and reference stimuli. By doing so, a comparative analysis between the original and experimental stimulus will be achieved. Similar to ACR, users are asked to give scores to the experimental stimulus on a 5-point Likert scale as illustrated in Table 2.4 .

Table 2.4: Degradation Category Rating based on 5-point Likert Scale

1	**2**	**3**	**4**	**5**
Very Annoying	*Annoying*	*Slightly annoying*	*Perceptible but not annoying*	*Imperceptible*

- **Mean Opinion Score (MoS)**: In the literature, MOS is the traditional metric used to assess the overall user-perceived quality of the evaluated stimulus or system. Therefore, MOS is reported by the participants via the identical scale as the rating scale that the examiner selects for their assessment (ACR or DCR) thus it serves as a collective score for all reported answers to the evaluation of QoE. MOS is universally employed to define the overall user attitude of an audiovisual stimulus, nevertheless, it has been extended to incorporate several fields covered by QoE to give comprehension into a user's perception of innovative technology, e.g, digital twin, systems such as teleoperation.

Even though, the outcome of these subjective tests would include diverging opinions and insights on the quality of the perceived multimedia content result in precise measures and a good comprehension of the QoE and its parameters such as enjoyment, usability, and immersion. In addition, QoE in haptics has always been evaluated through subjective tests with the user-in-the-loop[22]. This type of assessment can be expensive, time-consuming, and recency biases as the tests have to be performed by a substantial sample size of participants for statistically appropriate and revealing results. Further, subjective

Table 2.5: Generic Mathematical Models of QoE-QoS

	Model	Name	Trend	Form	Equation
Inherited from Psychophysics	Authors in [87, 88]	Stevens' Power Law	Stimulus-centric	Power	$QoE = \alpha.QoS^{\beta}$
	Authors in [89, 90]	Weber-Fechner Law	Stimulus-centric	Logarithmic	$QoE = \alpha \log(QoS_i)$
Adopted from Hypothesis	Authors in [91]	IQX	Perception-centric	Exponential	$QoE = \alpha \cdot e^{-\beta \cdot QoS} + \gamma$
	Authors in [92]	Linear Regression	Perception-centric	Linear	$QoE \propto QoS_i$

QoE testing in haptics is expensive since customized haptic hardware makes it challenging to parallelize tests. Furthermore, since the test persons are typically new to haptics technology, extensive experimenter monitoring is needed. To tackle these issues, objective QoE testing is desirable.

2.4.2 Objective QoE Metric

The evaluation of QoE through objective testing is based on algorithmic and mathematical models of human perception and/or the measurement of several parameters related to service delivery. Mathematical Modelling represents a Multi-tier model that includes higher-level weights, categorical weights, and individual parameters weights. Different weight techniques exist, among them are principal component analysis, correlation, regression, and equal weight distribution. Nonetheless, most of the proposed QoE-QoS correlation models in the literature are service-oriented and cannot be comprehensively applied to all service classes [86]. However, and using Least Squares Regression (LSR) method and regardless of network service level, four well-known mathematical models endeavour to establish a generic relationship between QoE and QoS, Table 2.5. In the following subsections, a brief overview is provided to cover these models.

Stevens' Power Law

In 1957, an American psycho-physicist by the name of Stanley Smith Stevens proposed a modest, but effective model that relates the human perception changes to the intensity

of a physical stimulus [93]. The law can be represented by the following equation:

$$P\left(S\right) = K.S^{b} \tag{2.1}$$

Where P denotes the human perception degree as a function of the strength of the physical stimulus (S), K and b can be quantified using the least square method where K is a constant parameter that varies based on the conditions of measurement environment, while the exponent b denotes the kind of stimulus and concurrently identifies the magnitude of upward or downward curving of the power function. For instance, humans cannot strongly distinguish the increasing odour of garlic when the quantity of odour in the air surpasses a certain threshold. Hence, the value of b should be less than 1. Nevertheless, If we assume that QoE is represented by the human perception and QoS as a stimulus intensity that alter and impact that experience then Stevens' law can be reformatted in this context as below:

$$\frac{\partial QoE}{\partial QoE} \propto \frac{\partial QoS_i}{QoS_i} \tag{2.2}$$

$$\log(QoE) \propto \log(QoS_i) \tag{2.3}$$

Solving the differential equations 2.2 and 2.3 formulates a power relationship between QoE and QoS as illustrated in equation2.4:

$$QoE = \alpha.QoS^{\beta} \tag{2.4}$$

Where $\alpha > 0$ and β is the exponent that determines the curve's shape i.e., the impact of the QoS parameter on QoE. The relation's power form implies that for different QoS parameters, the growth of perceived quality magnitude may experience various speeds. The authors in [87, 88] suggest that it is feasible to take into consideration the power function to describe the possible relationship between QoE (MOS) and QoS (packet loss) in video streaming services.

Weber-Fechner Law

Ernst Heinrich Weber is a German psychologist who performed some studies on human perception with regards to different physical stimuli. The outcome of his studies is published as a law referred to as Weber-Fechner Law (WFL) [94]. WFL stipulates that the stimulus is considered perceivable and needs to be updated or conveyed only if the relative difference between two successive stimuli exceeds the just-noticeable differences (JND) threshold. Fechner. et.al [89, 90], envisioned a modern elaborate explanation of Weber's law and articulated it as a differential equation between physical stimulus strength and the magnitude of human perception. This correlation is described as QoE gradient is proportional to reciprocal QoS parameter and can be denoted as:

$$QoE \propto \log(QoS_i) \tag{2.5}$$

$$\partial QoE \propto \frac{\partial QoS_i}{QoS_i} \tag{2.6}$$

if we integrate both side at equation 2.6, then we will come up with formula describing a logarithmic or partial logarithmic correlation between QoE and QoS, as shown in 2.7.

$$QoE = \alpha \log(QoS_i) \tag{2.7}$$

Where α is a constant. It should be noted that the logarithmic nature between QoE and QoS is proven through an experiment at which the speech quality is assessed by Mean Score Opinion (MOS) with a logarithmic function of bit rate or loss rate in Voice over IP (VoIP) service [89, 90].

IQX hypothesis (exponential interdependency of QoE and QoS)

authors in [91] proposed the IQX Hypothesis Model to outperform the logarithmic function described in [90] that links between the physical stimuli and human perception. In which, a natural and generic exponential function 2.8 2.9 is used to quantify the relationship between QoS technical parameters (packet loss, jitter, response and download

times) and the QoE factors in terms of MoS findings.

$$\frac{\partial QoE}{\partial QoS} \sim -\left(QoE - \gamma\right) \tag{2.8}$$

$$QoE = \alpha \cdot e^{-\beta \cdot QoS} + \gamma \tag{2.9}$$

Where:

$$\alpha, \beta, \gamma > 0.$$

The IQX Hypothesis Model has produced a 0.993 correlation coefficient which indicates a desirable matching between the ground truth data and the applied exponential model. This model was only tested on two different types of applications (VoIP and web browsing) therefore we do not know it's effectiveness on environments that have complex combinations of modalities. Further, the model was only evaluated using the (NIST Net) simulator which has some difficulties to capture the real IP network behaviour [95]

Linear QoE-QoS relationship

A linear mathematical model between QoE and QoS implies that the magnitude of the user perception depends on the linear scale e.g.,(increase in throughput) or (decrease in latency), of the QoS parameter. In other words, QoE gradient independent of QoE and QoS as illustrated in 2.10

$$QoE \propto QoS_i \tag{2.10}$$

The work in [92] came up with a linear formula that validates equation 2.10, hence linearly correlates QoE in terms of Opinion Score and QoS in term of packet loss ratio. It is worth noting that linear model implies that the QoE is nearly equal to the QoS, despite the fact that in several situations, efficient service is the first step toward achieving it. While a service may be operating admirably, it may not be effective enough to fulfil or meet the user's needs [96].

Linear Multiple Regression

It should be noted that the aforementioned generic mathematical models use the least-squares approach for a straightforward regression. To put it another way, they look at the effect of a single QoS parameter on QoE. Other works such as [97, 98], on the other hand, take into consideration a variety of QoS parameters,. The authors of these papers describe the relationship between QoE and QoS using linear multiple regression based on the least-squares approach. In this regard, the author in [97] suggested a basic function calculating the QoE (MOS) from a collection of QoS parameters i.e., (packet loss rate, frame rate, bandwidth, round trip time and jitter) in the context of video streaming management with heterogeneous access technologies. In addition, based on the IQX hypothesis, the authors in [98] proposed a comprehensive correlation model between QoE and multiple QoS parameters. The model was tested for Video on Demand (VoD) systems, at which video quality is represented by Continuous Scale Opinion Score (OSCS) as an exponential formula of latency and packet loss ratio.

2.4.3 Objective QoE in Collaborated Haptic VR

Objective quality assessment (QA) for haptic VR communication is still in its infancy. Despite the existing attempts, there remains significant room for improvement. One next step could be the creation of accurate QA models that reflect the requirements of the haptic domain and could generalize as a building block in the design of successful haptic QoE. In order to effectively develop a haptic QA model, it is imperative to find a large set of parameters that affect the perceptual quality of the displayed content by exploiting three types of knowledge. First, the source of the haptic signal which describes how the signal looks and feels when displayed on different haptic interface devices has to be evaluated. The second is the knowledge about the distortion from the communication system point of view. Finally, since humans are in-the-loop and also are the ultimate consumers of haptic applications, the third type of knowledge is about

human haptic perception originated from physiology and psychophysical studies. We believe that models derived from these three domains should be adequate to develop successful haptic QA algorithms. So far, only very few studies for QoE evaluation for haptic communications are available in the literature. They can be categorized into two groups based on how the quality is predicted.

Signal-Level Quality Prediction:

The first work in this research line was introduced in [99]. In this work, a haptic perceptually weighted peak signal-to-noise ratio (HPW-PSNR) was derived to account for the perceptual significance of haptic signal degradation using the just noticeable difference (JND). The mathematical formulation is described as follows:

$$\text{HPW} - \text{PSNR} = 10 \cdot \log_{10}\left(\frac{||v_{\max} - v_{\min}||^2}{\text{MSE}\cdot\text{HPW}}\right)$$

$$\text{HPW} = \begin{cases} C, & \text{if\$|v}-\widehat{v}| \leq \text{JND(v)\$} \\ k \cdot (|v - \hat{v}| - \text{JND}(v)) + C, & \text{otherwise} \end{cases}$$

where v_{min} and v_{max} are the maximum and minimum values of the haptic original signal v . C is a constant term that weights the signal degradations below the perceptual threshold. k is a penalty factor that weights the haptic degradations beyond the JND of the signal. JND (v)= a_v x abs(v), with a_v being the percentage of the tolerable degradation of signal values.

Another work in [100] proposed a quality prediction framework for the compression of kinesthetic signals for closed-loop teleoperation. However, the proposed approach is only able to qualitatively predict user ratings. In [43], the perceptual mean squared error metric (P-MSE) is introduced. A perceptual comparison of a compressed haptic signal relative to the uncompressed one is made based on the Weber–Fechner law [101] which relates the psychophysical sensation S with the magnitude of physical stimulus x

as follows:

$$pmse = \frac{1}{N}\sum_{i=0}^{N-1}(S(i) - \hat{S}(i))^2 = \frac{c^2}{N} \cdot \sum_{i=0}^{N-1}\left(ln(\frac{x(i)}{\hat{x}(i)})\right)^2 \tag{2.11}$$

Where:

$$S = c \cdot ln(\frac{x}{x_o}) \tag{2.12}$$

N denotes the number of samples, S is the sensation of the human experiences expressed as a function of the applied haptic stimulus, while $\hat{S}$ is the signal distorted by data compression, x and$\hat{x}$are the corresponding time-domain signals, c is a scaling constant that is calculated experimentally, I the magnitude of the applied stimulus, and I_0 the absolute detection threshold under which no stimulus can be perceived at all. The quality-prediction results show a (decreasing) quality trend equivalent to that from subjective tests, as the strength of the applied compression increases.

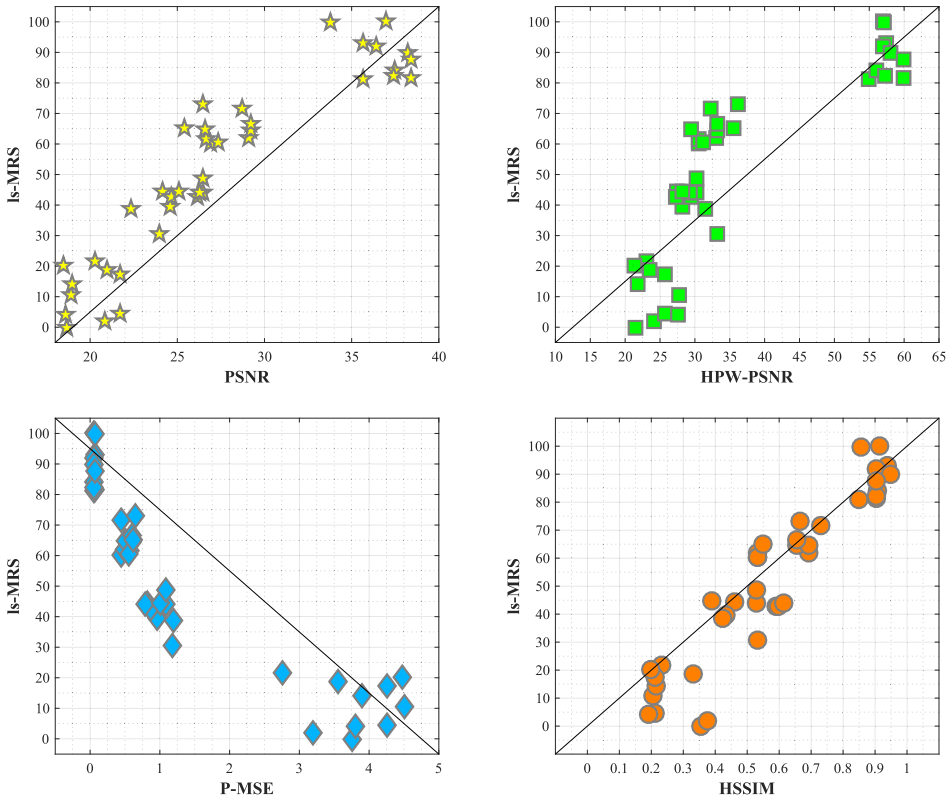


Figure 2.7: Scatter plots of linear-scale mean rank scores (ls-MRS) versus objective quality assessment methods. Each sample point represents one force-feedback signal. (a) PSNR. (b) HPW-PSNR . (c) P-MSE. (d) HSSIM [5].

All of the above objective quality measures focus on measuring signal fidelity by computing the "distance" between the two signals perceptually. However, they all assess signal quality based on error measures that operate solely on a sample-by-sample basis such that content-dependent variations are not considered. To fill this gap and by adapting The structural similarity index measure (SSIM)[102] which is a technique used for anticipating the perceived quality of digital television and cinematic pictures, including other sorts of digital images and videos, [5] introduced the haptic quality assessment measure haptic SSIM (HSSIM). HSSIM employs the generic definition of (SSIM) and in conjunction with Stevens power law as a haptic perception model. By assuming that X denotes the original haptic force feedback signal and Y is the correspondent compressed

signal of X. The overall quality assessment HSSIM mathematical model is derived as the following:

$$l(x_p, y_p) = \frac{2\mu_{x_p}\mu_{y_p} + C_1}{\mu_{x_p}^2 + \mu_{)_p}^2 + C_1}$$
$$c(x_p, y_p) = \frac{2\sigma_{x_p}\sigma_{y_p} + C_2}{\sigma_{x_p}^2 + \sigma_{y_p}^2 + C_2}$$
$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$
$$S_p = [l(x_p, y_p)]^\alpha \cdot [c(x_p, y_p)]^\beta \cdot [s(x, y)]^\gamma$$

Where:

$$x_p = k.x^b \qquad , \qquad y_p = k.y^b \tag{2.13}$$

by utilizing the Minkowski pooling approach:

$$HSSIM = \frac{1}{N}\sum_{i=1}^{N} S_p^r \tag{2.14}$$

where x_p,y_p denote the the perceived force signals after applying Stevens power law. ($\mu_{x_p}, \mu_{y_p}, \sigma_{x_p}$) are local means and standard deviations of the two perceptually manipulated signals x_p,y_p. σ_x,σ_Y,and σ_{xy} are the local standard deviations and correlation coefficient between original signal X and distorted replica Y. C_1, C_2, C_3 are positive stabilizing constants. $(\alpha_i, \beta_j, \gamma_k)>0$ and denote the variables used to tune the relative importance of the three comparison components. S_p is the perceptual quality map.

The underlying premise is that the signal quality is evaluated considering neighbouring sample dependencies and that only perceived distortions are penalized after accounting for human sensitivities. In order to quantitatively ascertain the potential of the above described objective quality measures to predict human haptic judgement, a force-feedback database was built and used to conduct subjective experiments. The database contains ten original and 40 distorted force-feedback test signals using the perceptual deadband-based data reduction technique described in chapter 3. The force-feedback

test signals are generated from a static interaction with objects in a virtual environment. The DB parameters are chosen to span the range of below, within, and above distortion detection threshold. Twenty-five participants participated in the subjective experiment. Fig.2.7 depicts the scatter plots for four objective quality measures against subjective scores. By comparing the overall performance using the three most used correlation coefficients and root-mean-squared error (RMSE). For most of these measures, HSSIM performs better with high correlation and low RMSE values, which suggests better prediction accuracy and monotonicity. However, it is worth noting that PSNR and P-MSE also perform quite well.

System-Level Quality Prediction:

The work in[103] provides a system-level mathematical model for Haptic Audio Visual Environment (HAVE) applications' QoE based on weighted linear combinations of QoS and User-Experience parameters as depicted below:

$$QoE = \zeta \cdot QoS + (1 - \zeta) \cdot UX \tag{2.15}$$

Where:

$$QoE = \frac{\sum_i \eta_i S_i}{\sum_i \eta_i} \tag{2.16}$$

and :

$$UX = A\frac{\sum_i \alpha_i P_i}{\sum_i \alpha_i} + B\frac{\sum_j \beta_j R_j}{\sum_j \beta_j} + C\frac{\sum_k \gamma_k U_k}{\sum_k \gamma_k} \tag{2.17}$$

where $\alpha_i, \beta_j, \gamma_k$ and (A, B, C) are the model weighing factors used to maintain the overall quality of experience between 0 and 1. S_i represents the QoS parameters in terms of delay, jitter, and packet loss whereas P_i, R_j, and Y_j denote the user experience parameters in terms of perception measures, rendering quality measures, and user state

measures. Lastly, ζ is used to control the relative priority of the QoS parameters versus user experience parameters. The authors' model was evaluated empirically using subjective testbeds on 30 participants who used a HAVE game. Further, they implemented a Fuzzy Logic Inference System (FIS) using the Mamdani MATLAB, that can anticipate the users QoE based on certain input parameters. Their FIS system maintains a percent error of 4.6% and a correlation coefficient of 0.92. Lastly, a recent study [104] tried to define a generic utility model to assess the QoE of multimedia content enhanced with sensory effects. The utility model is also referred to as Quality of Sensory Experience (QuaSE) and is given by equation 2.10:

$$QuaSE = QoE_{av}(\delta + \sum \omega_i b_i) \tag{2.18}$$

Where w_i indicates the relative weight given to sensory effect, b_i is binary variable used to confirm whether a given sensory effect has been used or not, and δ represents a fine tuning parameter. As noticed, the authors claim that the quality of sensory experience (QuaSE) is based on a linear relationship between the QoE of the audiovisual content (QoE_{av}) and the sensory effects number. As far as we know, the utility model is neither implemented nor validated.

2.4.4 QoE based on Machine Learning Systems and Implicit Metric

QoE based on Machine Learning can be defined as a systematic way of deriving an objective QoE metric depending on rules and relations obtained from training data. The system must be trained with historical user data that includes users' overall evaluation of an application. On one hand, It has been reported that the QoE-QoS relationship is strongly non-linear, and no single mathematical function can adequately explain this relationship [86]. Recently, it looks like non-linear, non-parametric regression trees or random forests are the best models in terms of forecasting precision and accuracy [105].

On the other hand, and as has been elaborated in the previous sections, subjective QoE methods rely solely on filling out surveys right after undergoing the system or environment of interest. However, this only considers the user's emotional and physical states immediately after the experience, and any variation in emotional or physical states during the experience is missed. Besides, these approaches depend on conscious responses and often fail to provide sufficient understanding of underlying perceptual and cognitive processes. To tackle these issues, QoE frameworks based on capturing physiological and physiological metrics have surfaced [106, 107, 108]. The authors in [106] propose using physiological methods to assess the quality of experience of virtual reality systems. Using stress as a use case, they suggest that the sympathetic nervous system is stimulated, resulting in increased blood flow, heart rate, and breathing rate. The stress level can be estimated by calculating the aforementioned parameters. Another study suggests the use of heart rate and electrodermal activity captured from the Fitbit and PIP biosensor respectively to assess QoE for immersive virtual reality environments. The outcome from this study revealed that HR and EDA are correlated with physiological arousal and hence impacting the user QoE.

A master thesis [109] has evaluated the network delay impact on QoE of the user according to a VR industry application. The thesis aimed to develop an application that helps to show the influence of network delay on the QoE of users. Technologies influence factors play an important role to consider immersive human-centric perception for incorporating in QoE. One of which is the network delay impact that can be measured subjectively by human perceptions. In other words, network delay is the latency caused by the undergoing network components. However, once we integrate the VR experience, other human factors (enjoyment, usability, immersion, etc.) must be accounted for and evaluated. These factors can be assessed after conducting the experience. Regarding measurable health, emotional states, and fitness, heart rate is a crucial measurable method for physiological metrics. This thesis integrated heart rate investigation in the experiment by placing an electrocardiography (ECG) device on subjects' bodies to mon-

itor their heart action as invasive means to detect pulses produced by heartbeat. For the sake of capturing accurate physiological data, using a wristband called Empatic E4, the heart rate signals were captured with the blood pressure during the experiment. Resulting from such physiological metric results, the experiment showed some emotional states like enjoyment immersion. Combing capturing heart rate and electrodermal activities, the results show a clear increase in attitude and stimulation. The problem is seen when more and different modality interactions are involved due to high delay as an impact on the virtual reality environment. The author followed ITU-T for the evaluation of the subjective metric. To collect audiovisual stimuli, the author used two methods of rating, ACR and DCR. The first one was built based on a 5-point Likert Scale that varies from Bad to Excellent. One stimulus is evaluated one at a time so the subject can assess their experience for that stimulus. Whereas the second method was used to assess varied interests according to visual characteristics. Also, this method was built on a 5-point Likert Scale that varies from Very Annoying to Imperceptible. The network delays that were used for the experiment were 0, 1, 2, and 3 seconds. The goal was to find out the noticeable delay a user can rely upon to decide the feasibility of the VR application to use or not. The network delay value between 2 and 3 seconds was perceivable by the subjects that affect their QoE. The result shows that some factors like experiencing new technologies, ease completion of a role, and subjects' expertise background have dominancy in network delay of virtual reality surrounding. This infers that subjects will admit the delay of virtual reality surrounding despite network delay value.

2.5 QoE in Modern Multimedia

A holistic study of recent QoE in interactive multimodal systems is presented in this section. The review complimented the discussion based on the related works in this field, specifying the advantages and disadvantages of existing QoE studies. Authors in [17, 18] conducted a study based on an online questionnaire to evaluate the effects of

user perception of specific multiple-sensorial media components like haptic and airflow. Fifty-four users participated in their experiments and filled up their feedbacks accordingly. 70% of the users found that haptic and airflow effects dramatically enhanced their sense of reality and enjoyment level. Whereas, 4% of users experienced some distraction and annoyance, the rest of users gave neutral feedback. This infers the importance of user experience on multimodal applications. However, the experiment was done locally, therefore the results cannot be extended to tele-multimodal systems where some challenges have to be overcome when disseminating multimodal content over the Internet. The authors addressed this issue in [71]. They use the MPEG-7 description scheme to integrate multiple sensorial metadata into the audiovisual stream. While evaluating this system, non-precise synchronization between modalities (especially when including olfaction) was reported as the main reason to reduce the user enjoyment levels. This finding was also obtained by [31]. Also, [71] found that reducing some sensorial effects while presenting content to users does not have a negative impact on user perception of the multimedia component. Further, in evaluating what aspects of the mulsemedia content would affect user perception of the delivered content, the result was tactile (62.5%), wind effect (31.25%) with olfaction (6.25%).

In the context of evaluating haptic using biometrics, machine learning techniques have been utilized to optimize the classification accuracy of users' tactile experience by measuring brain activities. One of the best measures to render brain data is through Electroencephalography (EEG) devices. Authors in [110] proposed a time-shifting solution using the neural machine learning classification-based method. This classification is aimed to enhance the distinguish of presence response of tactile experience during touch screen activity coming from neural responses. Using the EEG device, the raw data of the brain will be extracted with each user's touch to be matched with cognitive processing, awareness, and sensorimotor. Any irrelevant features will be dismissed, so the points following each touch are only counted. The EEG output is injected into a nonlinear SVM classifier to train the data coming from the EEG channels. To increase the classi-

fication accuracy, the authors used a voting classifier technique aiming to integrate the result according to the highest voting. This was achieved by using a validator called stratified 10-fold cross for each class of training and testing. Compared to 60% of other brain regions, the highest classification accuracy was achieved from the middle frontal cortex with 68% accuracy. The result shows that using this classifier along with the time-shifting, the classification accuracy enhanced to 85%. However, this relies on the EEG performance as a downside of this proposal. Also, the model is only deterministic subjectively on the user's touch features.

Another approach proposed by [111] to provide a texture classification using a 10-fold cross-validation technique as a machine learning classifier. Their work aims to extract the relevant signals to varied textures according to live surfaces touch using EEG features, so the roughness surface (flat, medium, or high) can be measured. Dozen of the participants were hired to do the experiment, 8 females and the reset were males who all were sure had no implants either metal or electronic, and were healthy people with ages above 23 and less than 32. The subjects were asked while mounting the EEG to tap on a transducer. Then the rendered EEG data has been analyzed using the Power Spectral Density. To extract the EEG features, three values were calculated, namely the spectral properties, the normalized power around 7-12 Hz, and around 13-30 respectively. After the extraction process, a 3 class of SVM machine classifier has been used to classify the features to distinguish the variations of textures. It shows from the results that the texture discriminating of EEG features is vulnerable to the roughness level increase. For example, in the friction case, the speed increase makes the accuracy increased. Not like the fast and medium speeds, in the tap case, the accuracy is higher due to the low speed. This might infer the capability of the participant to choose the textures according to the speed of touch. According to this inference, the EEG signals could affect the resulted features because of the motor components. Also, since the participants were right-handed, this might lead to different discriminating between textures once compared to left-handed participants.

The paper [4] examines the cross-modal effect between visual, gustatory and olfactory via a pseudo-gustatory display outlined above. The authors evaluated the efficiency of their work subjectively by asking participants to try different flavours using Meta cookies i.e a cookie with a neutral taste. 44 users experienced 6 cookie visual/scent combinations. Each user was first asked to try plain cookie enhanced with visual and olfactory effects and the second trial without those effects. For more than 79% of the trials, the participants indicate a change in the cookie's taste.

In [112] authors discuss a new assessment of olfaction QoE using a novel neuronal model called Functional Connectivity Map (FCM). In which, volunteering smellers, from Ecole Polytechnique Fédérale de Lausanne (EPFL), were asked to experience different scents while their brain activities were monitored using an Electroencephalography (EEG) system. The recorded signals are then passed to the FCM for analysis and further interpretations. The resulting output was used by a machine learning algorithm known as Support Vector Machine (SVM) to classify and anticipate the perceived level of pleasant or unpleasant odour by the participants. Versus the questionnaire's ground truth, the EEG approach can predict Quality of Experience with up to a 65% accuracy rate.

Authors in [113] examined the impacts of crossmodal QoE by capturing subjects' eye gaze and heart rate while the users conduct a mulsimedia experiment. The process is aimed to map between 12 users' experiences of video, smell, sound, and haptic to question them about their experience result of excitement and observation. The authors used some types of content in the experiments like video, audio, and olfactory. The experiment's participants were selected based on the experimental (EG) or control (CG) group. The latter group is used with visual content only excluding cross-modality content. The experiment results showed that there is a remarkable user observation caused by the cross-modality that mapped with mulsimedia. Another observable result was noted which is different users' moods generate higher heart rates in the EG compared to the one in CG. A downside of the study is that it didn't examine the content differences in both

the crossmodal and multisensory.

Author in [114] proposed a framework called PanSEM SER 2, aiming to deliver the effects of hybrid systems that have different sources of mulsmedia like vibration, smell, lighting, and, wind. The integration of the mulsmedia sources was done in different situations and constraints. This encourages the interoperability of multi-touch interfaces, voice processing, and brain-computer interfaces, so mulsimedia applications and devices can be accommodated in users' environments. The framework architecture involved an evolution of the PlaySEM SER platform that needed to tackle some challenges, so changeable, extensible, orthogonal, and adaptable features can be built in mulsimedia systems. One beauty of this framework is that it allows adding as many devices as possible from different sources with unique specifications. It provides flexibility for delivering multisensory effects with multi-communication and multi-connectivity protocols, multi-standards, and it is agile for new technologies. Also, the framework gives the freedom to design multimedia systems by allowing to use of SDKs and APIs. However, it still does not allow for new content adaptation concerning users' feedback. Also, safety and security were not considered in the framework design. Moreover, integrating infant sensory data like gustatory in mulsimedia systems is still not tackled.

Lastly, the impact of QoE differences from one individual to another might play an important role in indicating a better QoE outcome. This worthy observation is investigated in [115] by integrating haptic and olfactory as mulsimedia stimuli in heterogeneous environments putting into consideration versatile sources of users' differences like smell sensitivity, age, gender, and, education. The study was built based on two QoE questions, first what is the cross-modality impact that matched smell sense on user QoE of mulsimedia? And user's differences impact on user QoE of mulsimedia? The authors analyzed the impact of individuals' differences (smell, gender, age, education) on crossmodal QoE. One experimental group consists of 24 subjects (9 males and 3 females) and a control group consists of (7 males and 5 females) have been allocated for participating in the study. Some video clips, a haptic vest, and an olfactory device are employed in

the study. The feedback results showed that there is an influence of individuals' differences on the user QoE of mulsimedia. The age impact was clear on the groups; 18-21 years group was satisfied using the multisensorial experience more than the older group (above 21). Also, gender impact showed that females' smell sense is more sensitive than males'. Similarly, education level shows an impact on mulsimedia QoE where postgraduate participants watching video while wearing haptic vests were more enjoyable than the undergraduate level group. The outcome of this work shows that there are correlations between participants' feedback. This was clear when the pleasure is achieved once cross-modality matched smell as interpreted by the haptic vest. Although the study has examined some important individuals' differences, other individuals' differences were not considered in the study (like culture, personality, etc.), which no doubt will enrich mulsimedia QoE output.

2.6 Summary:

In short, QoE in multimedia can be done in three approaches. The first approach is based on subjective tests in which users explicitly give their opinion about the system they used. Then, the results are passed through regression analysis to come up with the optimized technical factors that enhance the overall experience. This approach is very expensive, time-consuming, and lacks repeatability. Also, it cannot be applied in real-time. The second method is based on algorithmic and/or mathematical derivations. In this approach, QoE augments QoS but does not totally replace it. Such an approach suffers from feasibility and accuracy issues as there is no comprehensive model that can quantify the multidimensionality and large individual variability. However, to precisely capture the QoE using objective testing, more developments are needed, such as the mapping of network performance metrics (intrinsic QoS factors) to user experience related factors (e.g., haptic perception), the integration of sophisticated models for human haptic control into the objective quality metrics, and the development of joint metrics for audi-

tory, visual, and haptic modalities. The third type is based on a machine learning-based approach. In the visual quality assessment domain, many universal machine learning models have been proposed [116, 117]. The main challenge for the machine learning approach is how to learn rules from a human semantic description of what they are experiencing. For example, humans can describe their experience as "very good," "fair," or "horrible." Directly mapping human linguistic descriptions to meaningful features that will represent the quality of the stimuli is a crucial step. Therefore, more efforts have to be carried out in optimizing a QoE model, for example, automated models, to capture the heterogeneous nature of multimodal systems in a smart and cost-effective way that satisfies the most important entity of the multimedia system (i.e. the user). This is where this thesis steps in. It incorporates the subjective, objective and biometrics of the user and puts it in a kaleidoscope to best assess the user's QoE and their interaction with the real and virtual environments of the DT architecture.

Chapter 3

The Proposed QoE Evaluation Taxonomy for Digital Twin Systems

The revolution of the Tactile Internet and its potential impact on society is expected to add a new dimension to human-to-machine interaction in a variety of different application fields, **including but not limited to** healthcare, education, industry, and entertainment. Manufacturers, content providers, as well as the communications infrastructure enabling the envisioned Tactile Internet have to meet a number of design requirements. As reported in the article titled "Why is multimedia quality of experience assessment a challenging problem?" [118], there is no standardized taxonomy of objective QoE aspects that are available (correlated with human perception and judgment). Basic insufficient attempts have been made, for example, [119, 120, 121]. As a result, it is critical to implement appropriate taxonomy standards, which will aid not only in the facilitation of quality measures but also in the standardization of the architecture of commercial QoE systems. In this chapter, we propose a taxonomy for the evaluation of the QoE for digital twin systems taking into consideration haptic and other sensory stimuli such as olfaction and VR. The taxonomy classifies QoE-related parameters into both objective and subjective influencing groups in order to consistently outline the characteristics and feature levels for both objective QoE engines and human users' physiopsychological aspects for

better use.

3.1 QoE Taxonomy for Digital Twin

To evaluate the (proposed) Digital Twin system, assessment factors have to be defined in such a way that addresses the end user's perception. Further, the assessment should take into consideration the bidirectional nature of haptics' exchanged information (the forward flow (i.e. the (x,y,z) positions and velocities data from the human operator to the robot) and the backward flow (i.e. the torques and forces information from the robot to the human operator). Consequently, humans not only feel haptic feedback, similar to audio/video but also physically act upon an environment [122]. In this context, QoE is defined as a multi-level paradigm of the users' perceptions and behaviours, representing emotional, cognitive, and behavioural reactions that are both subjective and objective, while dealing with services, products, or applications [123, 76]. Accordingly, our QoE taxonomy for DT applications should involve: both the subjective and objective quality influencing factors. Influencing factor (IF) in general is any characteristic, feature, attribute of a real user, digital twin, system, network, application, or context whose tangible state or setting might have an effect on the overall QoE of the end-user[124]. The objective quality factors aim to address technical metrics related to the DT system such as network, content and hardware and are referred to as Key Performance Indicators (KPIs). Whereas the subjective quality factors are addressing the user's physical/mental, the degree of perception, context, and the usability of the system. This higher-level architecture of our taxonomy is shown in Fig.3.1.

3.2 Objective Quality Influencing Factors

Objective Quality Influencing Factors to define the technical parameters that ensure the smooth flow of the DT application from the content, manufacturer, and service provider

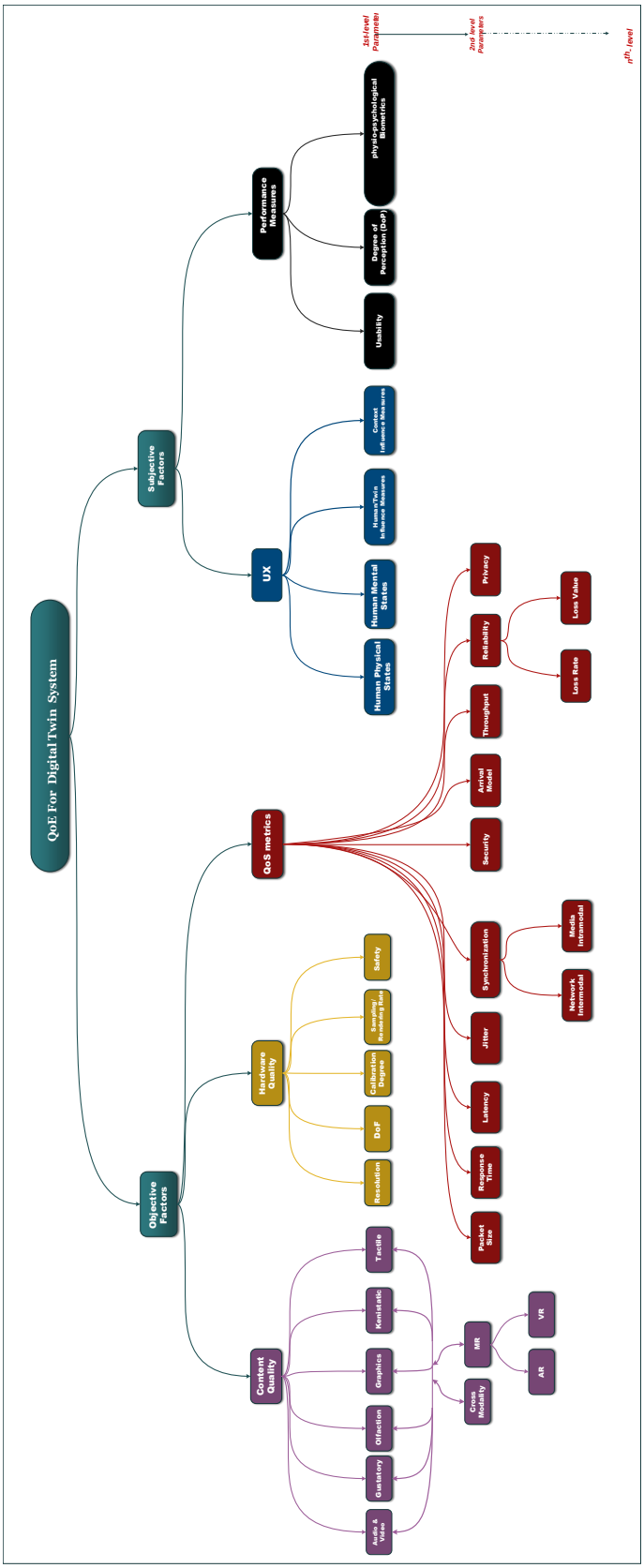


Figure 3.1: Proposed QoE Taxonomy for DT System Evaluation Model over TI

i.e network perspectives. A Key Performance Indicator (KPI) refers to a mathematical grouping of factors called performance counters, that allow for the system events, which prove that something happened in a network, application and, hardware level of the DT application, such as failure or success in a particular network procedure, and they will be placed in KPI formula. Consequently, we divided them into three classes: Network influencing Quality, Content Influencing Quality, and Hardware Influencing Quality.

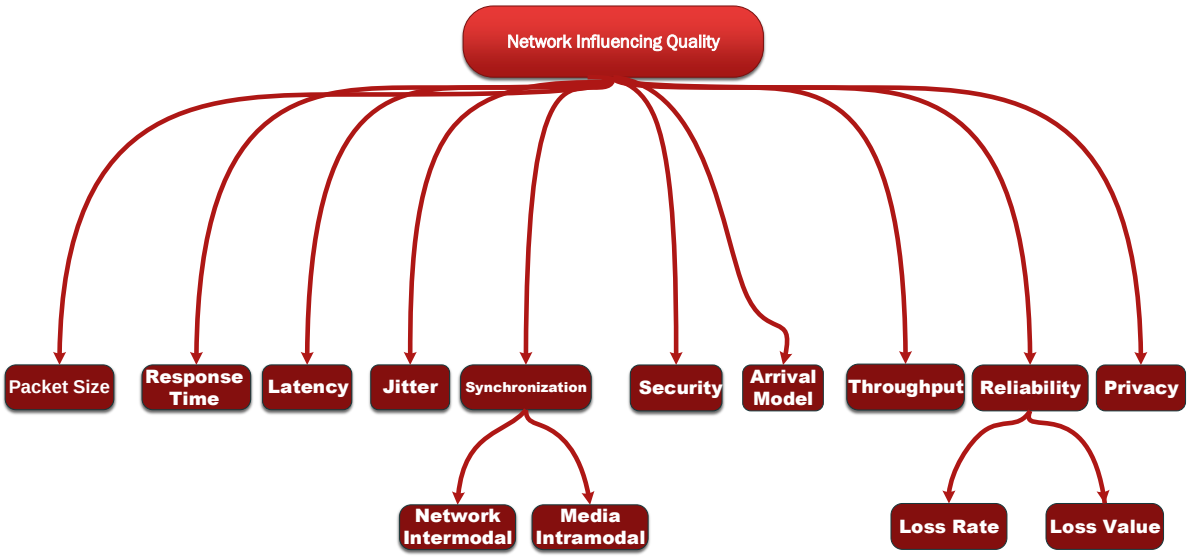


Figure 3.2: Network Influencing Quality

3.2.1 Network Influencing Quality

The Network influencing quality resembles is strictly related to the Quality of Service Key Performance Indicator. Key performance indicators utilized by service providers as average metrics that mask the real experience of users and are particular of a certain technology. Evidently, generic QoS difficulties (e.g., packet loss, latency, delay variation, reordering, bandwidth shortage) infer generic QoE problems (e.g., flickering, artifacts, instability, glitches, extreme waiting times). Therefore in this jargon, QoS-KPIs refer to the quantitative and/or qualitative performance aspects of media transmission over the

network from the network provider perspective. This class address the characteristics of a telecommunications service that bear on its capacity to fulfill the stated and implied needs of the user of the service. We highlighted below most of the metrics used to ensure the quality of service for the newly evolved DT networked application.

- **Packet Type and Size**: denotes how the modality is encoded and sent over the Internet. The size of the packet is defined based on the codec standard as illustrated in the next section.

- **Response Time**: The period taken by a system or the user to act/react to any changes in the local and remote environment, usually measured in milliseconds or microseconds.

- **Latency/Delay**: The end-end time required for the packet to be sent over the network and received by the destination is usually measured in milliseconds or microseconds. End-to-end delay can be measured in one direction from sender to receiver using the equation3.1:

$$T = T_{pack} + \sum_{h \in path} (T_h + Q_{h+} + P_h) + T_{play} \tag{3.1}$$

  At which T_h denotes is the transmission delay, Q_{h+} denotes the queuing delay required at each network traversing node, and P_h represents the propagation delay. Note to forget T_{pack} which denotes the packetization delay that occurred due to the encapsulation process at the source.T_{pack} and T_{play} depends on encoding/decoding delay that occurs while conversion of analog to digital (A/D) signal into samples and vice versa. For instance, in the IP phone telephony essence, $T_{pack} + T_{play}$ is equal to 1/20hz.

  1. Haptic Feedback Latency: In a teleoperation scenario via a DT architecture, the use of a force and vibrotactile sensors are required to obtain the actual force signal perceived from the communication/interaction between the robot

end-effector and the remote objects. The overall all end-end latency can be demonstrated as: $T_{overall-latency} = (t_{mux} + t_{net} + t_{DAQ} + t_{\text{int}erface})$ Where:
t_{mux} and t_{net} deonte the delay of multiplexing and network delay. To minimize the noise level of the force feedback, the data acquisition (DAQ) card with on-board Digital Signal Processing (DSP) filter has to be used, thus t_{DAQ} denotes the delay produced by the DAQ card. According to [72] $t_{DAQ} \cong \frac{1}{f_{cutoff}}$, where f_{cutoff} is the cutoff frequency of the DSP filter which is approximately= 31.25 as such t_{DAQ} is approximated by 32ms. $t_{\text{int}erface}$ denotes the latency that arises between the workstation and the haptic interface. Both the phantom and falcon have a refresh rate equal to 1KHz, hence, $t_{\text{int}erface}$ is equal to 1ms. Consequently, The overall latency on the force feedback can be defined as $T_{overall-latency}$= t_{mux}+t_{net}+33ms.

2. Audio latency: The cumulative latency experienced during acoustic feedback is written as follows: $T_{latency} = t_{env}+t_{render}+t_{encode}+t_{mux}+t_{net}+t_{decoder}+t_{display}$ where t_{env} depends on the delay caused audio propagation in the environment. it ranges from 5 to 20 ms. t_{render} is related to the hardware quality as it depends on the sound card propagation delay. According to [72] which used PortAudio I/O library t_{render} and $t_{display}$ is approximated around 12.7 ms.Nonetheless, the overall latency of tranmission the audio is approximated as $T_{latency} = t_{mux} + t_{net} + 39.4ms \pm 15ms$ based on the value of sound propagation speed in the environment.

3. Video latency: The cumulative latency experienced during visual feedback: $T_{latency} = t_{source} + t_{render} + t_{encode} + t_{mux} + t_{net} + t_{decoder} + t_{display}$ and it depends on both the codecs such as (H.264) of the video and the quality and speed of graphical processing unit. According to the YouTube and Skype statistics for both video streaming and conferencing, an acceptable $T_{latency}$ is margined around 125 ms.

4. Mobile Network latency t_{net}: In a cellular system, the total latency can be

reduced across different segments of the network, i.e., RAN, back-haul and core. According to [125], the total End-to-End (E2E) latency in an LTE-A cellular network is:

$$T_{\mathbf{E2E}} = \ 2 \ . \ (\tau_{RAN} + \tau_{Backhaul} + \tau_{Core} + \tau_{Transport}) \tag{3.2}$$

where τ_{RAN} denotes the transmission time over radio Access network, $\tau_{Backhaul}$, τ_{Core}, and $\tau_{Transport}$ correspond, respectively to: (1) the packet transmission time between the User Equipments (UEs) and the network nodes that are part of the mobile radio access network such as ENodeB or Ng-ENB in 5G network, (2) the time required to make connections between the core network (i.e., Evolved Packet Core - EPC) and the eNodeB, (3) the processing time needed at the EPC, and (4) the data communication time between the EPC network and the Cloud/Internet. While τ_{RAN} depends on network coding schemes, and $\tau_{Backhaul}$ depends on the medium used.

- **Jitter**: According to RFC 4689, jitter is defined as the latency fluctuation between two consecutive packets belonging to the same flow or stream between two systems [126]. This usually occurs because some packets take a longer time to travel from one point to another in the system mainly due to network congestion (which is totally random and time-variant), timing drift and route changes. Jitter can be estimated by the following formula:

$$Jitter = \left|(t_{i,r} - t_{i,s}) - \left(t_{[i-1],r} - t_{[i-1],s}\right)\right| \tag{3.3}$$

Where:

$t_{[i-1],s}$,$t_{i,s}$, $t_{i,r}$, and $t_{[i-1]}$ denote the sending time of packet $(i-1)$, the sending time of packet i, represents the receiving time of packet $(i-1)$, and represents the receiving time of packet i respectively. Generally speaking, The average or mean value of the Jitter over a long period of observation is often the point of interest.

- **Throughput**: The actual available bandwidth offered by the multimedia channel. it is also referred to as the amount of data transferred from the sender to the receiver. Measured typically in bits/second or bytes/second. Bandwidth refers to the maximum transmission rate that a medium or network will handle, but throughput is not always the same as bandwidth. In other words, throughput is the rate at which data can pass through a medium or network in one second despite the fact that it is a slow medium or network. When throughput equals usable bandwidth, a network is said to be maximally efficient. The throughput efficiency formula can be calculated more than one way, but the general formula $E = RxT$, where R is the flow rate and T is the flow time.

- **Reliability**: The capability of the digital twin system and its factors, i.e., haptic audiovisual environment to steadily operate based on various specifications and network conditions. In this context two KPIs can be defined:

  1. Packet loss rate: The percentage of data that have been sent and not delivered to the receiver due to loss, reordering or re-transmission. it can be quantified as $PacketLossRate = \left(\frac{\text{Packets sent} - \text{Packets received}}{\text{Packets sent}}\right) * 100$.
  2. Error/loss Magnitude: Occasionally DT modalities such as collaborated haptic VR packets are distorted because of bit errors caused by interference, congestion, and noise. The receiver has to distinguish that and, if the information included in the packet is vital to correctly render CHAVE scene, might request for this data to be resent from the sender.

- **Update rate**: How many times per second the data of each modality needs to be refreshed. It is different than the sampling rate and is measured in Hz. for instance, the most efficient audible frequency is ranging between 16 Hz to 20,000 Hz.

- **Arrival model**: Defines the arrival and losses characteristics of the modality

packet on the Internet. It can be either Periodic, Aperiodic, Markov Model, or Bernoulli Model.

- **Synchronization**:

  1. Intermodal Synchronization: measures the temporal correlation of the multimodal presentation. For instance: the sequence of sensorial effects in a 4DX^{TM} scene.
  2. Intramodal Synchronization: measures the temporal correlation of the same modality presentation. For instance, when intermittent stimuli are presented to the eye they are perceived as separate if the rate at which they are presented is below a certain value. If the rate of presentation of the intermittent stimuli is slow, it appears to stay on but with changes in intensity, producing the sensation called flicker. Examples of Intermodal Synchronization, a Youtube video streamed at 50 fps implies that each of the frames has to be displayed for 20 msec. Similarly, a haptic kinesthetic refreshed with 1 kHz, each of the data samples must be captured and displayed for 1 msec.

Synchronization[45, 127, 22, 128] is another important challenge to consider in order to address different media streams that will traverse the multi-modal communication module of the DT. In fact, some studies have proven that the olfaction stream is very fragile to this condition [31]. In order to achieve an acceptable QoE, video, voice, olfaction, gustatory and haptic data streams will need to be synchronized accordingly without affecting the end user's perception. Different synchronization techniques have been proposed over the years, which include synchronization buffer schemes, time-stamping and time adjustment algorithms, each with its own uses and requirements. To maintain a synchronized stream of media, three types of levels need to be addressed: (1) Intra-Stream, (2) Inter-Stream and (3) Group Synchronization Control (also known as Inter-Destination) synchronization. One of the most prominent synchronization schemes, known as Adaptive synchronization, can

be used to address multiple channels of communication, where both intra-stream and inter-stream synchronization can be taken care of. In addition, according to the findings in [129], adaptive synchronization is immune to clock offset and/or clock frequency drift does not require a global clock and can support optimal delay and buffering for a given quality of service requirements.

- **Price or Cost**: The amount of compensation paid or awarded from one entity to another in return for service or good delivery/ deployment, it can be determined via factors associated with monetary, energy, automation or another productivity of the service.

- **Security**: The level of protecting the data exchanged through the use of communication channel technology. it is the KPI characterized by the demand for extremely secure communication, and the protection of privacy and user integrity

- **Privacy**: Deals with what personal information can be shared with whom and whether messages can be publicized or not.

- **Accessibility** is the KPI which in the planned coverage area represents the measure of service availability to the end-user.

As has been discussed in the related work (chapter 2), to derive the mapping relationships between QoS and QoE, researchers rely on the quality comparisons between what is referred to as the reference and the outcome. The former denotes the undistorted content of the multimedia content such as the original haptic signal before transmission, original audiovisual content, or the undistributed delivery service such as download action. While the latter denotes the form of any possibly impaired or even demolished multimedia stimuli at the receiver side e.g distorted audiovisual or haptic signal after the transmission took place. It also can represent a distorted service such as a delayed download action. Hypothetically, the nearer the quality of the outcome goes to its correspondent reference, the better the QoE. In an ideal scenario of multimedia communication, both the refer-

ence and outcome match, as such the communication channel between the sender and the receiver can be referred to as a transparent network. A distortion of the outcome will adversely influence the quality of the multimedia (e.g., audio clarity, smoothness of a video) and/or effectiveness of the service (e.g., download rates, service activation time). In that case, the QoE correlates to the remaining outcome's quality after such a distortion. The effect of distortions can, along with other approaches, be expressed with the help of utility functions [91]. Generally speaking, quality comparison can take place in two common situations. A laboratory-controlled setup as a result of which both the reference and the outcome are observed and ready for offline comparison which gives more accurate details. While the other situation is referred to as a communication setup in which only the outcome is obtainable for online comparison. Based on the metrics to study, observation of the multimedia content and related network traffic can occur at different levels. Application-level observation includes the inspection of the payload, which allows obtaining a comprehensive view of the content as well as the timing of outcome and reference. Concerns and issues related to the latter may occur from the implementation of the application as well as network nodes, connections, and the network stacks in the end-infrastructure. Consequently, assessments on the network level should be performed to inspect the flow of packets in the context of completeness, timeliness, and pattern analysis involving bursty losses or correlated latencies. For testing any mathematical QoE-QoS model, it is very recommended to consider the observation level at the application or system, context, and network levels. System of measurement can be categorized according to the below outline and as illustrated in Figure 3.3.

- **Full Reference (FR) Metrics** In this category, both the outcome (Y) and reference (X) are obtainable and can be subjectively and objectively compared offline at any observation level (application, e.g., tactile signal and network levels e.g packet traces). $FR(X, Y)$ is the most effective schema to estimate the quality of experience for any given system as it allows the extraction, assessment, and comparison of both QoS and QoE parameters in a controlled laboratory setup and at

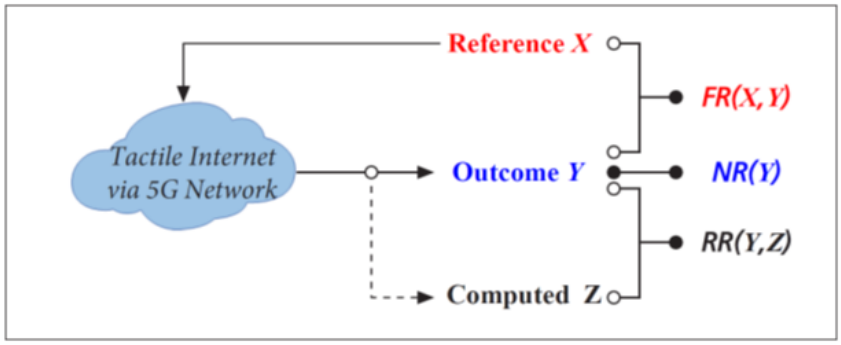


Figure 3.3: Categorization of System of Quality measurement. (Metrics) used to derive QoE-QoS models

any time. In the context of image quality, A well-known model of an FR metric is Peak signal-to-noise ratio PSNR [130]. In the context of haptic signal quality, mean squared error (MSE) compares original and decoded signals and provides numerical values. Both PSNR and MSE are automated calculating procedures, hence they are repeatable and produce statistically significant findings.

- **No Reference (NR) Metrics** In this category, only the outcome is available with no reference to be compared with. This is a normal online observation situation that concentrates merely on the resultant quality as perceived by the end-user and can be achieved subjectively via surveys and observations or objectively via an algorithm or a mathematical model on behalf of an actual user, aiming to mimic or expect user perception by extracting the key features of the outcome. When $NR(Y)$ QoE is applied in the networking domain, it does not tackle and distinguish between the quality problems issues yielding from the very reference and other disturbance by the network. Consequently, $NR(Y)$ has a hindrance when it is used to assess the QoE-QoS relationships seeking at capturing the effect of the network. However, applying $NR(Y)$ as a generalizable experience according to user-acceptance-based QoE parameters such as video quality (resolution, contrast, color dynamic range (HDR)) or response times of a given multimedia application may fairly serve as a

foundation for QoE-QoS relationships.

- **Reduced Reference (RR) Metrics** At the reduced reference metrics, the same set of parameters are extracted from the reference and outcome. For example, image quality assessment at the application level can be represented using the Quality Index With Adaptive Sub-Dictionaries (QSAD) metric [131], whereas throughput variations and packet losses can be used to describe QoS parameters at the network level. As a matter of fact, RR metrics are based on measuring, summarizing, and interpreting results from FR research. Consequently, the function $RR(Y, Z)$ is used to assess the QoE in which Z denotes the measured parameter attainable at the receiver side and potentially has its root conveyed from the sender side. Hence, RR metrics describe key QoE and QoS parameters in a very concentrated approach, conveying and/or extracting them from the sender and receiver make them applicable in online service scenarios hence comparing them can be utilized to tackle quality issues. In essence, they can be candidly used to derive formulate QoE-QoS relationships.

Figure 3.4 illustrates a hypothetical general pattern of the mapping curve describing the effect of QoS disturbance on QoE. The schematic quantitative $QoE = F(QoS_i)$ can be represented by the following function 3.4:

$$QoE(QoS) = \begin{cases} 5 & QoS_i \leq T1 \\ e^{\alpha_0} + e^{\sum_{i=1}^{n} -\alpha_i QoS_i} & T2 \leq QoS_i > T1 \\ user - based & QoS_i > T2 \end{cases} \tag{3.4}$$

Where the x-axis represents the impairment of QoS_i of the i parameter e.g., (decreasing download rate, increased latency, growth in the packet loss and jitter), α is the weighting factor of the QoS impairment, and the QoE rating e.g, (ACR or MoS) is denoted by the y-axis. As can be noticed from the figure Figure 3.4, the mapping curve between the impairment of QoS and QoE can be divided into three zones. In zone 1,

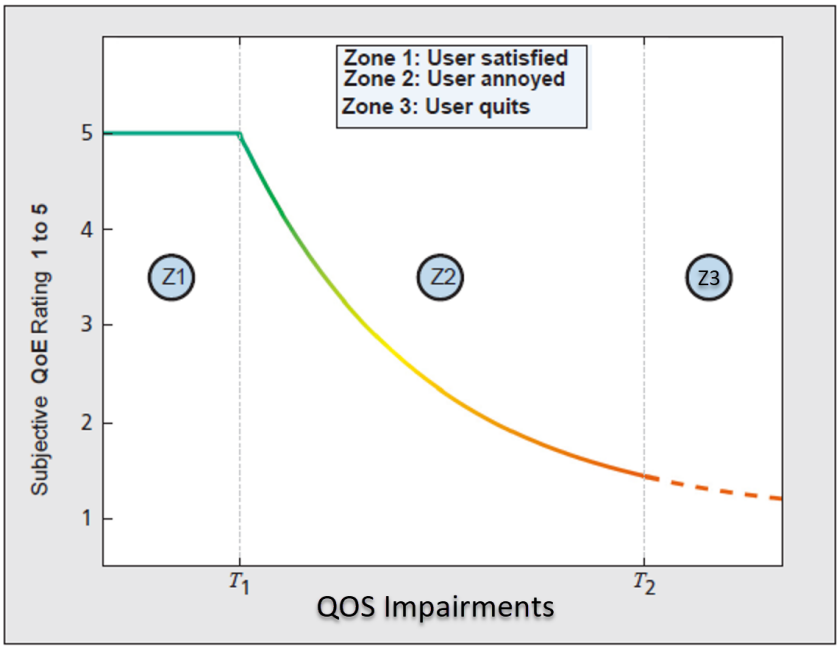


Figure 3.4: Hypothetical curve fitting of generic QoE-QoS dependency

for a fading QoS disturbance (i.e., in the situation of a transparent network) A minor increase of the QoS disturbance until it is equal to a given threshold $T1$, QoE has a high constant and optimal value in which the user's satisfaction and enjoyment are not affected at all. For example and in the context of web browsing, if the web page is loaded faster than half a second ($T1$= response time =0.5 seconds), the average user will not notice that at all. Zone 2 represents a sinking QoE. In more detail, as soon as the QoS disturbance exceeds $T1$, the QoE rating as well as the user's appreciation and satisfaction, decrease exponentially. As can be noticed in zone 2, if QoE is high, any growth at the QoS disturbance will adversely impact the user's rating. On the other hand, further QoS disturbance might not be that critical anymore in the case of a low QoE rating. Lastly, zone 3 represents an unaccepted QoE. More specifically, as soon as the QoS disturbance exceeds threshold 2, the QoE may fall, i.e. the users' satisfaction will be negatively impacted and they quit using the service completely. Normally, once the QoS disturbance parameter grows significantly, the QoE metric and user's perception

of the quality sink. It should be noted that while $T1$ denotes a significant threshold that thoroughly might be collocated with the y-axis, $T2$ denotes a user-dependent threshold, for instance, some users may still watch a YouTube video with 144P resolution with others do not. Consequently, the QoE model has to consider the end-user degree of experience as argued in the coming few sections.

3.2.2 Content and Hardware Influencing Quality

Content quality is the core of any multimedia system. It addresses the codecs' aspect of the DT modality as generated by the content creator.In this context, High-Efficiency Video Coding (HEVC) also referred to as H.265, Dolby vision/Atoms, and IEEE P1918.1.1 are nowadays the most effective and widely used codecs for video, audio, kinesthetic and tactile information exchange across communication networks. Note that both the work to provide olfaction and gustatory codecs is still in their embryonic stage and needs more attention. For MR modality, both VR and AR are considered separately first and eventually blended and mixed modalities are considered. Cross-modality, is highlighted in red, denotes the impact of one modality on another, for instance, the impact of video quality on the perception of audio quality and the impact of VR and olfaction codecs in the gustatory system. The detailed factors for each modality are illustrated in Figure 3.5.

With regards to QoE studies in the field of content quality, multimodal data fusion and compression are areas of interest. As the access to data continues to grow, we need to incorporate both hard data and soft data to improve real-time decision-making processes. Multi-source data and information fusion is a technology to enable combining data/information from several sources in order to form a unified picture of the DT. As mentioned previously, Digital Twin's data can be selectively collected from different hard data sources and soft data sources. Clearly, the accuracy, biases, and levels of observation of soft data are quite different from those of hard data. On the other hand, soft data contain valuable inferences and observations not available in standard

sensors. Multimodal data fusion considers the incorporation of multiple modalities to be processed while maintaining the spatial, temporal and contextual dependencies of the media scene. Fired by the need for real-time simultaneous recording and transmission of voluminous data that are produced by multiple sensors, multimodal compression has received significant attention. The compression research for audio, video, and 3d graphic data has attained a matured stage (including MPEG, H.261, to H.265). However, despite the stringent need for haptic data compression, this area is still in its infancy and many open research directions have emerged. So far, the effective approach for haptic data compression is accomplished by minimizing the amount of data using perception-based data reduction. For instance, not all haptic data need to be sampled with 1 kHz; tactile information can be fully sensed with a 30 Hz sample rate. In addition, extracting semantic information from the kinesthetic scene will drastically decrease the data volume needed to resemble the force-feedback scene, and thus allowing efficient haptic storage and transmission. For instance, the perceptual dead band (PD)-based data reduction methods for kinesthetic information have been deployed in [2].

These schemes are inherited from Weber's law of just-noticeable differences (JND). In this context, the law stipulates that the haptic signal is considered perceivable and need to be sent only if the relative difference between two successive haptic stimuli exceeds the JND threshold as illustrated in Figure 3.6.

With regards to the Hardware influencing quality, it relates to the performance aspects of the end/rendering devices used to generate the multi-modal scene as perceived by the end-user. Eventually, the taxonomy will cover new emerging modalities such as olfaction and gustatory (digital smell and taste) when their hardware interfaces reach the mature level of manufacturing. Rendering hardware implies that the hardware device must embed technologies to make sure that the user synchronously perceives the immersed media, i.e. with proper spatial and temporal relationships. The rendering hardware needs to convert the digital media signal into a form perceivable by humans. Examples include a humanoid robotic, haptic engine, and olfactory display, 3D touch

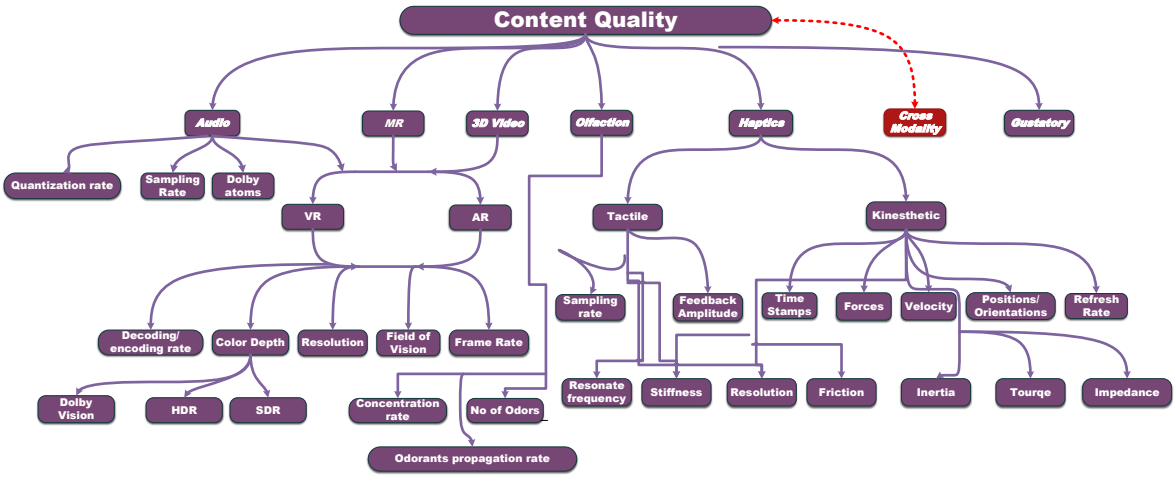


Figure 3.5: Content Influencing Quality

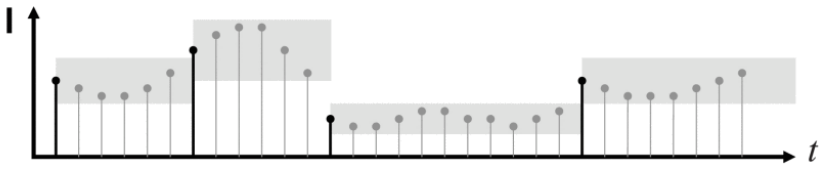


Figure 3.6: Illustration of the perceptual deadband principle. The size of the perceptual deadband depends on the stimulus intensity I. Samples falling within the perceptual deadband are considered as perceptually unimportant, hence can be dropped

and Omega. X devices, Head mounting display (HMD) for VR and AR, etc. As seen in Figure 3.7, we have focused on the haptic and VR interfaces for both (Tactile and kinesthetic) as they are the newest modalities to be effectively used in DT systems to interact with other humans and machines. As such, they hold very stringent requirements in terms of feedback loops which can impact the stability and transparency of the application. In a master-slave teleportation system, a system is stable iff error remains small when executing various desired trajectories/velocity even in the presence of some moderate disturbance and network parameters changes. Therefore, while interacting with a real or virtual environment, there exists a bidirectional exchange of information (position/velocity, force/torque) between the master and slave sides. Thus, we have a control loop between master and slave, which may get destabilized even because of a small amount of delay (transmission delay) or latency variations. In the presence of delay, hard objects will be perceived as softer ones. In order to avoid this situation, the stability of the control loop needs to be maintained. Stability-fidelity trade-off is commonly agreed that stability and fidelity/transparency of a haptic interface are contradicting requirements, a problem yet to be solved by the research community[132]. With regards the transparency factor, teleportation or remote interaction is said to be transparent if there is a perfect match between master and slave position and force signals. Stability and transparency are important features of a haptic teleoperation scenario, as well as the primary emphasis of system control techniques. A fully transparent system is one in which telepresence is a faultless and seamless experience. To attain transparency, the system also demands to be stable for an expected (bounded) performance of the operator and the remote environment. An unstable master-slave system could cause damage to the device or harm the human operator. Therefore, stability is guaranteed by reducing the range of dynamic impedances. Hence, the stiffness of the object cannot perceive as real and the transparency is regarded. Accordingly, there is a conflict between stability and transparency, hence a compromise requires to be rendered. Full transparency means that the operator feels as he/she manipulate the remote environment directly, hence the

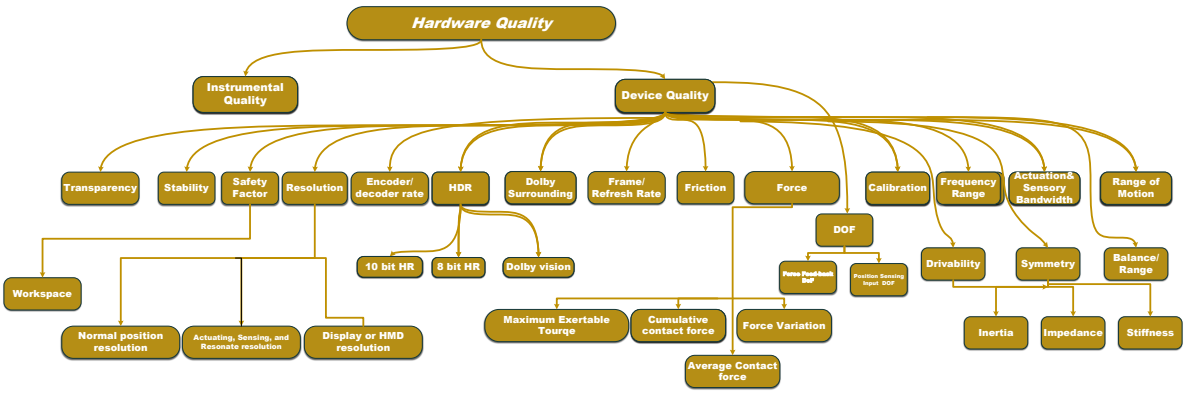


Figure 3.7: Hardware Influencing Quality

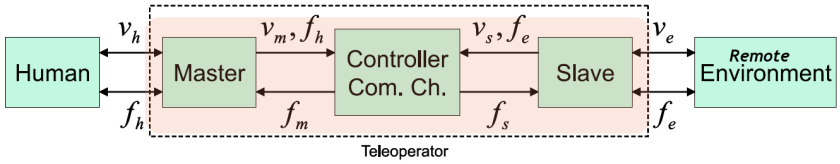


Figure 3.8: A Model of a Master-Slave Teleoperation Architecture

impedance should be felt as same:

For full transparency and by analyzing the two-port network model of MST systems as illustrated in Figure 3.8 The remote environment, the slave's velocity V_e and interaction force F_e are related by impedance Z_e:

$$F_e = Z_e.V_e \tag{3.5}$$

$$V_e = V_h \tag{3.6}$$

$$F_e = F_h \tag{3.7}$$

$$Z_t = Z_e \tag{3.8}$$

by applying the Hadamard matrix:

$$[F_h V_h] = [h_{11} h_{12} h_{21} h_{22}][V_e - F_e] \tag{3.9}$$

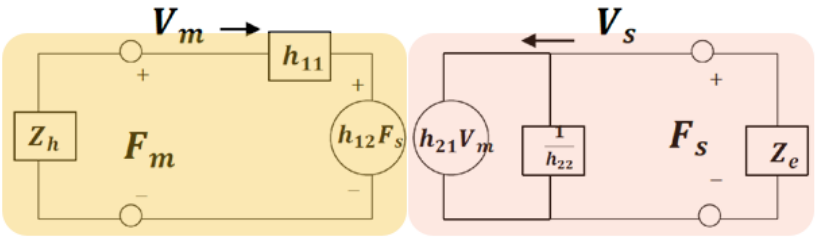


Figure 3.9: Two-port Hybrid Network Model of Master-Slave Teleoperation System

by eliminating F_e and V_e, we have:

$$F_h = (h_{11} - h_{12}Z_e)(h_{21} - h_{22}Z_e)^{-1}V_h \tag{3.10}$$

Hence:

$$Z_t = (h_{11} - h_{12}Z_e)(h_{21} - h_{22}Z_e)^{-1} \tag{3.11}$$

Recall that full transparency means Z_t=Z_e, hence h_{22}=0.

For stability we analyze two-port hybrid network model, Figure 3.9. From the right port, we can get:

$$V_s = -\frac{F_s}{Z_s} = F_s h_{22} + V_m h_{21} \tag{3.12}$$

$$\rightarrow \frac{F_s}{V_m} = -h_{21}(\frac{Z_e}{h_{22}}) \tag{3.13}$$

From the left port, we can get:

$$\frac{V_m}{F_s} = \frac{h_{12}}{Z_h + h_{11}} \tag{3.14}$$

Combining both equations 3.12 and 3.14, we can calculate the loop-gain as:

$$G_e(s) = \frac{-h_{12}h_{21}Z_e}{(Z_h + h_{11}) + (1 + h_{22}Z_e)} \tag{3.15}$$

Consequently, a trade-off between transparency and stability can be illustrated as below:

$$\uparrow h_{22} \Rightarrow \downarrow |G_e(s)| \Rightarrow \uparrow Stability \tag{3.16}$$

$$\uparrow h_{22} \Rightarrow (Z_t \neq Z_e) \Rightarrow \downarrow Transparency \tag{3.17}$$

To resolve the MST stability and reliability contradicting objectives, a control architecture referred to as Model-Mediated Teleoperation and sometimes called Telemanipulation (MMT) is proposed [133]. When combined with the Perceptual Deadband-Based (PD) haptic data reduction approach (MMT+PD) [133, 134] can guarantee the transparency as well as stability of the teleoperation system even with the presence of given communication impairments such as slightly low data rate, end-to-end latency, and jitters. More details about teleoperation control schemes are discussed in chapter 5

3.3 Subjective Influencing Quality

Optimizing the objective factors of a DT application does not evidently lead to the fact that the user of this application is satisfied. As the DT application embodies haptic and VR modalities, users might find the application not easy to be used or experience negative aspects such as stimulation sickness i.e., the user might feel dizzy a condition referred to as cybersickness. Consequently, it is very important to take into consideration the subjective experience while evaluating DT applications and services. As a matter of fact, the authors in [135] reported that there is no comprehension model of human behaviour that exists as a result of its multi-dimensionality and large individual randomness, variability and unpredictability.

Subjective Quality Influencing Factors can be described as dynamic, context-dependent and fuzzy; deriving from a wide range of potential benefits users may stem from a product. Hence, user characteristics change the performance of entities and their perception. Those influencing factors are also compromising what is referred to as Key Quality Indicators (KQI). We break down the subjective categories into three parts namely physical and mental factors, as well as usability and degree of perception. This is an important evaluation category for the overall quality of the DT application as it reflects the emerging state as well as the individual's stream of perception and interpretation of one or multiple stimuli which is referred to as the user experience (UX) [136, 124].

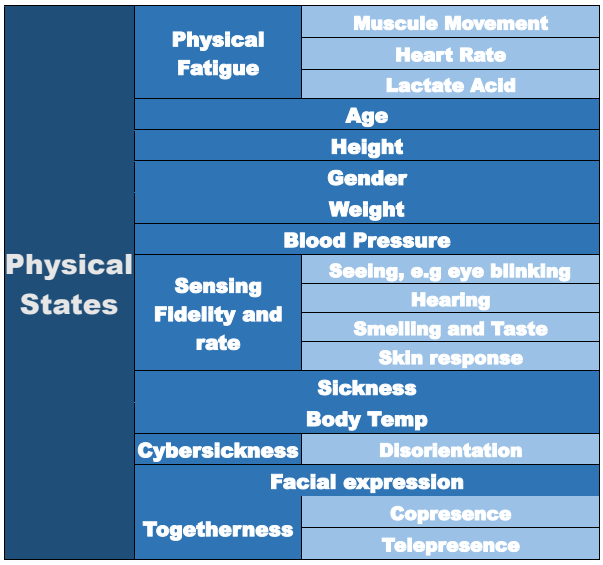


Figure 3.10: Physical State Factors

Physical State influencing factors, figure 3.10, used to describe the biological measures of the users such as height, weight, age, gender, endurance, etc. It also includes some Physiological aspects of the user such (fatigue, Cybersickness, heart rate, disorientation, postural instability, sweating, nausea, drowsiness... etc.). Physiology is involved with the detection and recording of signals provided by the body as it performs its normal tasks. Several research studies [107, 108, 137] have gone into the analysis of neurophysiology and the signal ascribed to the brain, as well as their association with physical reactions. Although we do not completely comprehend the distinctions of cognition and the conscious neural functioning of the brain, we do comprehend brain mechanisms and their interconnections with physical systems in other areas of the human body. This involves mechanisms such as increased respiration rate when excited or increased temperature when aroused, stressed, and embarrassed. By obtaining these signals and processes, insights into facets of the human condition can be captured. Physiological aspect can be measured using Electromyography (EMG), Electrocardiography (ECG) tests, thermometers, biomedical sensors such as smart wearable devices, etc. Biometric devices, such as the Fitbit, Delyses, Emotiv helmet, the E4 armlet, enable the collection

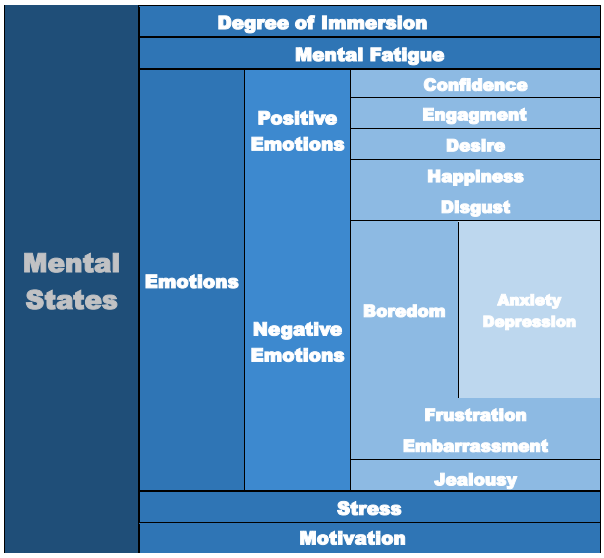


Figure 3.11: Mental States Factors

of biomedical data such as heart rate, heart rate variability, temperature, biomotion, respiration rate, electrophysiological, and electrodermal behaviour. These variables, along with contextual judgments, have been proven to align with a user's QoE in [107] and [108] and serve as estimators to quantify overall satisfaction or interaction comparisons with a stimulus or technology. Furthermore, studies have been undertaken to determine the viability of eye monitoring in assessing a user's understanding of an area or technology [137]. On the other hand, mental state factors, figure 3.11, reflect the state of the user's mind and emotions through observation and user feedback with and/or without direct measurements. Observation can assess the psychological behaviour of users, such as stress, without hindering the user's movements by including measuring devices. Examples of this category are positive emotions such as (enjoyment and appreciation) and negative emotions such as (fear and phobia). It should be noted that the deployment of none invasive electroencephalogram (EEG) test is very effective to measure the electrical activity in the user's brain. Consequently, it is very useful to quantify the user's mental influencing factors and to gauge emotional arousal.

In terms of QoE research, an emphasis has been placed on capturing processes as-

sociated with the central nervous system (CNS), the autonomic nervous system (ANS), and perceptual mechanisms, like brain, heart, and eye measurements. As depicted in Figure 3.12. The ANS is a branch of the peripheral nervous system (PNS), which relays information and sensations from the CNS to the rest of the body. The ANS is concerned with monitoring the human body's unconscious or spontaneous reactions. Each of these systems has its own set of signals, and by recording these signals with biometric scanners, stakeholder in the field of QoE can get insight into a subject's emotional, mental and physical states.

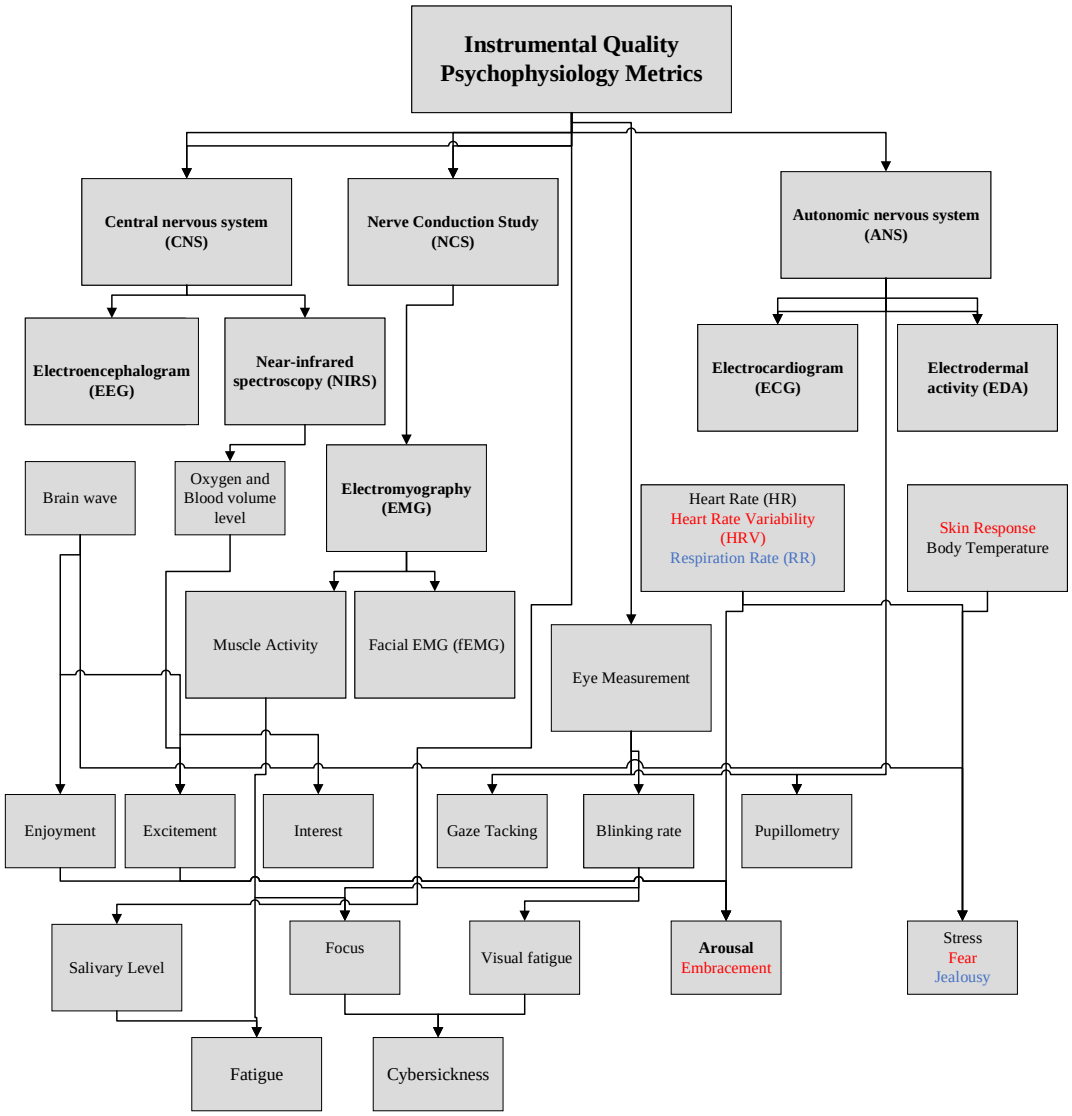


Figure 3.12: Psychophysiology Metrics

With regards to QoE research, the sympathetic nervous system (SNS) subset is of particular importance. This subset is governed by physiological responses that are commonly correlated with variations in an emotional state, such as growing body temperature and heart rate during a situation of anxiety or excitement. Moreover, this subset has the ability to affect the heartbeat, skin conductance, blood pressure, and eye dilation. Previous QoE article has used recording these kinds of bodily metrics to help assess im-

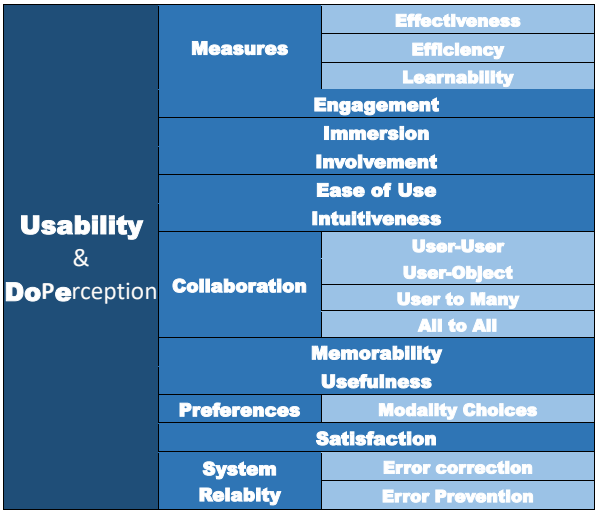


Figure 3.13: Usability and degree of perception

pacts on emotional and physical arousal [107]. Arousal is a psychophysiology condition characterized by the stimulation of sensory organs to the moment of perception.

As regards of the degree of perception, reception denotes the process of accepting an energy stimulus from the external real/virtual world and translating this signal into a form that can then excite perception. Perception denotes the encoding of the received energy into a neural message which is then transferred to the appropriate brain center which can achieve cognition. Cognition demotes the assimilation of the information contained in the neural message concerning data extracted from the memory, thus achieving identification based upon some aspect of past experience. Immersion describes the degree to which the media environment sub-merges the perceptual system of the user in computer-generated stimuli. The more the system generates stimuli from the physical world, the more the system is considered to be immersive [138]. Context influencing factors are the metrics that incorporate any situational property to illustrate the user's environment in terms of physical, temporal, social, economic, task, and technical characteristics[1, 76]. For instance, the user location and time of the day are an example of physical and temporal context respectively whereas the costs of device/system/service are examples of the economic context. Perceiving the experience lonely or collaboratively is an example of

a social or task context. The technical and information context describe the correlation among the user of interest and other related systems and services including applications, networks rendering hardware, or extra informational artifacts. For instance, your DT recommends[139] a related multimedia e.g specific movie on a specific device over a 5G network.

Usability, Figure 3.13, is defined as the extent to which the application can be used by the users to achieve satisfaction effectively and efficiently. DT application designer has to take into consideration how well is the application features fit the user's need and context. For the sake of simplicity, we will go over some of the influencing factors that should be considered when designing such an application. Effectiveness measures the accuracy at which the users can complete the tasks of the application, Efficiency measures how fast the user can perform a given task or interact with the application, learnability is a feature that allows users to be familiarized with the application tasks and objectives, system reliability refers to how easily users can avoid errors and/or recover from those errors. Engagement, immersion, involvements, collaboration and intuitiveness are all the influencing factors that contribute to the user's degree of perception.

3.4 Deploying the Taxonomy to a Case Study

To achieve the design requirements of a QoE metric for a Digital Twin system, user testing would be conducted. The main methodology for designing experimentation with human subjects can be divided into the following [121, 140, 141];

- **Selecting participants**: based on the experiment's nature some questions must be addressed. Are the subjects will be placed into groups or not? some common practices will be to divide the subjects based on their gender, age, background, or expertise in the application. The study objectives are often used to decide the gender, age, and other features of the users. Further, to confirm the correctness of the study, each subject has to be interviewed. To do so, an invitation can be

broadcasted(via email, phone, web, etc) to whoever is willing to participate in the experiment however conflict of interest has to be taken into consideration to confirm that the subjects are not biased.

- **Sample size and population**: it is based on the expected confidence interval from the data. In addition, the experimental protocol plays an important role in determining the size of the population e.g, the exact number of trials per experimental scenario, number of sessions conducted by the subjects in a day, and the total period (time/days) the study will last. Another important point is to consider a wide range of demographics in human subjects (age, gender, handedness, anthropometric measures) as stimulus sensitivity threshold e.g., tactile or hearing sensitivities varies between individuals.

- **Within subjects or between subjects' study**: each type of these studies has its pros and cons. Within-subject denotes comparing results of different data sets of the same subject, while between-subject is based on comparing data sets of different groups typically each group is experimenting with different controlled conditions. The latter solves the learning-bias disadvantage but introduces individual variability bias.

- **Counterbalancing task and stimuli order**: refers to changing the sequence of the study's tasks if possible to reduce the subject's learnability. In the context of haptic, changing the order of the perceived test and original signals is important to avoid time order effects (TOE): which can be implemented by comparing 2^{nd} or control stimulus with a value slightly shifted towards the mean of all presented stimuli instead of 1^{st} stimulus. On the other hand, sequential stimuli presentation is very crucial to avoid Spatio-temporal attention effects. counter-balance stimuli/task order is not always beneficial to the study especially when the order of the tasks plays a significant role in determining some outcome of the study.

- **Selecting Parameters** For both technical and non-technical categories, application-

specific parameters for evaluation are selected. This provides a contextualized view of the QoE evaluation. The parameters can be divided into Objective parameters which refer to measurable and verifiable aspects of a thing or phenomenon, expressed in numbers or quantities, such as lightness or heaviness, thickness or thinness, softness or hardness. Subjective parameters denote the attribute, characteristic, or property of a thing or phenomenon that can be observed and interpreted, and may be approximated (quantified) but cannot be measured, such as beauty, feel, flavour, taste. Objective parameters can be observed through intrusive methodologies such as biological feedback or non-intrusive methodologies such as Network/Application parameters (e.g. QoS metrics) or performance metrics (error rate, time to completion, number of input.. etc.). Subjective parameters can be observed through Questionnaires, Rating scales, Annotations, Non-/verbal methods. They measure the following aspects:effectiveness, efficiency, satisfaction, enjoyment, and social presence. One of the popular used subjective methodologies is the Mean Opinion Score.

- **Independent and Dependent Variables**: denote defining what are the controlled variables and what are the response variables to be quantified. The dependent variable is the final QoE value of the DT application; this is the parameter that the experiment is attempting to calculate and or predict. On the other hand, the independent variables are those parameters of the application that impact the user's decision on perception and experience. For instance, the precise timing of the application's stimuli delivery has to be considered to avoid masking, sensory persistence, temporal integration, or habituation effects.

- **Consent**: an approval from each participant should be considered when conducting the experiment and after completing any subjective survey related to the experiment.

As we will show in the next chapters, we have strictly followed the aforementioned

guidelines in our experimental studies. The work in this thesis was based on recruiting a large number of subjects from the University of Ottawa and the University of Qatar with wide demographics considering different subjects' physiopsychological features. As such we confirm that we have achieved a high confidence rate for our results. Further, we have considered both within-subjects or between subjects' studies to assess the QoE of DT system in different scenarios. Furthermore, we have asked our subjects to do multiple trials per experiment throughout a different day/time period, keep in mind that we have randomly selected the order of the control scenarios in order to eliminate any bias in the results. Furthermost, we have and considered and conducted subjective, objective, and physiopsychological quality assessments. Lastly, in each experiment we have conducted, we took consent forms from the subjects and we let them quit the experiment at any time they want.

3.5 Summary

Since different multimedia systems, applications, and devices have different architectures which are specifically tailored to the needs of different users, there is currently no well-known QoE model that can be applied to suit all systems and devices. Motivated by the fact that Digital Twin is still in its fancy stages and its QoE parameters and metrics are not yet well established and defined, we present a taxonomy to assess and comprehensively capture the performance of the said system. The taxonomy is designed on a hierarchical based taking into consideration subjective, objective, and physiopsychological quality influencing and assessment factors. In the following chapter, the QoE for applications based on the DT paradigm is quantified and predicted using machine learning regression and classification-based algorithms. In both cases, the QoE value obtained from each assessment paradigm is compared to the ground truth QoE value reported by the participants in the case study and investigated.

Chapter 4

Experimental and QoE Model for DT Systems

This chapter presents a description of the technology and hardware used in this research as well as a summary of the different factors that went into the creation of the digital twin applications. The research methods and experimental techniques used to address the research question which is how to assess and predict QoE for DT applications based on incorporating the influencing factors inherited from the taxonomy described in the previous chapter are also presented. This chapter also shows the impact of using machine learning engineering to build a QoE model for DT applications that can automatically estimate their degree of perception under different controlled scenarios.

4.1 Devices and Technologies

4.1.1 Haptic Hardware

Touch haptic device developed by $3Dsystems$ [6] corporation was selected in our work. The touch haptic device improves productivity and reliability by allowing the most natural human-computer interface possible as well as enhancing the ability to solve problems

by touch and force feedback interactions. This haptic interface was preferred in this research over other commercially available interfaces since it provided a decent balance of functionality, service, and feature set at a competitive price. The Touch system has a high precision force feedback and detects motion in 6 degrees of freedom (DoF), resulting in the best, most natural 3D Touch sensation for any haptic VR application with 3 DoF kinesthetic and tactile feedback. Users can sense the stylus's point in the (X,Y,Z) axes and follow its orientation (pitch, roll and yaw). The Touch's lightweight architecture, small footprint, and USB connectivity allow for easy installation and use. The stylus has two buttons for ease of use and end-user customization. Not to forget the device's compact work-space and wrist-rest which maximize the user comfort and usability.

SPECIFICATIONS	TOUCH™
Workspace	-6.4 W X 4.8 H X 2.8 D in > 160 W x 120 H x 70 D mm
Range of motion	Hand movement pivoting at wrist
Nominal position resolution	> 450 dpi -0.055 mm
Maximum exertable force and torque at nominal position (orthogonal arms)	0.75 lbf/3.3 N
Stiffness	x-axis> 7.3 lb/in (1.26 N/mm) y-axis> 13.4 lb/in (2.31 N/mm) z-axis > 5.9 lb/in (1.02 N/mm)
Force feedback (3 Degrees of Freedom)	x y z
Position sensing/input (6 Degrees of Freedom) [Stylus gimbal]	x , y , z (digital encoders) [Roll, pitch, yaw(± 5% linearity potentiometers)]
Interface	USB 2.0

b)

Figure 4.1: 3D System Touch haptic device with technical specifications [6].

As shown in Figure 4.1, the Touch haptic device supports six degrees of movement provided by six-axis points. Functional limits exist for all degrees of motion. Once the user crosses one of these thresholds, he/she will experience an unexpected halt; this is the device's mechanical stop. As we will show later on in this chapter, we were capable to afford two touch haptic devices to ensure network collaborations for our testing scenarios.

We have also used the tactile glove industrialized at the [7] research work. The tactile glove has five Force Sensitive resistors (FSRs), each represents 0.5 diminished feeling area. The FSRs are used for assessing the amount of contact force for each finger. It also has five compact vibrotactile actuators implemented as shaft-less vibratory motors. As seen in Figure 4.2, these actuators are positioned on the fingernail locations. The motors are

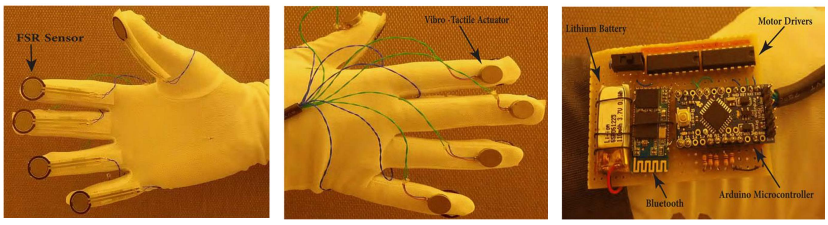


Figure 4.2: Vibro-tactile glove adapted from [7].

used to stimulate tactile sensations in case of gripping or 3D texture variations. A control board connects these sensors and actuators. The control board is powered by a Lithium Polymer based rechargeable battery supplying a DC voltage of 3.3V. The control board includes the following electronics:

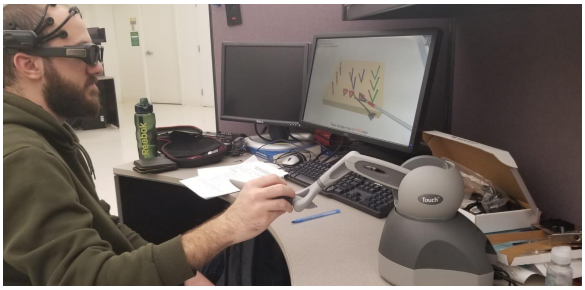
- (Arduino pro-mini) Microcontroller which represents the heart of the board.
- Bluetooth (Linvor-HC06) modem for wirelessly conveying the tactile sensations over the Bluetooth channel (with a baud rate 9600)
- Motor drivers (SN754410 H-Bridge)

The tactile glove works as follows: the FSRs sensors capture the analogue voltage in case of object's contact with the user's fingers. The measured voltages are converted to digital via 10-bit analog-to-digital converters. The Arduino micro-controller is responsible for the operation of the actuators. The amount of grip and contact force applied to the actuators determines the strength of the vibration. The tactile raw data is sent through the Bluetooth modem to the experiment workstation for further analysis.

4.1.2 3D Vision and Virtual Reality Equipment

Nvidia 3D vision [142] stereoscopic gaming kit was used in this research work. The kit contains LCD shutter glasses (SKU 942-10701-0003) that use a wireless IR protocol to communicate with the PC hosting the Nvidia Graphical Processing (GPU) card to allow stereoscopic vision for any Direct3D application, with different degrees of compatibility.

The wireless emitter plugs into a USB port and communicates with the driver software. Each pair of glasses has two lenses the function at 60 Hz frequency and alternates to produce a 120 Hz three-dimensional experience.



a) Subject 1: VR haptic Interaction with 3D glasses

b) Subject 2: VR haptic Interaction with Oculus rift-s HMD

Figure 4.3: 3D Vision and Virtual Reality Equipment.

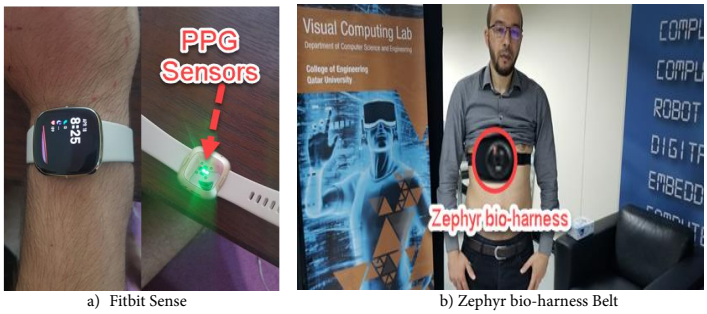
Oculus rift-s [143] is a virtual reality headset branded by (Oculus VR, LLC) and Facebook Corporations. The Oculus rift s was selected as the VR HMD for the experiment testbed (see Figure 4.3). The Oculus rift was chosen over other commercially available headsets because it offered a good equilibrium of support, performance, and feature set at an affordable price. The HMD headset, controllers, and oculus sensor track constellations, which are IR LEDs to translate user's movements in VR, are the three hardware elements that make up the VR system. The HMD is a true 6 DOF headset with inside-out monitoring that allows for a high-performance VR experience. The rift s HMD has LED-based screens with a resolution of 2560 x 1440 pixels and a refresh rate

of 80 HZ. It supports a native colour Space: sRGB/Rec.709 gamut, 2.2 gammas, D75 white point. The Oculus rift-s was deployed in our control scenarios to the impact of the incorporation of VR to the degree of immersion for the DT-based application.

4.1.3 Physiopsychology sensors

Fitbit sense and Zephyr bio-harness Belt

To build a biomedical data set for further machine learning analysis, the dissertation discussed in this study has focused on measures related to the autonomic nervous system and central nervous system, namely electro-dermal activity, heart rate, HRV, breathing rate, and brain signals KPIs. These measurements were selected because they have been proven to have close ties to a person's physical and emotional condition, and previous researches have shown associations between these signs and a user's subjective experience [144, 145, 146].



a) Fitbit Sense b) Zephyr bio-harness Belt

Figure 4.4: Fitbit sense and Zephyr bio-harness Belt to continuously measure Heart/Respiration and EDA metrics.

Electrodermal activity (EDA) or galvanic skin reaction (GVR), heart rate, heart rate variability (HRV), respiration rate, blood pressure, acceleration, brain signals, and temperature are all important metrics in the QoE physiopsychology research domain. In this thesis, EDA, heart rate, HRV, breathing rate, and EEG bio-signals are particularly relevant because they are considered good markers of physical and emotional feelings

such as arousal and stress in humans [110, 109]. The aforementioned heart, breathing, and skin-related metrics are captured while the users are experimenting using the Fitbit sense as well as Zephyr bio-harness belt, as illustrated in Figure 4.4.

The BioHarness 3.0 [147] is a non-invasive telemetry system for physiological surveillance of adults in the home, office, and other care environments. A chest strap and an electronics module that connects to the harness make up the unit 4.4.b. ECG, heart rate, respiration rate, body position, and movement are all stored and transmitted by the system. The BioHarness 3.0 can detect and relay single-lead ECG signals, which can then be processed by Bluetooth/USB qualified workstation. It should be noted that heart Rate, Posture, and Movement metrics are collected and transmitted by the BioHarness 3.0 during both sedentary and strenuous activity. The heart rate belt is marketed by zephyr bioharness technology and was selected because 1) it is easy to be adjusted providing comfortability for our subjects 2) it can wirelessly measure the ECG data such as the Heart Rate and R-R Interval with Bluetooth Smart (4.0). The technical specification of the belt sensors is as the following: 28 (Diam) x 7 mm (1.84 x 0.44 inches), with ECG, heart rate, respiration rate encoder/decoder, supplied by rechargeable lithium polymer with 3.7 volts and Bluetooth module with 115,200 baud, 8 data bits, 1 stop bit, No parity.

Similar to the zephyr bio-harness belt, we have used the Fitbit sense to measure heart rate and electrodermal activity. It should be noted that our research got approval from google to access the raw data of Fitbit sense measurements through a special API. Fitbit measures the heart rate blood volume pulses (BVP) which can be used also to measure the variability of heart rate. The sense is equipped with 2 photoplethysmogram (PPG) green LEDs and photodiodes positioned underneath the main body of the watch. These sensors measure the subject's heart rate with a high degree of accuracy.

In the case of our research and due to the inconsistency in time-to-task completion among all subjects, it was better to demonstrate comparative visual data of whole groups by normalizing the length of the experiment when presenting results. Hence we have

adopted the work in [108] by displaying the heart rate and respiration rate in normalized time slides and then storing them in the data set. To do so, we have divided both rates measurement into five frames; the first frame indicated the heart rate when the subject is rested i.e., before conducting the experiment test. The remaining four frames indicate the measurement of the heart rate when the subject has finished 25%, 50%, 75% and 100% of the experiment's task accordingly.

Skin conductivity is the parameter used to measure EDA, and the variation in conductivity over time is referred to as the galvanic skin reaction (GSR). The Fitbit sense is equipped with EDA sensors that basically track electrical changes in response to the sweat levels of the subject's skin. The sense measures EDA on the palms of the subject's hand as well as the fingertips. In front of the Fitbit sense, two silver electrodes are clearly separated and as the subject is asked to place his hand on them the sense will apply a low undetectable constant voltage to the skin and measure the conductivity of the skin in a range of 0.01 μ-Siemen to 100 μ-Siemens. The synchronization of EDA and HRV measurements enables precise monitoring of a users' variation in physiopsychology behaviour across time. It should be noted that EDA was interpreted visually and stored in our data-set in the same way as heart rate is, i.e., by using a normalized slide window to display change throughout the experiment time.

EMOTIV EPOC X 14 Channel Mobile EEG Brainwear®

In addition to the heart and skin-based physiopsychology metrics, we have used the EMOTIV EPOC X to capture the user's brain signals while conducting our experiment tests as illustrated in Figure 4.5. .

EMOTIV EPOC X is a portable and contextual human brain science platform that provides access to professional-grade brain data in a more user-friendly format. EMOTIV EPOC X was selected because of its compact size, non-invasive, wireless connectivity, and capabilities to capture all 14 electroencephalogram EEG signals namely ($AF3, AF4, F3, F4, FC5, FC6, F7, F8, T7, T8, P7, P8, O1, O2$). The whole-brain signals

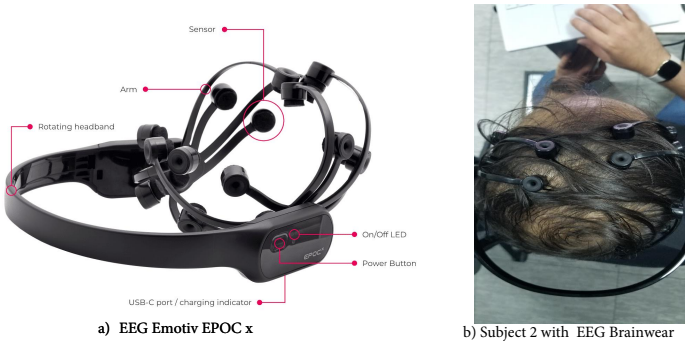


Figure 4.5: EMOTIV EPOC X to capture electroencephalogram siginals

are captured using saline-based electroplated. Electrodes connected to the scalp are used in EEG systems to detect electric potentials produced by the brain. EEG electrodes are attached to the skin with conductive gel, paste, or cream, which is mostly based on saline. Since certain metals corrode quickly, resulting in bad data, the right mixture of electrode metal and conductive paste is critical. Consequently, before the starting of our experiment session, we made sure that each of the Emotiv electrodes are wet using a saline-based solution to ensure high-quality recordings as indicated in Figure 4.6.a.

Figure 4.6.b shows a line graph of electroencephalogram data obtained during the experiment test. The graph depicts the signal amplitude in each electrode of the Emotiv EPOCx. Each of the electrodes is represented by a different colour on the line. Fortunately, The Emotiv X comes with an autopilot (Emotive Pro platform) for data processing – they take the lead and apply automated de-noising procedures or automatically generate high-level cognitive-affective metrics which can be used in order to get to conclusions much faster. The data collected from the EEG electrodes are transferred to the frequency domain using the prominent Fourier transform (FFT) and then filtered using a butter-worth filter with Hanning window as shown in Figure 4.7. We have deliberately shown the FFT of the frontal electrodes $AF3$ and $f7$ as they significantly correlate with the user's physical and emotional recognition. [148]. In the QoE regards, the *theta* band (4–8 Hz), *alpha* band (8–12 Hz), low *beta* band (12–16 Hz), high *beta*

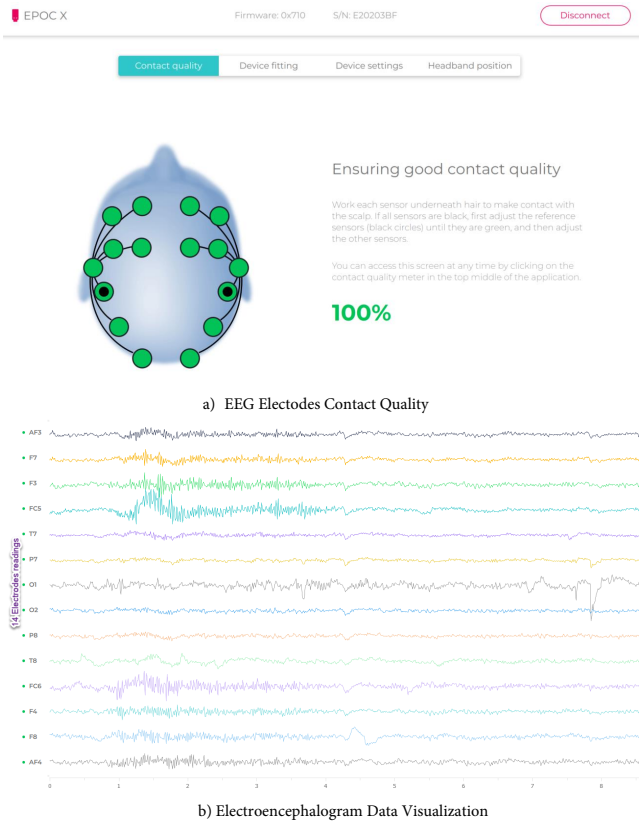


Figure 4.6: EMOTIV EPOC X to capture electroencephalogram siginals.

band (16–25 Hz), and *gamma* band (25–45 Hz) are the frequencies of interest. Frontal theta behaviour is reliably linked to the complexity of mental operations, such as concentrated attention and information uptake, processing and learning, or memory retrieval, according to studies [149, 150]. With rising mission complexity, theta frequencies become more prevalent. This is why theta is commonly correlated with mental workload or working memory processes in the brain. On one hand, there are many functional correlates of alpha waves that represent sensory, motor, and memory functions [151]. In the user's resting or meditation stage, one can see elevated levels of alpha-band strength during mental and physical relaxation. Alpha capacity, on the other hand, is diminished or suppressed during mental or physical activity while the eyes are open. For example,

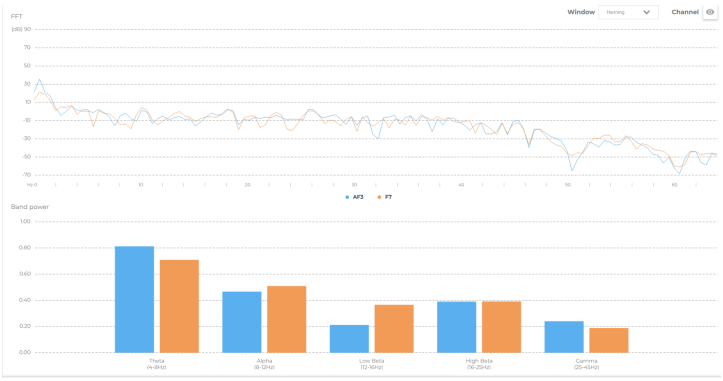


Figure 4.7: EEG FFT and Preprocessing for Electrodes AF3 and F7.

alpha suppression is a true signature of states of mental behaviour such as stress and engagement. Beta band activity is described as oscillations that occur between 12 and 25 Hz. This frequency is produced in both the posterior and frontal areas of the brain. Higher beta power is associated with active, busy, or nervous thought, as well as active focus [152]. As we will show later on. users will execute haptic movements, hence beta power in the central cortex increases, particularly when reaching or grasping is needed. It should be noted that till now gamma frequencies are currently the black holes of EEG science, as it is vague where these oscillations are produced and what they reveal. According to some researchers, gamma, like theta, acts as a carrier frequency for tying together various sensory perceptions of an object into a coherent shape, thus representing an attentional mechanism. Others claim that gamma frequency is a by-product of other neural processes including eye movements and micro-saccades, and therefore does not represent cognitive processing.

In this work, we relied on the Emotiv Pro algorithms to interpolate the aforementioned frequency bands to quantify the performance metrics of the user's cognitive states, Figure4.8. These metrics are namely stress, engagement, interest, excitement, focus, and relaxation which correlates with the psychological metrics address in the taxonomy presented in the previous chapter. A brief description of these metrics is provided below:

- **stress**: the level of satisfaction with the current challenge is measured by stress.

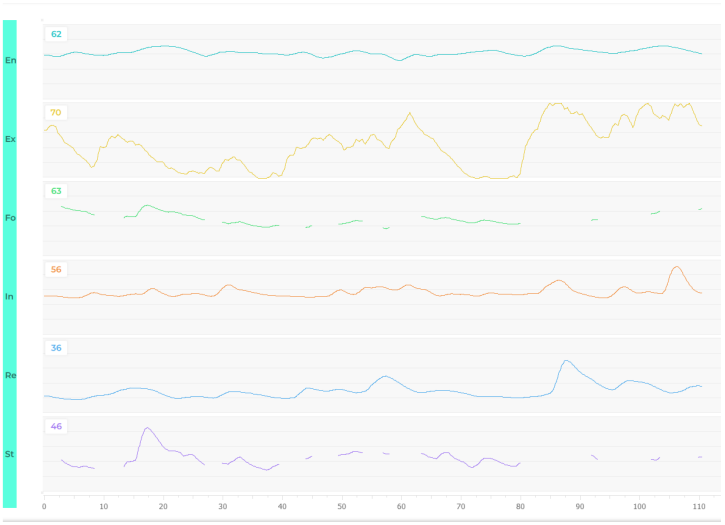


Figure 4.8: EEG KPIs.

Inability to complete a challenging mission, feeling frustrated, and expecting negative repercussions for failing to meet task criteria can all lead to high-stress levels. In general, a low to moderate level of stress can boost efficiency, while a higher level can be harmful to one's health and well-being in the long run.

- **Focus**: is a metric for paying specific attention to a single mission. The depth of concentration as well as the frequency at which attention moves between tasks are measured by focus. A high rate of task switching indicates a lack of concentration and distraction.

- **Relaxation**: is a metric for the ability to turn off and recover from periods of intense focus. Familiar and trained users with the experimental protocol as well as the experiment environment can achieve exceptionally high levels of relaxation.

- **Engagement**: is characterized by alertness and the deliberate focus of attention on task-relevant stimuli. It is a mixture of focus and concentration that compares with boredom and tests the amount of immersion at the moment. Increased physiological arousal and beta frequencies, as well as attenuated alpha frequencies,

describe engagement. The performance score recorded by the detection increases as attention, concentration, and workload increase.

- **Excitement**: It is exemplified by triggering of the sympathetic nervous system (SNS) which outcomes in a variety of physiological reactions such as pupil dilation, eye enlarging, sweat gland stimulation, heightened heart rate and muscle tension, blood diversion, and digestive inhibition. In most cases, the higher the rise in physiological arousal, the higher the detection performance score. Excitement detection has been fine-tuned to produce performance scores that represent short-term shifts in excitement over time spans as short as a few seconds.

- **Interest**: also known as valence which is a term used to describe the degree of attraction or aversion to current stimuli, atmosphere, or behaviour. Low interest implies a strong dislike for the task, while high interest signifies a strong affection for the task, and mid-range ratings suggest that the subject does not like or dislike it.

The EEG performance metrics are supplied at 0.1Hz for data export and we used the same sliding principle (rest slides, 25%, 50%, 75% and 100% of the task completion time (TCT)) to store the KPIs for each subject in our data-set.

4.1.4 3D Haptic Virtual Environment

3D Development Engine

Unity 2020 was used to build the haptic virtual reality environments for our experiments. Unity is a free game development engine that gives developers access to a wide range of tools for designing and developing games and virtual worlds. Unity comes with a comprehensive collection of tutorials that help newer developers quickly learn how to use the engine. Furthermore, all processes and functions scripting and coding are done in the C-sharp programming language, using a Unity-specific library. C-sharp is a good

starting point for developers since it is commonly used in other fields, such as windows, web, and Android development. Unity provides a high degree of support for the creation and delivery of third-party modifications and software, in addition to its ease of learning and use. Users who create their own Unity plugins and toolkits will share them on the Unity store for others to use in their own projects. This enables for the use of a wide range of resources without the need to design these tools for each project from scratch. In addition, the unity engine is compatible with third-party toolkits namely OpenHaptics® Developer Edition SDK and Virtual Reality Toolkit (VRTK). The combination of the two SDKs allows us to append haptic and true 3D navigation to our testing application. The OpenHaptics® toolkit is based on the OpenGL® API, allowing for easier integration of OpenGL applications. Developers can use the OpenHaptics toolkit to complement existing OpenGL code for defining geometry with OpenHaptics functions to model haptic material properties and stimulations including friction and stiffness. On the other hand, VRTK [153] is a third-party toolkit that includes scripts and assets to help us building and executing a virtual reality world in Unity. VRTK has a variety of simplified modules for locomotion and gestures including grasping and touching. It likewise included augmented reality user interface features as well as physics-based assets including gravity-affected hinges, rotators, and buttons.

4.2 Digital Twin Architecture

The main objective of our research is to investigate the effects of the objective, subjective, physiopsychology, user experience, and usability taxonomy of measurements on the QoE while using DT applications such as telemanipulation and remote tactile applications. As a result, it was critical that the program be designed in such a manner that it mirrored a real-world example of this sort of environment, such as one that could potentially be used for Digital twin purposes. Consequently, we have designed two main applications. A brief description of each application is provided as follows:

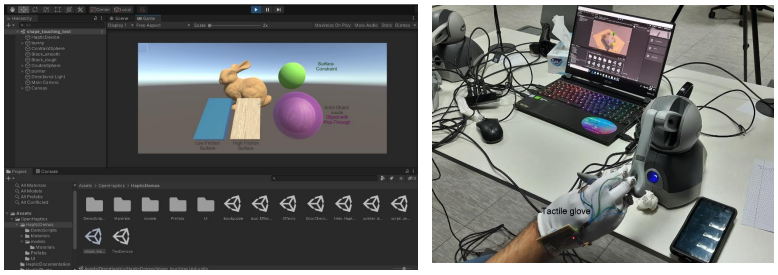


Figure 4.9: Tactile Application: A user can interact with different material probes and texture and receives and tactile feedback.

- *Remote tactile application:* Figure 4.9 shows a snap-shot of the tactile application used in our experiment. As can be seen, users can interact with the haptic VR environments using both the tactile glove and the Touch haptic device to switch between different 3D surfaces and objects. Upon moving the virtual tool that represents the master (styles tip) user can feel and experience different tactile touch stimuli (shape, friction, stiffness, and roughness) of the textures and materials. From now on and for the sake of simplicity, we will refer to the remote tactile application by (RTA).

Figure 4.10: Haptic VR tele-game: two user are competing to finish the rings throwing game while manipulating the VR objects and feeling the kinesthetic and physics feedbacks.

- *Telemanipulation Virtual Reality with kinesthetic and physics feedbacks*: Figure 4.10 demonstrates the second application in our experiment which denotes a net-

worked virtual game of the real-life Muckers, also known as ring toss game. As can be seen, the goal for each subject is to grab and place the rings at each correspondent coloured post using the Geomagic touch which is six (DOF) haptic devices. Two Geomagic Touch haptic devices were used in our platform as they are designed with a linkage-based system, which itself involves of stylus attached with a robotic arm. The robotic arm can generate the force at the stylus tip by tracking the appropriate position of the stylus. The subject should remedy that by using the force feedback, the weight of objects, and the 3D graphics to apply their judgment in placing the rings again to another pole. From now on and for the sake of simplicity, we will refer to the remote fame application by the term of ($Muckers$).

In both tactile and kinesthetic applications, we have added some auxiliary features i.e., auditory and visual support for the user's actions during the experiment. To enhance the user's immersion, audiovisual feedback was produced in the form of lighting the objects of interest when the user presses one button of the Touch styles device as well as a 3D sound when a contact is made. In addition, and to increase the usability of the testing software, several guidelines were also placed in bold text within the haptic virtual environment to instruct the subjects during the experiment without the intervention from the invigilator.

4.2.1 Experiment framework and protocol:

The experimental framework we used to conduct the experiment is depicted in Figure 4.11. Photon [154] is a software development kit (SDK) that can be integrated as a third-party asset for the Unity Engine to allow programmers to add collaboration i.e., multiplayer and togetherness features to their application. Photon also includes communication SDKs (PUN, BOLT and QUANTUM) by which the server for the Immersed haptic VR environment could be hosted locally, hence reducing undesirable (uncontrolled) network side effects such as latency and jitters. The Network Emulator for Windows

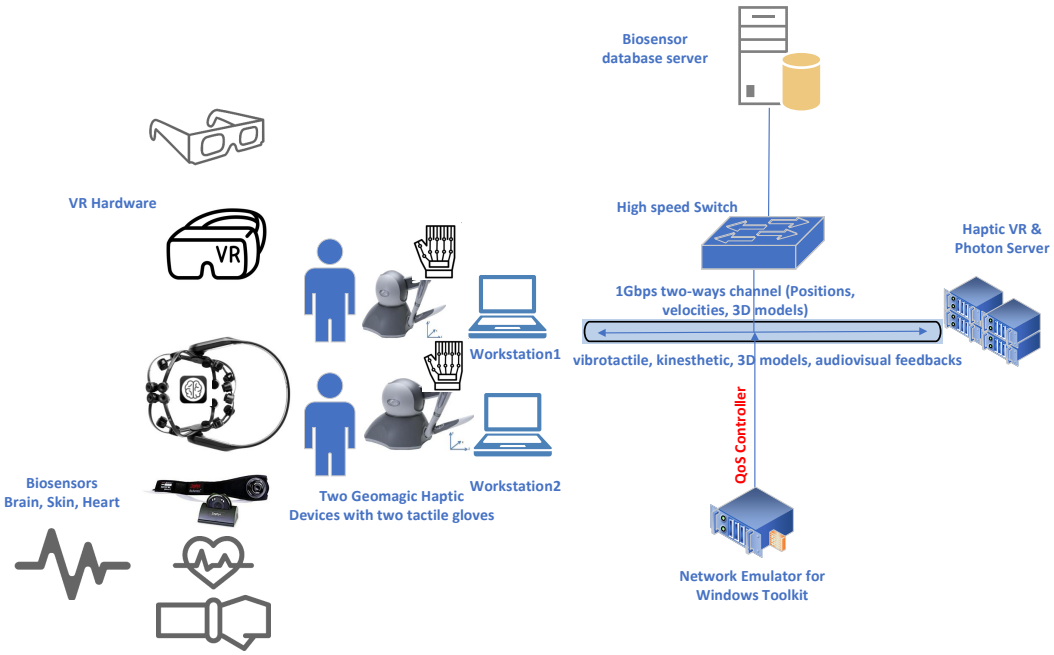


Figure 4.11: The Experimental Framework.

Toolkit (NWT), figure 4.12, was hosted in a server that incorporates a software-based model to emulate a multiplicity of network conditions by varying the QoS attributes that include but are not limited to latency, incoming and outgoing throughput, reordering of packets, a specified degree of packet loss, and error propagations. In addition, it offers flexibility in filtering network packets according to IP addresses or protocols for instance TCP, UDP, and ICMP. Consequently, we have used it because of its capability to simulate a poor network connection by degrading the QoS metrics, e.g. injecting artificial delay, packet loss between the VR server and the subjects' workstation that runs the immersed haptic VR application.

As can be seen in Figure 4.11, the experiment setup was based on a client-server architecture. On the server-side, a haptic virtual environment (HVE) instance was running, while subjects interacted haptically with a client-side mirror of (HVE) on their workstations. The server, which ran the exact instance of Unity and the same environment as the user but without the VR components, managed the HVE machine state and mirrored

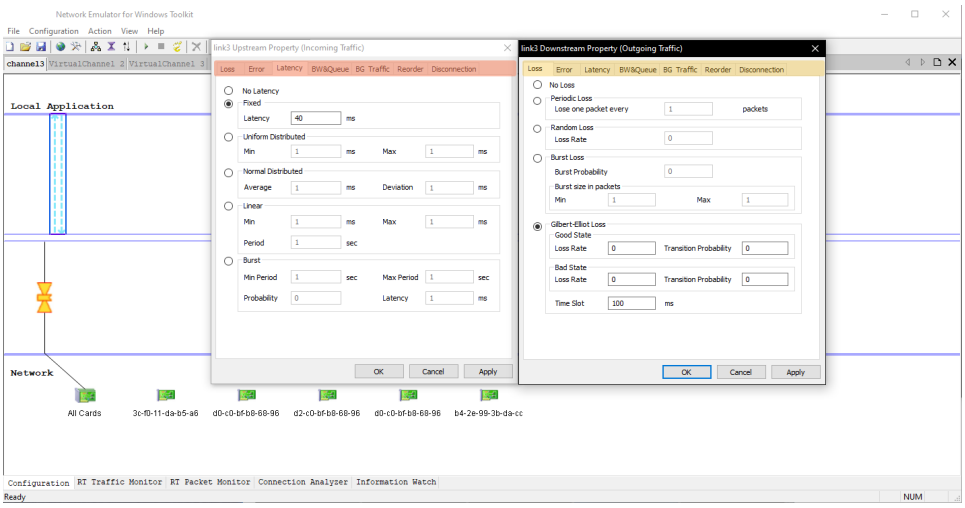


Figure 4.12: Network Emulator for Windows Toolkit.

it on the client-side workstations. **This is how a pseudo-digital twin resemblance was accomplished**. More specifically, The server provides the remote HVE model updates that reflect a real-world experience of teletouching and telemanipulation. When the subject interacts with a local model of the HVE, an update of (positions, speeds, and 3D models) is transmitted to the server via a remote process call (RPC) over UDP protocol which informs the server to execute a process (tactile, force, VR feedback). The server will then handle the necessary process and instruct the user's HVE to adjust its state to match the server's model. An emulation of network impairments was accomplished by controlling the communication between the subject and the server such as delaying the RPCs using NEWT. With regards to sound updates, RPC calls were also used to handle audio in Unity. When a user's action needs a sound to be played, the call to start the sound will be sent via RPC, thus the server would then ask the environment to play that sound. This permitted sounds to be delayed under the same latency conditions as the HVE processes and state updates. The sound was purposefully delayed in this manner in order to be harmonized and synchronized with the visual and haptic delays. The workstations used in the experiment were from ALIENWARE AURORA and GIGABYTE AERO 15 - 15.6" 4K/UHD OLED -running Windows 10 with Intel I7 3.4 GHz, 32GB

of RAM, and a GPU from Nvidia RTX 2080TI for smooth and buttery 3D rendering. The server's machine had identical specs and was also running Windows 10 with a clone of Photon's Server program and a process of Network Emulator for Windows Toolkit (NEWT).

The experiment protocol is as the following:

1. Subjects were guided to wear the biosensors instruments including the Emotiv-X, Fitbit sense, as well as Zephyr bio-harness belt.

2. The invigilator of the experiment runs the required software to capture the EEG and ECG biomedical data while making sure that 100% contact quality is achieved.

3. Subjects are provided with a training session to get familiar with the experiment's types of equipment, e.g. haptic and VR devices, as well as the RTA and muckers applications.

4. Subjects are asked to try the reference scenario of each haptic VR application (RTA and muckers). It should be noted that the reference scenario of each application is intended to achieve the best performance and stimulus experience. Therefore, no network impairments are applied in these scenarios with a 1Gbps bilateral communication channel as well as the 3D models of HVR are set to be displayed with a 4K ultra-high definition resolution i.e., (3840 x 2160 pixel).

5. Subjects are asked to fill the experiment survey after trying the reference scenario of each application. More details about the experiment's questionnaire are provided shortly.

6. After taking a rest period of 5 minutes, subjects are asked to try three controlled scenarios for each application.

    - *Control scenario one* which represents impairing the communication channel between the server and the subjects' workstations using NEWT with 25 ms

latency, 30% drop in the throughput, and packet loss rate of 0.1 %. The resolution of the 3D objects is set to be Full-High definition (FHD), i.e., 1,920x1,080 pixels.

- *Control scenario two*: which demonstrates impairing the communication channel between the server and the subjects' workstations by injecting 125 ms latency, 60% drop in the throughput, and packet loss rate of 0.1 %. The HVE 3D objects are displayed with an HD resolution, i.e., 720P.
- *Control scenario three*: which resembles impairing the communication channel between the server and the subjects' workstations by injecting 225 ms latency, 80% drop in the throughput, and packet loss rate of 0.1 %. The resolution of the 3D objects is set to be Standard Definition (SD) i.e., 576p.

7. After completing each controlled scenario, subjects are asked to fill the experiment's survey by reporting their experience of the control scenario against the reference scenario. In case of ambiguity and hesitation, it should be noted that the subjects can recall the reference scenario at any time. Further and in order to avoid any bias in the results and to reduce the subject's learnability effect, the order of the tested control scenarios are randomly selected. In other words, the subjects are unaware of the parameters of the controlled scenario for each application and their sequence. Furthermore, the QoS impairments values chosen for the control scenarios were based on pilot examinations to gauge subjects' reactions to such disturbance. For instance, the latency values are to be set to be 20 ms, 100 ms and 300 ms. In our pilot experiments, we discovered that users did not notice or experience delays of less than 25 ms, despite the literature review indicating that delays of much lower duration would be perceivable, i.e., 1ms. This prompted us to expand our latency thresholds beyond that. When we discovered that over 25 ms of latency was already being reported as appropriate, we agreed to measure excessive amounts of delay at 100 ms intervals.

8. While trying any of the controlled and the reference scenarios, the EMOTIV-X, Fitbit sense, as well as the heart belt are set to capture biometric of the subjects in order to keep track of any variation in the subject's physical and emotional state during the full time of the experiment. The biometric data are stored in the database server for further analysis.

Given the unique pandemic case presented at a worldwide level, the imposing "stay-at-home" order brought down by the Canadian Federal Government has forced the University to administer strict only procedures for all of its on-campus activities. This has negatively impacted the sample size in our experiment which demands a physical attendance of the subjects to be in the loop. Consequently, part of the experiment was done at the University of Qatar in order to recruit a total of 45 subjects in our test. Keep in mind that with the COVID-19 situation, subjects have voluntarily participated in the experiment. We confirmed that all given consent forms had been signed by the subjects before we initiating the experiments. Moreover, the research was approved by the University of Ottawa and the University of Qatar while considering the public health precautions such as wearing the mask and physical distancing. Most of the subjects are graduate and undergraduate students and researchers from both universities with an age range of 10-59 years. Some of the subjects already have some haptic VR background while the reset has no experience in this domain. In all cases, we have given the subjects detailed instructions about the experiment scenarios and objectives. We did not record any data until each subject felt confrontable with the environment setup and tried each scenario at least twice. After finishing each scenario, each subject was asked to complete a survey as discussed in the next sections.

4.2.2 A Post-Experience Questionnaire

The survey in this study was used to build the ground truth of our dataset based on a five-point Likert scale using the MoS rating system, Appendix 8. More specifically, each

question is followed by two anchors with five points to be selected. Fundamentally, most of the 26 questions are followed by a five-point scale in which the users are needed to select the point that matches their level of agreement. We used a descriptive label to mark the extreme poles of the scale to help the user understand the question properly. Specifically, values ranging from "1" to "5" reflect various answers (e.g., "not at all" to "completely"). The questionnaire is posted online using the google forms platform (https://forms.gle/41mLKFppEimKoXF46).

The questions for each reference and control scenario of application are similar, with the exception of explanations to help subjects understand questions chosen specifically for the given application. The survey begins by gathering basic user information such as name, gender, age-range, nationality, physical/mental state, and prior HVR experience. The questions of the survey were designed to capture the QoE parameters related to each high-level influencing qualities in our taxonomy namely content, hardware, and network qualities, user experience and usability.

Questions about low-level parameters are posed first, followed by a general ranking for the corresponding high-level parameters. For example, (Q14) asks the subject to rank the synchronization between 3D graphics and haptic feedback; where 1 indicates (very poor) and 5 indicates Excellent. (Q15) tackles the stability of the application by asking the users to what experience the system responds precisely (low jitter) to your hand movement; similarly (Q16) concerns the latency impact on the application while (Q17) addresses the collaboration impact on the user experience. Following these four questions, a general query *NWTQ* is posed to score the high-level parameter network quality based on Q14-Q17. Therefore, the subject will have a better comprehension of high-level metrics, e.g, content quality, hardware quality, etc. Nevertheless, in order to minimize ordering effects, all questions were randomized for each individual user. Furthermore, questions were posed in either a positive or negative syntax, ensuring that subjects were not swayed into responding disproportionately positive or negative with any question. Finally, a rating score (0-100) for the overall standard HVR QoE is asked

to be reported with some constructive comments from the users.

4.3 Machine Learning Model

Machine Learning (ML) algorithms, when employed in the QoE modelling, use a series of observations representing the network state and the user's insight to derive inference rules that forecast the QoE value inevitably. To solve this modelling challenge, the right learning model and features must be chosen. In general, Figure 4.13 categorized the most

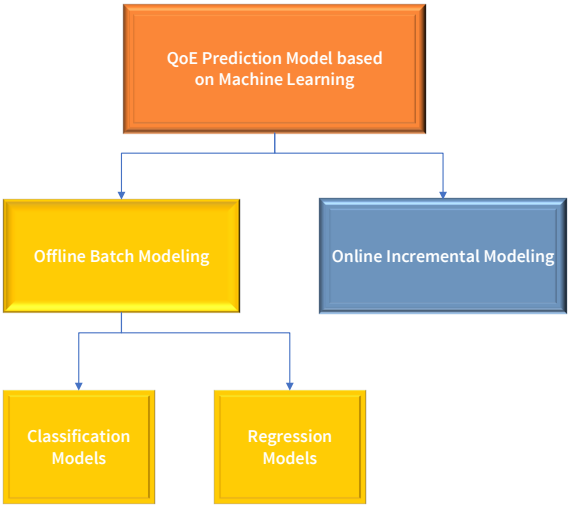


Figure 4.13: Categorization of QoE Prediction Models.

well-used ML approaches to predict QoE. Using offline ML, the model is conditioned on a series of collected data before being deployed on new data. However, online learning does not distinguish between model testing and model implementation. It keeps learning as soon as new data becomes available in order to better its efficiency. Moreover, offline batch models can be split into two classes based on the form of QoE values. The first party employs regression analysis to estimate QoE as a mathematical formula of the taxonomy parameters, i.e., subjective reported score and biosensors data. The second group predicts the QoE class using classification techniques. In our work, we have deployed

regression and classification techniques. In more detail and based on the best practice of classical machine learning engineering, we have utilized most supervised ML algorithms used in the literature [155]. These algorithms can be split into three kinds. The first type resembles the traditional ML algorithms such as Support Vector Machine (SVM) and Decision Tree (DT). The second one uses the boosting-based scheme such as (boosting based on DT). The third one uses bagging-based methods like Random Forest. In our work, we compare many of them (classic, bagging and boosting) in order to choose the right one based on our collected dataset. More specifically, for classification modelling, we used (SVM) and Lagrangian Support Vector (LSVM) as a linear-decision boundary algorithm, random forest (RF) as an ensemble classifier, k-nearest neighbours algorithm (k-NN) as an instance-based algorithm and Classification and Regression (CR) Tree) as a rule-based classifier and finally Bayesian networks (BN) as a probabilistic graphical model that demonstrates a set of variables and their conditional dependencies via a directed acyclic graph (DAG). For regression purposes, we have used generalized and linear regression as well as SVM.

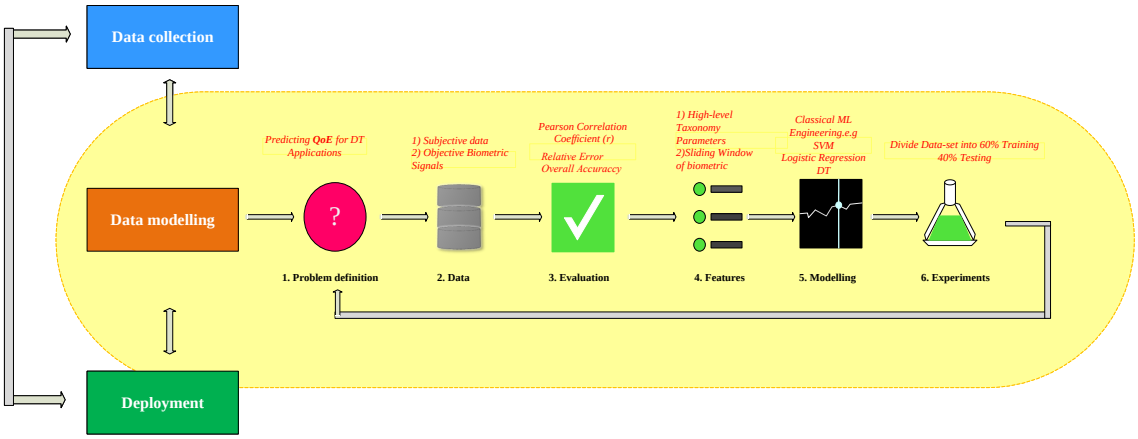


Figure 4.14: Description of the ML based QoE prediction model.

Figure 4.14 shows our framework as well as a description of the procedures used to predict the QoE for DT applications. It should be mentioned that Hinton and Krizhevsky [156, 157] demonstrated in 2012 that deep learning functions are generally superior to

other machine learning approaches for massive datasets requiring complex computation. Since we have linear classification and regression methods as well as a comparatively small dataset (data are collected from 45 subjects), deep learning analysis and inference are held beyond the scope of this work.

4.3.1 Data Set

The experiment's results were captured and assessed in a variety of ways. A post-test questionnaire was used to build the ground truth based on collecting subjective data on QoE. During the each session (reference and control) of the experiment, the EMOTIV-X, Fitbit sense, and ECG belt were used to collect objective physiological/ physiological data and biometric signals such as heart rate, HRV, breathing rate, EEG performance KPIs and EDA. Recall that we used sliding windows (frames) to represent and visualize most bio-metrics when the subject is rested i.e, (before starting the experiment scenario) and when he/she finished (25%, 50%, 75% and 100% of the scenario's tasks respectively). HRV (heart rate variability) is the difference in successive inter-beat periods (R-R intervals). The sympathetic and parasympathetic parts of the autonomic nervous system (ANS) are also involved in heart rate control (HR). It is well established that parasympathetic nervous system (PNS) action (vagal stimulation) lowers heart rate and raises heart rate variability. The sympathetic nervous system (SNS) function has an almost inverse impact on heart rate and heart rate variability, increasing HR while decreasing HRV. As a result, when we are at rest and thoroughly healed, our HR is lowest and our HRV is highest. When stressed, sympathetic nervous activity raises, resulting in a growth in resting heart rate and a reduction in heart rate variability. A high HRV at rest is generally desirable, while a low HRV is generally adverse. When in an active state, a lower relative HRV is normally preferable, whereas a high HRV is not favourable.

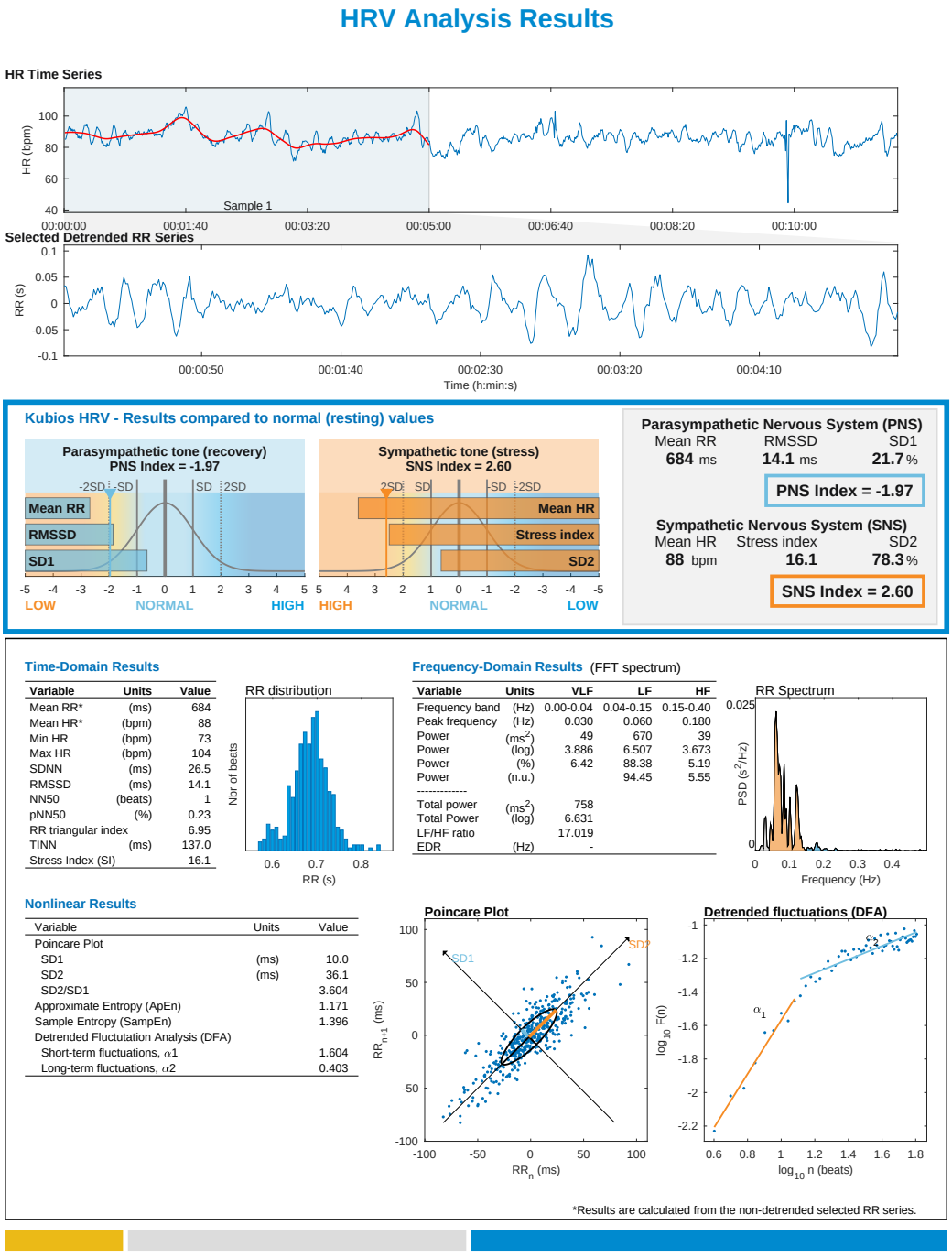


Figure 4.15: HRV analysis for subject 1 while conducting control scenario 2 of the Muckers application.

For the HRV analysis, we used the Kubios platform as shown in Figure 4.15. The metrics of interest are PNS and SNS indexes. Since the PNS index has a negative value while SNS index has a positive value of 2.60, this implies that the subject was under stress which can be confirmed by the depicted stress index = 16.1. Nevertheless, for the collected HRV from both data from the Fitbit sense and the R-R heart rate analysis, we used a binary format of 0 (indicating that the user was not stressed and 1 indicating otherwise). The data collected by the experiment was used to train the proposed machine learning model to predict the QoE for both DT applications. The data consisted of numerical and non-numerical features. Moreover, data entry errors and incomplete users' questionnaires were detected in the data set. Consequently, pre-processing steps were taken place in order to clean the abnormalities.

4.3.2 Data-Prepossessing

Several preprocessing tasks were needed to clean and formulate the data for the function extraction steps and to boost efficiency even further. Two anomalies reduction techniques have been applied to the collected data set, the deleting missing data technique or the imputing missing data technique. The former method is the simplest and most widely used technique. In our work, we have adapted this technique since the number of rows with missing data is small compared to the total number of records in the data set. The sliding window principle was used to prepare the barometric data from the EEG and ECG, EDA sensors. In addition and to make the dataset more manageable, we have converted (NQ), user experience (UX), and usability (US) from their MOS score (1 to 5) to a percentage vales using equation 4.1:

$$H_i = \frac{(H_{MOS-r} - 1)}{(N - 1)} * 100\% \tag{4.1}$$

Where H_i represents the taxonomy high-level parameter such as (NQ), H_{MOS-r} denotes the subject rating for that high-level parameter, and N reflects the maximum number of response choices in the MOS scale, which is 5. Consequently, the MOS score of 1 is

mapped to 0%, 2 is mapped to 25%, 3, 4, and 5 MOS scores are mapped to 50%, 75%, and 100% respectively.

4.3.3 Feature Selection

Feature selection is the process of representing each captured biometric as well as reported MOS score as a feature vector. Each entry position in a vector corresponds to a feature type extracted from the within the experiment biometric data and post-experiment survey results. Removing irrelevant features is an important step in data preparation for the following reasons: (1) achieving more accurate results, (2) faster model deployment, (3) and lastly, the generated model is easier to be interpreted. A total of 83 features for each user per scenario were collected in the dataset however only 53 features have been utilized to build our models. This is because there is a high correlation between some sets of questions in the survey and the summarized one. For instance, the first four questions are confounded with the fifth question which tackles the content quality. Moreover, the height and weight features were mapped into one feature which is the Body Mass Index (BMI). Nevertheless, Table 4.1 demonstrate the features we used to feed the ML to create the QoE prediction model for each DT system.

Table 4.1: Features used to develop the QoE model.

Features Name	Meaning	Type	Category
Age	Age of the Participant	Range	Subjective
Phyical_ST	Physical state before the experiment	Nominal	Subjective
Mental_ST	Physical state before the experiment	Nominal	Subjective
BMI	Body Mass Index of the user	Value	Subjective
L/R	Left or right handed	Flag	Subjective
BR_R	Breathing rate captured at rested condition	Value	Objective physiological
BR_25%	Breathing rate captured at 25% of the application scenario	Value	Objective physiological
BR)50%	Breathing rate captured at 50% of the application scenario	Value	Objective physiological
BR_75%	Breathing rate captured at 75% of the application scenario	Value	Objective physiological
BR_100%	Breathing rate captured at 100% of the application scenario	Value	Objective psychological
HR_R	Breathing rate at captured at rested condition	Value	Objective physiological
HR_25%	Heart rate at captured at 25% of the application scenario	Value	Objective physiological
HR_50%	Heart rate at captured at 50% of the application scenario	Value	Objective physiological
HR_75%	Heart rate at captured at 75% of the application scenario	Value	Objective physiological
HR_100%	Heart rate at captured at 100% of the application scenario	Value	Objective physiological
HRV	Hear rate variability	Binary	Objective physiopsychology
EDA_R	Electrodermal activity captured at rested condition	Value	Objective physiopsychology
EDA_25%	Electrodermal activity captured at 25% of the application senario	Value	Objective physiopsychology
EDA_50%	Electrodermal activity captured at 50% of the application senario	Value	Objective physiopsychology
EDA_75%	Electrodermal activity captured at 75% of the application senario	Value	Objective physiopsychology
EDA_100%	Electrodermal activity captured at 100% of the application scenario	Value	Objective physiopsychology
Haptic_FR	Haptic Familiarity	Value	Subjective
VR_FR	Virtual Reality Familiarity	Value	Subjective affected by objective QoS
CQ	Content quality	Value	Subjective affected by objective QoS
HWQ	Hardware quality	Value	Subjective affected by objective QoS
NWTQ	Network quality	Value	Subjective affected by objective QoS
UX	User experience	Value	Subjective affected by objective QoS
US	Usability	Value	Subjective affected by objective QoS
Enga_R	Engagement EEG KPI captured at rest	Value	Objective physiopsychology
Enga_25%	Engagement EEG KPI captured at 25% of the application scenario	Value	Objective physiopsychology
Enga_50%	Engagement EEG KPI captured at 50% of the application scenario	Value	Objective physiopsychology
Enga_75%	Engagement EEG KPI captured at 75% of the application scenario	Value	Objective physiopsychology
Eng_100%	Engagement EEG KPI captured at 100% of the application scenario	Value	Objective physiopsychology
EX_R	Excitement EEG KPI captured at rest	Value	Objective physiopsychology
EX_R_25%	Excitement EEG KPI captured at 25% of the application scenario	Value	Objective physiopsychology
EX_R_50%	Excitement EEG KPI captured at 50% of the application scenario	Value	Objective physiopsychology
EX_R_75%	ExcitementEEG KPI captured at 75% of the application scenario	Value	Objective physiopsychology
EX_R_100%	Excitement EEG KPI captured at 100% of the application scenario	Value	Objective physiopsychology
Fo_R	Focus EEG KPI captured at rest	Value	Objective physiopsychology
Fo_25%	Focus EEG KPI captured at 25% of the application scenario	Value	Objective physiopsychology
Fo_50%	Focus EEG KPI captured at 50% of the application scenario	Value	Objective physiopsychology
Fo_75%	Focus EEG KPI captured at 75% of the application scenario	Value	Objective physiopsychology
Fo_100%	Focus EEG KPI captured at 100% of the application scenario	Value	Objective physiopsychology
In_R	Interest EEG KPI captured at rest	Value	Objective physiopsychology
In_25%	Interest EEG KPI captured at 25% of the application scenario	Value	Objective physiopsychology
In_50%	Interest EEG KPI captured at 50% of the application scenario	Value	Objective physiopsychology
In_75%	Interest EEG KPI captured at 75% of the application scenario	Value	Objective physiopsychology
In_100%	Interest EEG KPI captured at 100% of the application scenario	Value	Objective physiopsychology
Stress_R	Interest EEG KPI captured at rest	Value	Objective physiopsychology
Stress_25%	Interest EEG KPI captured at 25% of the application scenario	Value	Objective physiopsychology
Stress_75%	Interest EEG KPI captured at 75% of the application senario	Value	Objective physiopsychology
Stress_100%	Interest EEG KPI captured at 100% of the application scenario	Value	Objective physio-psychology

4.3.4 QoE Estimation-Prediction Model building

Python based libraries and IBM SPSS modeler for machine learning have been used for data cleaning, statistical test, and model building. The collected data have been randomly split into two sets: A training set consisting of 70% of the data and a testing set consisting of 30% of the data. With the help of GRAPHLAB, CREATE, NUMPY, and Pandas python libraries, the classification models' performance is measured using the $F1_{score}$ as given by equation 4.2.

$$F1_{score} = \frac{2 * Precision * Recall}{Precision + Recall} \tag{4.2}$$

The regression models are evaluated using the Pearson's correlation (r) of model predictions and the mean absolute relative error re, equation 4.3.

$$re = \frac{|QoE_P - QoE_R|}{QoE_R} \tag{4.3}$$

Where QoE_R is the subjective QoE reported by the user and QoE_P is the predicted QoE of the given ML algorithm. It is worth noting that Numpy, SciPy, and other libraries were also used for data manipulations and obtaining results.

4.4 Results and Analysis

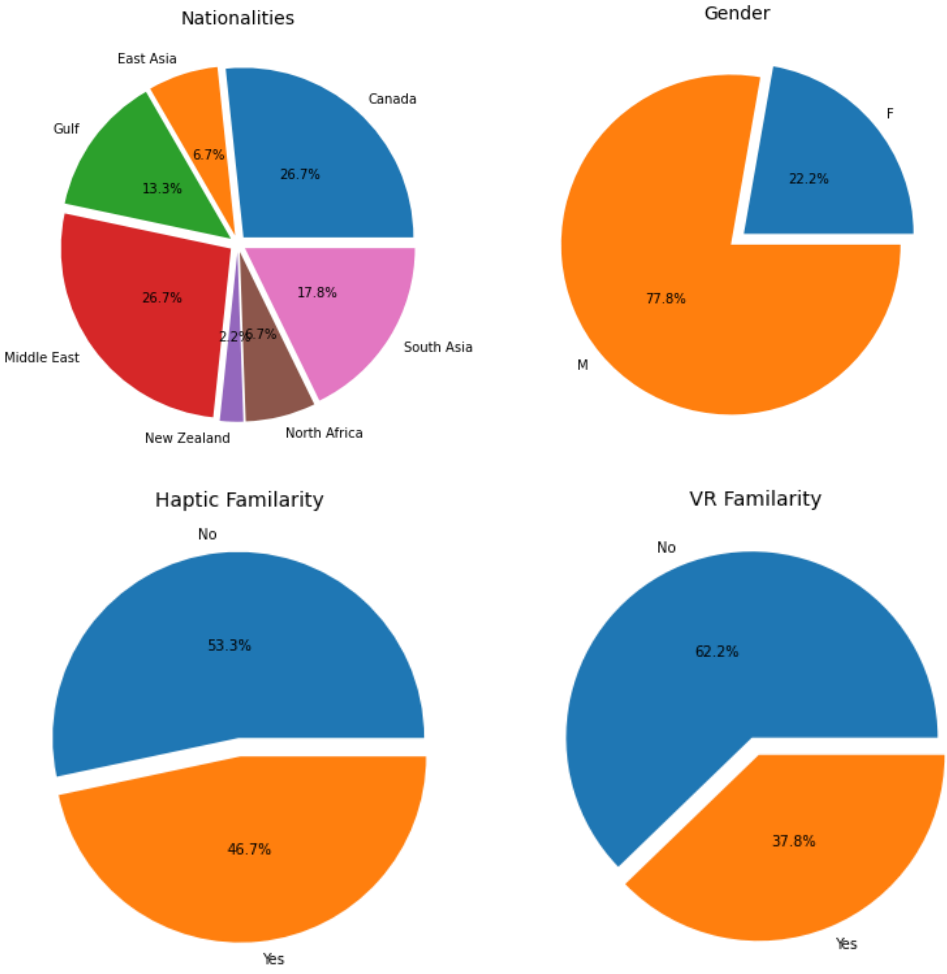


Figure 4.16: Information aspects of the subjects' demographics.

Before digging deeper into our ML evaluation criteria, it will be to the best of the reader's interest to take a look at the population's aspect of our subjects. As mentioned before the Covid-19 negatively impacts the participant recruitment for this work. Consequently, it was done based on convenience sampling. A visual breakdown of the population overview is illustrated in Figure 4.16. Out of a total of 45 subjects, 77.8% were males while 22.2% were females. As can be seen, the experiment included subjects from a wide variety of

nationalities. In terms of subject familiarity with the HVR technologies, 46.7% were familiar with haptic and 37.8% were familiar with VR. This might suggest that the findings could be skewed due to the sample size's tendency toward first-time novel use of haptic immersive technology. However, the majority of the subjects are unaware of the aforementioned technologies, 53.3% and 62.2% of the subjects are not familiar with haptic and VR at all, which confirms that most probably our results will not have any bias.

In the next section, we will discuss in-depth the results of our ML QoE modelling for both (RTA and Mukers) DT based applications. It should be noted that in order to build the best-generalized model that captures and forecast the users' features and ratings at all possible combinations and under different QoS conditions, we have blended the results from both the reference and control scenarios for each application to create more vectors before feeding it to ML pipeline. In this way, the selected ML algorithm will be trained to provide a crisp output of QoE under a variety of objective biological and subjective conditions. Furthermore, three evaluation cases are considered for each application's results. In the first case, only the subject's reported high-level scaled parameters, e.g, $(CQ, HWQ, NWTQ, UX, US)$ are used as inputs to build the QoE prediction model. In the second case, the QoE estimation model was built solely based on the collected *physiopsychology* features(such as heart rate, breathing rate, EDA, and EEG KPIs). In the third case, both subjective and physiopsychology results are used as inputs to build the QoE ML model. Nonetheless, Table 4.2 and Table 4.2 illustrates the descriptive statistics (min, max, mean,etc) about the features used to build the QoE model for each application.

Table 4.2: *Mukers*: Features Demographic for Both Subjective and physiopsychology data.

		Minimum	**Maximum**	**Mean**	**Std. Deviation**
Dependent Variable	QoE	10	100	58.19	27.503
	BMI	22	67	43.78	9.413
	BR_R	11	13	11.95	.594
	BR_25%	12	17	13.95	1.433
	BR)50%	13	17	14.95	1.366
	BR_75%	14	22	18.63	2.794
	BR_100%	14	24	20.68	4.211
	HR_R	65	75	71.05	2.894
	HR_25%	75	90	84.15	4.692
	HR_50%	75	98	89.30	6.554
	HR_75%	77	110	95.48	9.815
	HR_100%	77	121	101.83	13.718
	HRV	0	1	.31	.465
	EDA_R	1	2	1.52	.285
	EDA_25%	1	5	2.85	1.194
	EDA_50%	1	5	2.91	1.205
	EDA_100%	1	5	2.98	1.215
	Haptic Familarity	0	1	.41	.494
	VR Famlarity	0	1	.55	.500
	CQ%	25	100	63.48	34.353
	HWQ%	0	100	75.20	28.838
	NW%	0	100	70.48	29.353
	UX%	25	100	70.31	29.221
	US%	25	100	79.30	28.601
Covariate	Enga_R	0	1	.41	.081
	Enga_25%	0	1	.57	.230
	Enga_50%	0	1	.59	.301
	Enga_75%	0	1	.61	.332
	Eng_100%	0	1	.60	.343
	EX_R	0	1	.48	.136
	EX_R_25%	0	1	.53	.217
	EX_R_50%	0	1	.57	.307
	EX_R_75%	0	1	.57	.332
	EX_R_100%	0	1	.62	.324
	Fo_R	0	1	.58	.166
	Fo_25%	1	1	.77	.108
	Fo_50%	0	1	.70	.240
	Fo_75%	0	1	.68	.328
	Fo_100%	0	1	.61	.328
	In_R	1	1	.68	.069
	In_25%	0	1	.63	.270
	In_50%	0	1	.63	.316
	In_75%	0	1	.62	.338
	In_100%	0	1	.65	.365
	Stress_R	0	0	.33	.084
	Stress_25%	0	1	.44	.199
	Stress_75%	0	1	.54	.208
	Stress_100%	0	1	.62	.241

Table 4.3: *RTA*: Features Demographic for Both Subjective and physiopsychology data.

		Minimum	**Maximum**	**Mean**	**Std. Deviation**
Dependent Variable	QoE	30	100	76.11	20.729
Covariate	Height (cm)	137	190	169.73	11.884
	Weight (KG)	30	110	74.97	19.632
	BR_R	10	13	11.58	.671
	BR_25%	11	17	12.97	1.179
	BR)50%	11	18	13.78	1.719
	BR_75%	12	19	15.69	2.924
	BR_100%	13	22	16.49	3.288
	HR_R	60	75	70.63	3.629
	HR_25%	64	90	78.80	6.054
	HR_50%	75	95	83.35	5.769
	HR_75%	80	98	84.82	5.416
	HR_100%	80	115	88.37	7.171
	HRV	0	1	.08	.270
	EDA_R	1	2	1.62	.327
	EDA_25%	1	4	2.12	.874
	EDA_50%	1	4	2.13	.900
	EDA_100%	1	5	2.21	.946
	Haptic Familarity	0	1	.43	.498
	VR Famlarity	0	1	.47	.502
	Q01	2	5	4.13	.875
	Q02	2	5	4.12	.904
	Q03	2	5	4.14	.821
	CQ%	25	100	86.03	19.131
	HWQ%	50	100	87.25	17.500
	Q17	2	5	4.38	.868
	NWQ18	2	5	4.25	1.029
	NW%	25	100	81.13	25.717
	UX%	25	100	75.49	25.607
	US%	25	100	84.56	25.134
	Enga_R	0	1	.46	.141
	Enga_25%	0	1	.70	.193
	Enga_50%	0	1	.70	.248
	Enga_75%	0	1	.70	.272
	Eng_100%	0	1	.70	.276
	EX_R	0	1	.47	.132
	EX_R_25%	0	1	.65	.251
	EX_R_50%	0	1	.66	.271
	EX_R_75%	0	1	.66	.270
	EX_R_100%	0	1	.73	.237
	Fo_R	0	1	.62	.146
	Fo_25%	1	1	.81	.114
	Fo_50%	0	1	.77	.234
	Fo_75%	0	1	.79	.235
	Fo_100%	0	1	.72	.264
	In_R	1	1	.68	.074
	In_25%	0	1	.72	.215
	In_50%	0	1	.76	.228
	In_75%	0	1	.73	.272
	In_100%	0	1	.77	.293
	Stress_R	0	0	.29	.098
	Stress_25%	0	1	.37	.178
	Stress_75%	0	1	.43	.177
	Stress_100%	0	1	.49	.227

4.4.1 Results of the Mukers Telehaptic VR Application

Regression Models results for *Mukers*

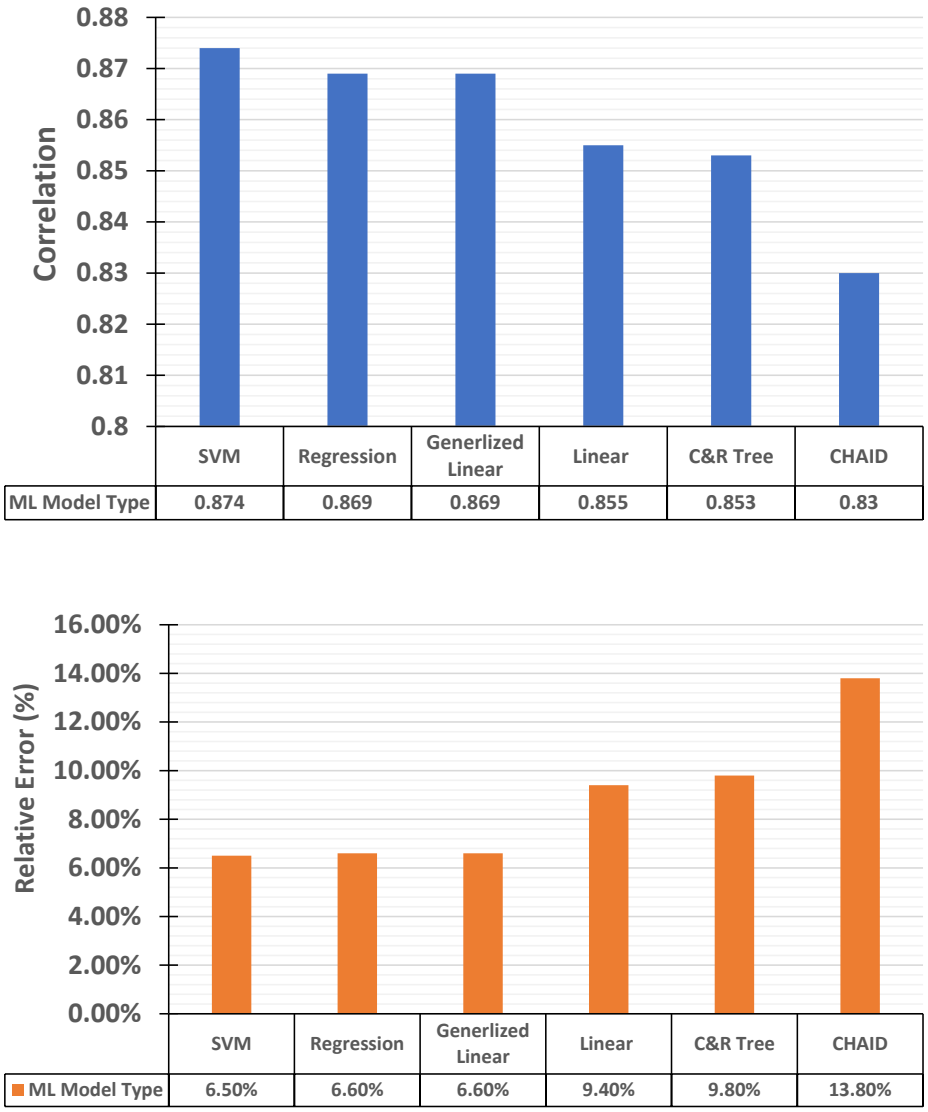


Figure 4.17: QoE based ML regression based Models result.

According to Figure 4.17, we observe that both SVM and normal and generalized regression are more efficient in prediction QoE for the Mukers when the inputs of the ML

algorithm are based solely on the high-level reported *Subjective* parameters of the taxonomy. Both models outperform the others with a Pearson correlation for the SVM=87.4% and 86.9% for the normal and generalized regression. In terms of how the reported QoE varies from its model-predicted level. The error rate for the SVM was the minimal with er=6.5%. This means that SVM can predict QoE for Mukers HVR game with results that only deviate 6.5 points on a scale of 100 from the reported subjective QoE values. The linear regression model also performs well and was capable to quantify the predicted QoE with equation 4.4:

$$QoE_P = \alpha_1.CQ + \alpha_2 HWQ + \alpha_3 NWQ + \alpha_4.UX + \alpha_5.US + \beta \tag{4.4}$$

Where α_i represents the weighting factor for each influencing quality and β is the intercept. With coefficient of determination= 91.0% , equation 4.4 can be realized with a 95% confidence interval as equation 4.5:

$$QoE_P = 5.751CQ + 0.1848HWQ + 0.3766NWQ + 0.1723.UX + 0.0372US - 11.29 \tag{4.5}$$

As can be noticed from equation4.5 has the most significant factors due to the fact that the users were was not aware of the simulated QoS disturbance for the controlled scenarios and most of them reported that there was a noticeable latency that impact the responsiveness of the game as well as physics and the force feedback from the haptic devices were abrupted during the HVR interaction. Keep in mind that the $Mukers$ is a dynamic Telehaptic VR environment that demands more communication with the server in order to update the objects' 3D model and its haptic and locomotion parameters.

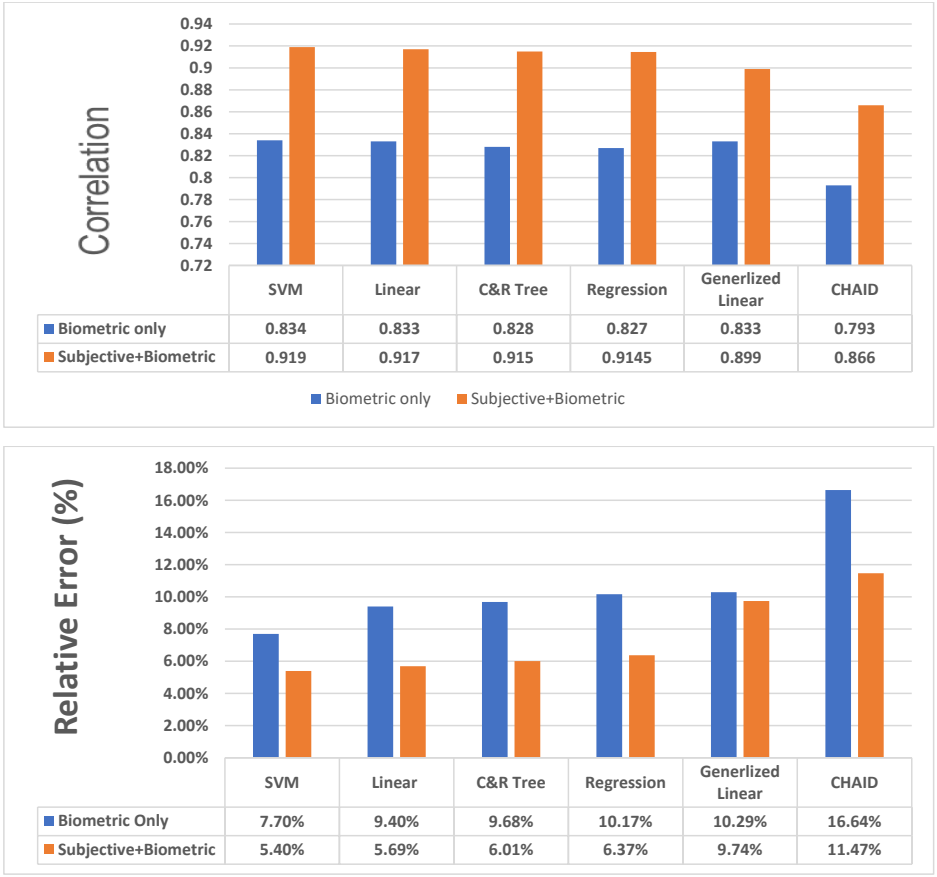


Figure 4.18: Evaluation Performance for ML QoE based on biometrics only (red) and Subjective+biometrics.

In terms of QoE prediction based on biosignals only. As can be seen in Figure 4.18, the SVM again outperforms the other ML algorithms. It should be noted that even with using biosignals as the input features for the ML algorithms, we have achieved an accepted prediction level with the largest relative error for Chi-squared Automatic Interaction Detection algorithm (CHAID) equal to 16.64%. It is obvious that incorporating both subjective and physiopsychological inputs improves the performance of all 6 ML algorithms reaching almost 92% correlation rate for the case of SVM. More statics about the performance of the 6 ML algorithms is depicted in Figure 4.18. Lastly, it should be reported that when we feed all the attributes to the ML regression pipeline,

the predictor's importance are depicted in Figure4.19.

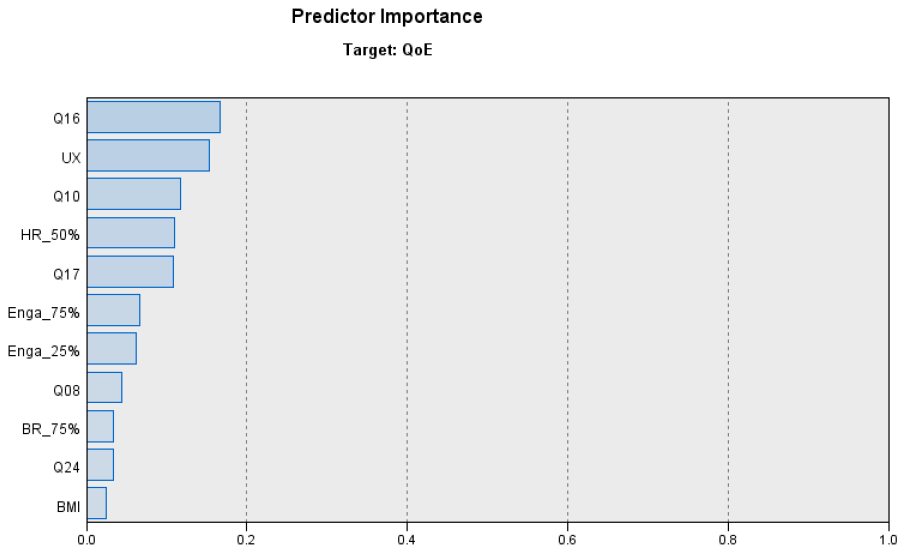
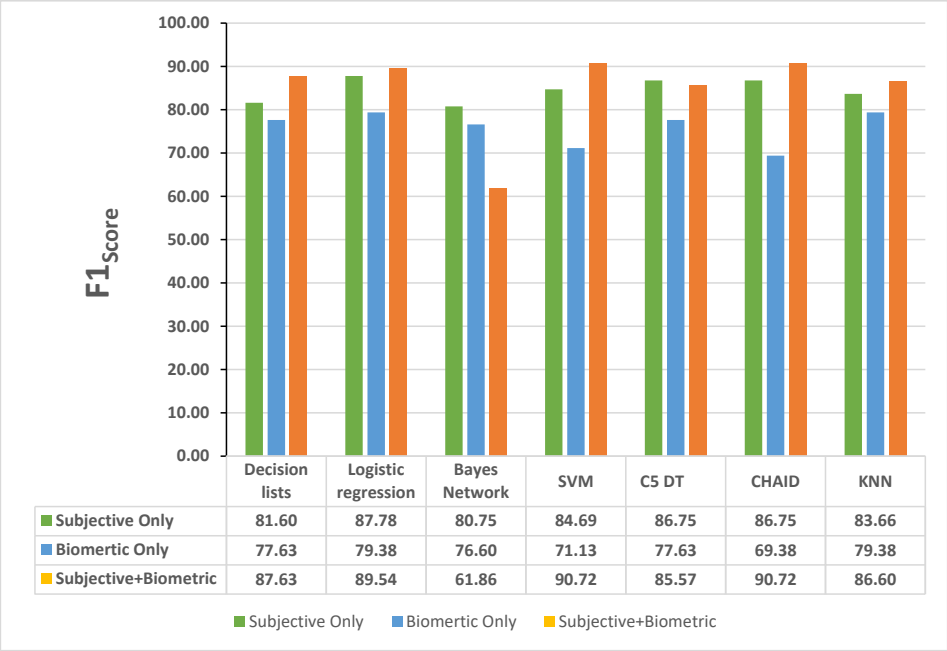


Figure 4.19: Predictor Importance of the linear regression QoE model.

As can be noticed, the user experience, the collaboration aspect of the game (Q16), the responsiveness of the HVR (Q17), as well as heart rate, breathing rate, and EEG engagement are the most important predictor in the linear regression model.

Classification Models Performance Evaluations for *Mukers*

In order to estimate QoE using classification-based algorithms, we have converted our dependant variable QoE into a binary format. More specifically, if the user reported a MOS equal to 3 i.e., accepted scaled of 50% QoE score, we set the classifier to 1 otherwise the value of QoE is set to equal 0.



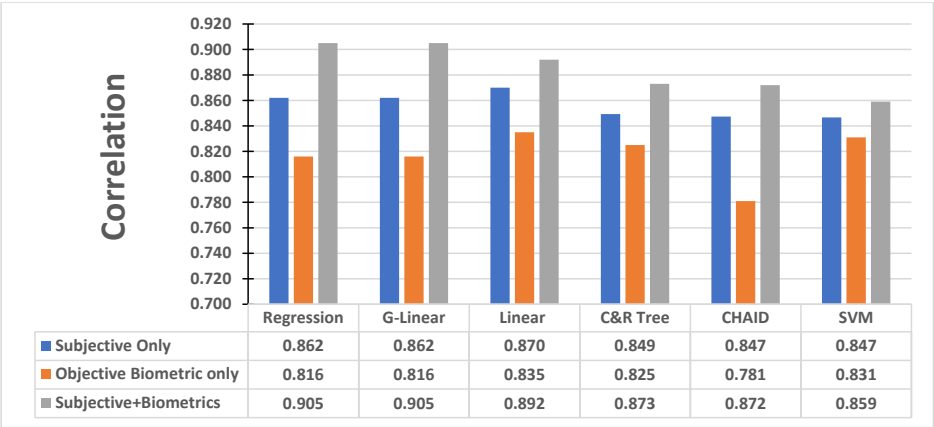
	Decision lists	Logistic regression	Bayes Network	SVM	C5 DT	CHAID	KNN
Subjective Only	81.60	87.78	80.75	84.69	86.75	86.75	83.66
Biomertic Only	77.63	79.38	76.60	71.13	77.63	69.38	79.38
Subjective+Biometric	87.63	89.54	61.86	90.72	85.57	90.72	86.60

Figure 4.20: The impact of different QoE features on $F1_{score}$.

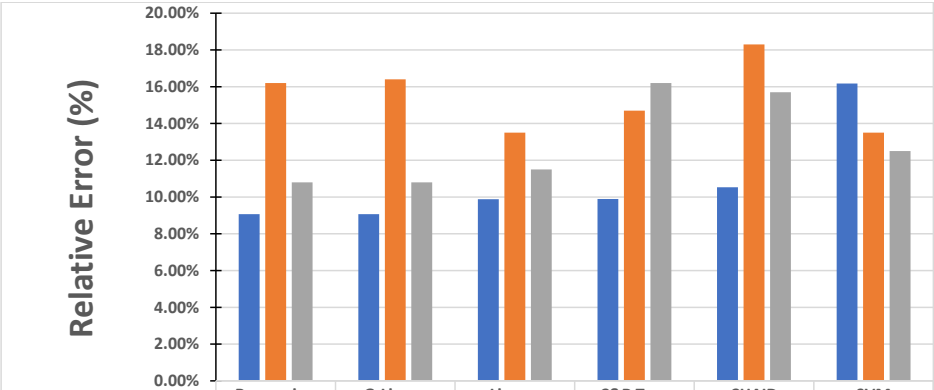
In essence, Figure 4.20 depicts the Evaluation performance of the QoE classification model based on 6 ML algorithms for our three cases of interest, i.e., subjective only, biometric only, and a combination of both. A point of interest is that most of the ML in the case of prediction QoE using the bio-metrics factors were capable to reach a threshold of $F1_{score}$ >= than 70%. In general, incorporating both biometrics and subjective ratings in the QoE classification model increases $F1_{score}$ of all ML algorithms, hence boosting the overall prediction performance. An exception to that is the Bayesian network is a boosting-based method that struggles in the case of large numbers of Feature's vectors. On the contrary, SVM can handle a large number of vectors configurations hence it can achieve an $F1_{score}$ approaching to 91%.

4.4.2 Results of the Remote Tactile Application

Regression Models results for RTA



	Regression	G-Linear	Linear	C&R Tree	CHAID	SVM
Subjective Only	0.862	0.862	0.870	0.849	0.847	0.847
Objective Biometric only	0.816	0.816	0.835	0.825	0.781	0.831
Subjective+Biometrics	0.905	0.905	0.892	0.873	0.872	0.859



	Regression	G-Linear	Linear	C&R Tree	CHAID	SVM
Subjective Only	9.07%	9.07%	9.88%	9.90%	10.53%	16.18%
Biometrics Only	16.20%	16.40%	13.50%	14.70%	18.30%	13.50%
Subjective+Biometrics	10.80%	10.80%	11.50%	16.20%	15.70%	12.50%

Figure 4.21: RTA Prediction Models: Performance evaluation In terms of correlation and er for the deployed ML algorithms.

In the case of prediction QoE for the RTA using the subjective parameters, both the generalized linear and regression models represent a good equilibrium in terms of correlation which is equal to 0.86 and relative error rate which is equal to 9.07%. The linear model outperforms all the other 5 ML models with a correlation equal to 0.87. It can be used

to quantify the predicted QoE in terms of the taxonomy high-level reported parameters as given in equation 4.6 with 95% confidence interval:

$$QoE_P = 0.2193CQ + 0.2809HWQ + 0.236NWQ + 0.1443UX + 0.1348US - 8.575 \quad (4.6)$$

Note that the hardware quality has the most significant weight. This can be justified by the fact that the HVR in case of RTA was static (3D models do not need frequent update from the server) and even in the controlled scenario 2 (100 ms latency, 60% drop in the throughput), subjects were still capable to feel the texture and the stiffness of the 3D objects. Not to forget that the tactile glove increases the subject's degree of perception. This confirms that the tactile modality is more immune to network impairments compared to other haptic such as force feed-backs and physics. With regards to predicting QoE using the user's physiopsychology features, as shown in Figure 4.21, all models did well with correlation results margined above 0.78 (the case of CHAID) and the best models were the linear and SVM with relative error rates margined around 13.5%. It is worth noting that the most significant biometric predictors are shown in Figure 4.22.

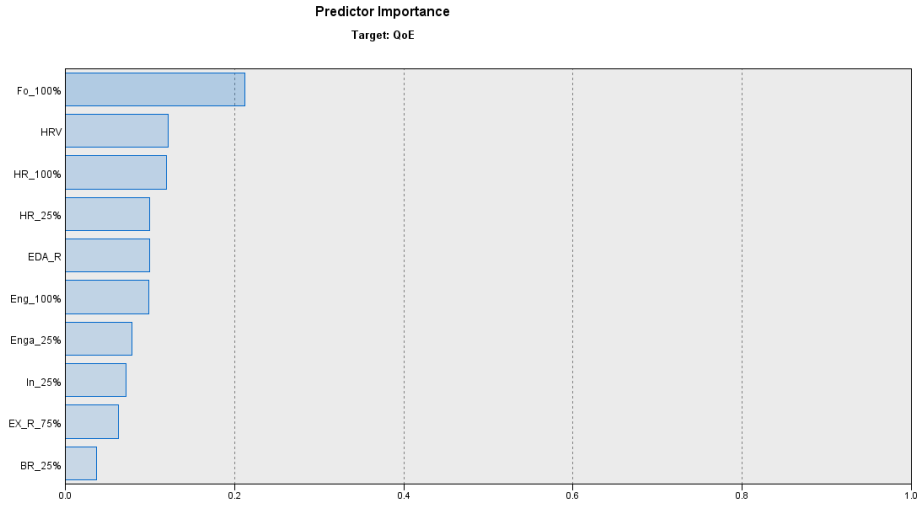


Figure 4.22: Predictor Importance in case of QoE regression ML for RTA

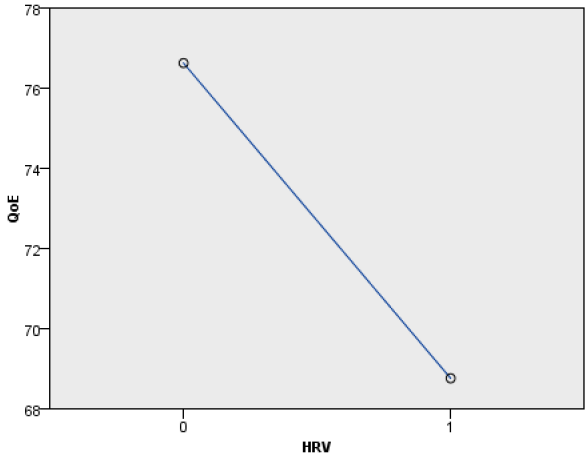
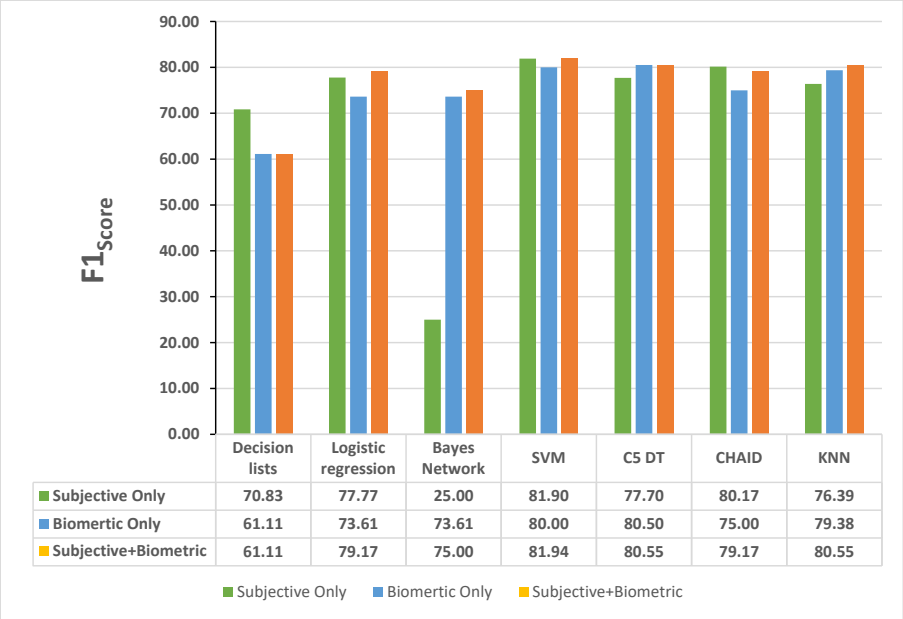


Figure 4.23: Impact on the HRV predictor on QoE.

Note that the predicted QoE linearly decreases, Figure 4.23, with the growth of HRV which indicates that subjects are under stress and overwhelming condition. The same observation is reinforced in the case of the telehaptic game. Lastly and as confirmed by Figure 4.21, combining both subjective and physiopsychology features boosted the prediction performance for most ML methods. In fact, the generalized linear and regression were the best estimation method with a correlation equal to 0.905 whereas the CR tree was the least efficient ML with er= 16.20%.

Classification Models Performance Evaluations For RTA



	Decision lists	Logistic regression	Bayes Network	SVM	C5 DT	CHAID	KNN
Subjective Only	70.83	77.77	25.00	81.90	77.70	80.17	76.39
Biomertic Only	61.11	73.61	73.61	80.00	80.50	75.00	79.38
Subjective+Biometric	61.11	79.17	75.00	81.94	80.55	79.17	80.55

Figure 4.24: Performance Evaluation of RTA model classifiers in terms of $F1_{score}$

As we discussed earlier, the tactile interaction was less susceptible to networks impairments. This leads all of our subjects in most scenarios to give an overall QoE over 60%. Only, A few subjects with haptic and VR backgrounds perceived the latency at control scenario three (the most disturbed QoS one) and reported that the impedance and as well as textures of the VR 3D objects become softer and marginally differ from the reference scenario. Unfortunately, this negatively impacted the data set. As a result, we have changed the classifier to report 1 if the user rated his/her tactile interaction above 75%. Even with this tweaking, the overall performance of the ML classifiers for RTA in terms of $F1_{score}$ is comparably lower than its correspondent at the tele-haptic game. This can be observed in the case of the subjective-based boosted Bayesian network classifier which gives a very low performance ($F1_{score}$=25) as the model failed to develop effective discriminative learning algorithms. A point of interest is that the Bayesian network classifier performed well when it is fed with the biometrics only features. Lastly, one can

observe that using the biometrics and the subjective results create more feature vectors which enhance the overall performance of most ML classifiers especially in the case of SVM and logistic regression.

4.5 Summary

This chapter included a discussion of the technology employed in the experiment as well as the research methodologies used to realized our DT QoE taxonomy. We wanted to elaborate and show the reasoning and justifications that drove the implementation of this experiment. We also comment about the variables that were important for this QoE modelling and why they were selected. We also build multiple QoE prediction models based on machine learning. These models can automatically estimate the QoE based on explicit users' subjective and implicit biomedical features. The models were benchmarked and assessed using multiple metrics. In the next two chapters, we will emphasize the QoE of DT teleoperation on a large scale as well as we will model an important physiological attribute i.e., fatigue when designing a DT system.

Chapter 5

Use Case 1: QoE for Large Scale Immersive DT Haptic Virtual Reality Communication over the Internet

5.1 Stabilized control Schemes for Haptic Communication

As discussed earlier, a standard teleoperation system over the digital twin architecture consists of slave (teleoperator) and master (operator) devices that interchange haptic data (positions, orientation, forces, speed, torques, impedances), video content, 3D texture, and audio content through a network channel such as the internet. The conveying of haptic information, in particular, places significant demands on the communication network because it closes a global control loop between the operator and the remote robot. As a consequence, the system's stability is extremely vulnerable to network latency. In particular and hypothetically, any non-zero value of latency or jitter leads to instabil-

ity of the DT system. Furthermore, high-fidelity teleoperation necessitates a high-level sampling rate for haptic signals of 1 kHz or greater in order to guarantee high-quality interaction and the stability of the system As a result, teleoperation systems necessitate the exchange of 1000 or even more data packets per second (PPS) among the master and slave devices. Such high packet rates are difficult to sustain in Internet-based networking. It should be noted that the user's quality-of-experience (QoE) and the teleoperation performance degrade with increasing communication impairments. Consequently, several techniques have been established by incorporating various stability-assuring control methods with haptic signal reduction systems to address the three key communication-related challenges of teleoperation systems (i.e. latency, jitter, and high data rate). While all control schemes can, in principle, guarantee system stability for random latencies, several control solutions have diverse latency tolerance, leading to various artifacts that reduce the user's degree of perception and necessitate various quantities of communication network resources. In the literature, the most well-known control schemes are the incorporation of the Perceptual Deadband-Based (PD) haptic data reduction approach with Time-Domain Passivity Approach (TDPA) [158, 159] or Model-Mediated Teleoperation (MMT)[133, 134]. Both of them are referred to as TDPA+PD and MMT+PD respectively [160], the operation of each control scheme is illustrated in Figure **??**.

The TDPA is proposed to tackle the stability of the teleoperation system. The arguments for stability are founded on the principle of passivity, which exemplifies energy transmission throughout a two-port network model and offers a necessary requirement for input-output stability. As depicted in Figure 5.1, the combination of passivity controllers (PC) as well as passivity observers (PO) will assist in ensuring the TDPA-based teleoperation system's stability. The former is used to contextually tune the adapted damper's parameters (α and β) in order to properly dissipate energy, hence ensuring the system's passivity. The latter is used to calculate existing system energy. In the TDPA+PD control approach, the PO is positioned right after the blocks of haptic data reduction and used to occasionally downsample the exchange of haptic packets using

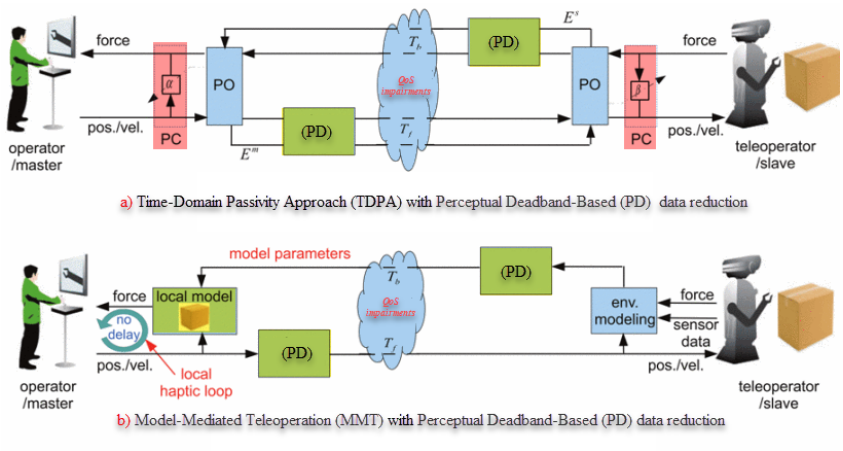


Figure 5.1: A demonstration of the state of the art stabilizing control approaches that use haptic data reduction.

perceptual limits. The TDPA is described as a conservative control scheme. As the delay increases, the demonstrated remote environment attributes turn out to be more distorted (e.g., the impedance of the objects irregularly changes, hence solid objects are experienced softer than they really are). Furthermore, when the PCs are triggered to dissipate the system output energy, the TDPA+PD method may cause abrupt force changes. With growing network latencies, this outcome becomes greater leading to an abrupt experience. As a result, the TDPA+PD solution is appropriate for short-distance teleoperation applications that may run at the communication network's edge in order to satisfy the necessity of repeated haptic updating among the operator and the teleoperator. Thus, it can deal with high ranges of interaction forms as well as frequent object dynamics such as deformation and motion.

One major problem with passivity-based control systems e.g., TDPA is the trade-off between the system passivity and transparency. This implies that the control scheme achieves stability at the expense of reduced transparency. To tackle this problem, an MMT control scheme is introduced to assure the transparency as well as stability of the teleoperation system even with the existence of network latencies. As can be seen in Figure **??**, on the operator side, a local object model is used to estimate the teleoperator

environment in the MMT control method. Once the teleoperator gets a new model, the model parameters representing that object in the teleoperator environment are always approximated concurrently and sent back to the operator. On the operator side, the local model is regularly rebuilt or modified frequently based on the obtained model parameters of the remote environment, and the haptic feedback is calculated without significant lag based on the local model. Stable and transparent teleoperation can be accomplished if the anticipated model is a precise approximation of the teleoperator environment. In the MMT-PD control scheme and using perceptual thresholds, the data reduction method is used to irregularly downgrade the sampling rate of the velocity signals in the forward channel and the model attributes in the backward channel. These thresholds for model attributes decide whether a model change results in a perceptible variation in the displayed signal. If this is not the case, the model update does not require to be sent back to the operator. The MMT scheme does not display hard objects softer with increasing delay, as the TDPA scheme does. As a result, when the network latency is considerably high, the MMT approach outperforms the TDPA in terms of teleoperation efficiency and quality. However, keeping the local model in harmony with the remote environment is difficult to infrequently updated scenes. As a result, the MMT solution is advantageous for medium-long distance teleoperation applications and situations with a low degree of scene dynamics. To test the aforementioned outcome, we have conducted a use case experiment for both teleoperation methods as discussed in the next sections.

5.1.1 A kinesthetic IEEE P1918.1.1. Reference Setup Description and QoE Experimental Protocol

We have used the example hardware and software setup for the evaluation of the haptic communication described in [161]. Further, we have also incorporated the implementation of TDPA+PD as well as MMT-PD schemes for control stabilizing communication with haptic perceptual data reduction. Ten subjects with and without haptic background

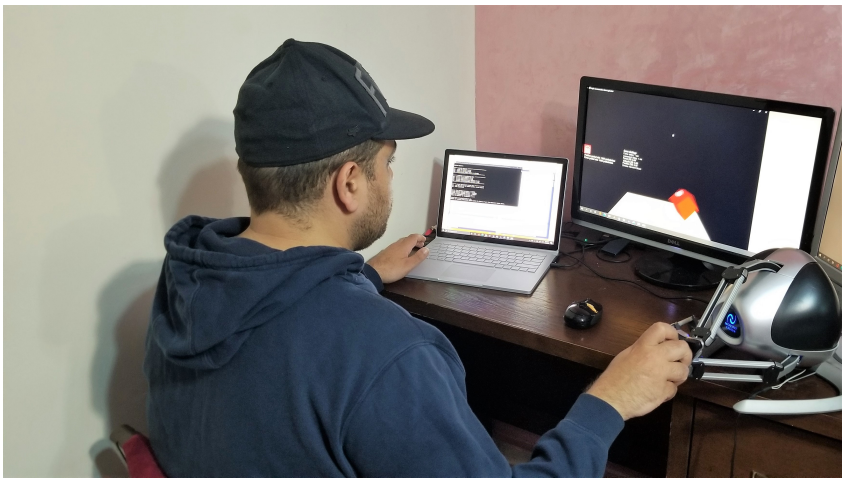


Figure 5.2: Experimental Setup for stabilized control teleoperation System. The master is the Omega.3 device while the slave is the VE.

are guided to use the teleportation system as shown in Figure 5.2. The subjects' ages range from 15-40 years old while all of them are right-handed.

The teleoperator i.e., the salve is represented as a virtual world consisting of a rigid (movable) cube sitting on a rigid planar surface The haptic device i.e the master is portrayed in virtual space by the small grey coloured ball (virtual tool). The omega.3 system is used to control the orientations and velocity of the virtual ball, and the subject perceives three degrees of freedom force feedback whenever the virtual ball makes contact with objects in the environment. Based on this configuration, we aim to achieve both static and dynamic teleoperation interactions: interactions with the rigid planar surface are classified as static, whereas interactions with the movable cube are classified as dynamic. Regardless of that, it should be noted that the omega.3 (www.forcedimension.com/products) is built around a one-of-a-kind kinematic architecture that has been designed for high-end force feedback. Its high mechanical stiffness, merged with its embedded fast USB 2.0 controller, allows it to make crisp contact forces. The omega.3 was developed and produced in Switzerland especially for challenging applications such as medical and space robotics, micro and nano manipulators, and tele-

operation research where performance and reliability are important. When the Omege.3 haptic device is powered on for the first time, it must be calibrated in order to display force feedback correctly. Else, there would be no force feedback. If the device is not calibrated, a red colour appears on the device's front light-emitting diode (LED). To calibrate the device, push/and pull it so that all joints meet their limits. When the LED turns blue, the device has been calibrated appropriately and read to be used.

The teleoperation environment was hosted in a workstation with an Intel i7 processor, 16 GB of RAM, Windows 10 Pro edition, and Nvidia Geforce 1060 TI GPU for smooth 3D frames and the best haptic rendering experience. with regards to communication network disturbances i.e, increasing latency and reducing the data rate, were simulated in the local work-station. To minimize learning effects and to increase the confidence

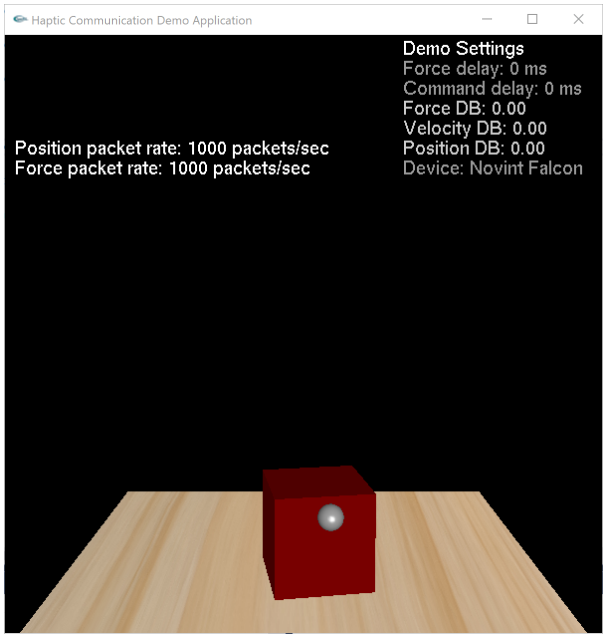


Figure 5.3: Teleoperation Reference Scenario.

of interval, each subject was first given a training session and we did not begin the experiment until the user felt comfortable with the device setup and the experimental procedure. After that, the subjects were asked to try the reference scenario at which no network impairments were deployed as well as the original environment haptic impedance

was shown, as well as the best output (no data-reduction, with no latency, as well as 1000 packet/s for both the positions and force-feed-backs) for this setup as shown in Figure 5.3.

Typically, throughout a closed-loop kinesthetic communication with a VR world, position-velocity signals are normally transmitted from the master (subject controlling the omega.3 device) to the salve i.e the virtual environments objects and 3D textures, via the forward channel, and force feedback signals are naturally transmitted from the slave to the master backward channel so it can be perceived by the subject. This is how the bilateral channel is established. Consequently, the haptic traces for the reference session are encoded using the kinesthetic IEEE P1918.1.1. reference codecs as illustrated in Figure 6.3 and Figure 6.4.

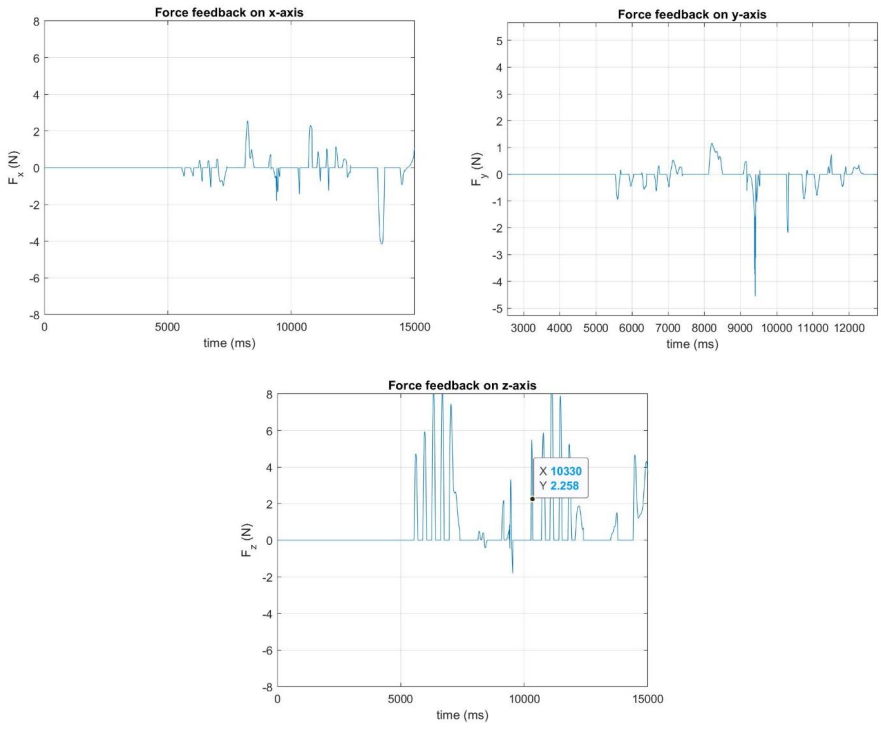


Figure 5.5: Measurements of the force feed-backs signals from the the slave to be perceived by Subject1

After the training session, each subject was asked to try control scenarios at which we

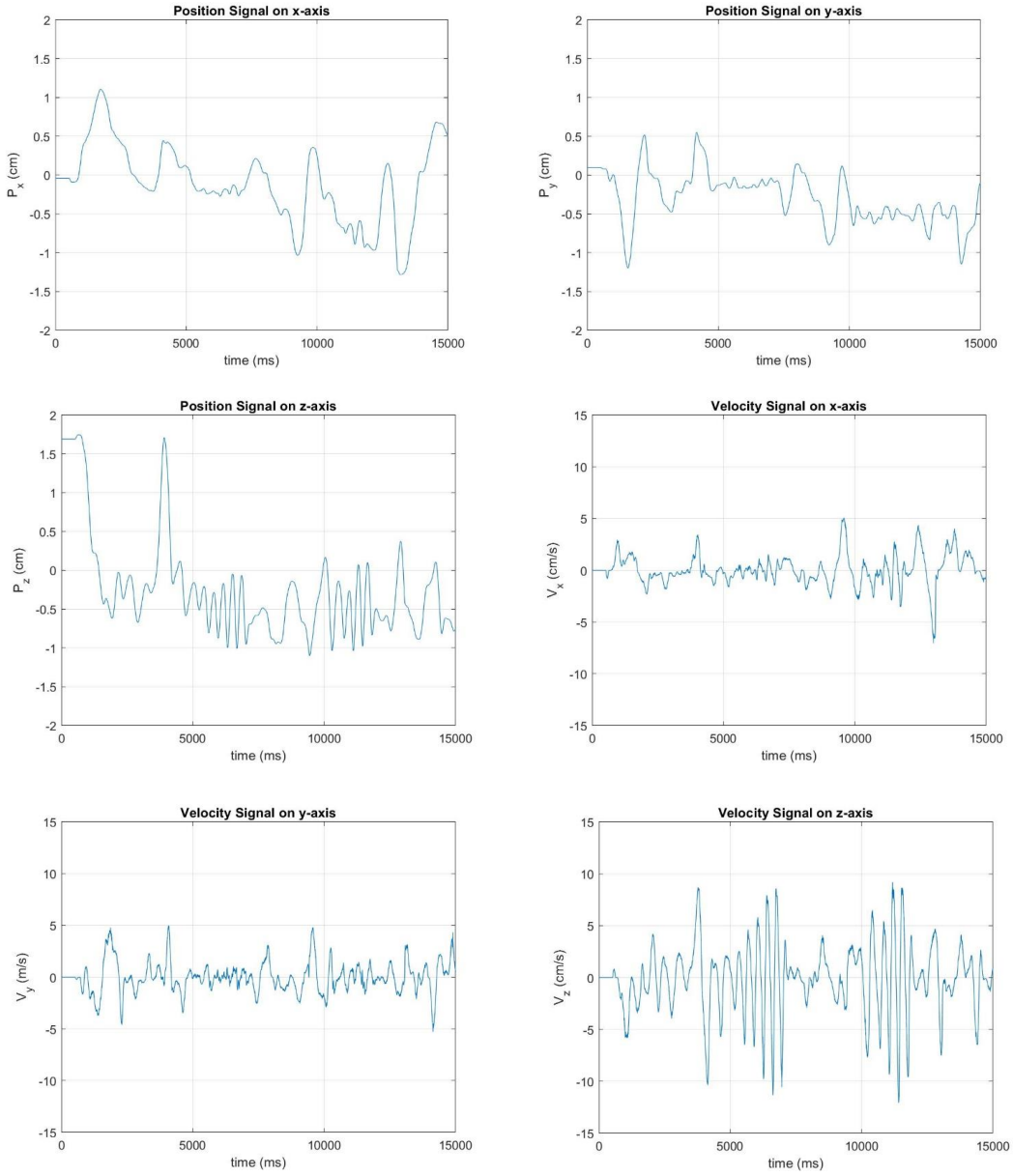


Figure 5.4: Measurements of the position and velocity haptic signals from the Master of Subject1.

injected artificial delays of 20 ms, i.e., 10 ms network delay on commanding channel plus 10 ms on the network delay on Force feedback channel, and overall 40 ms, and 80ms for the TDPA+PD teleoperation scenarios and then dynamically switch to MMT-PD control teleportation scenarios with a total of 100 and 200 ms, 300ms delays respectively. We have selected these thresholds very carefully as any latency $\leq 10ms$ is considered unperceivable for the TDPA+PD teleoperation whereas the MMT-PD based teleportation can hold a latency $\geq$ 100 ms with a haptic data rate reduction of 10% [162]. The data rate was reduced by increasing the value of positional and force DB from 0.01 to 0.1 accordingly. All control scenarios are demonstrated in Figure 5.6. ‘

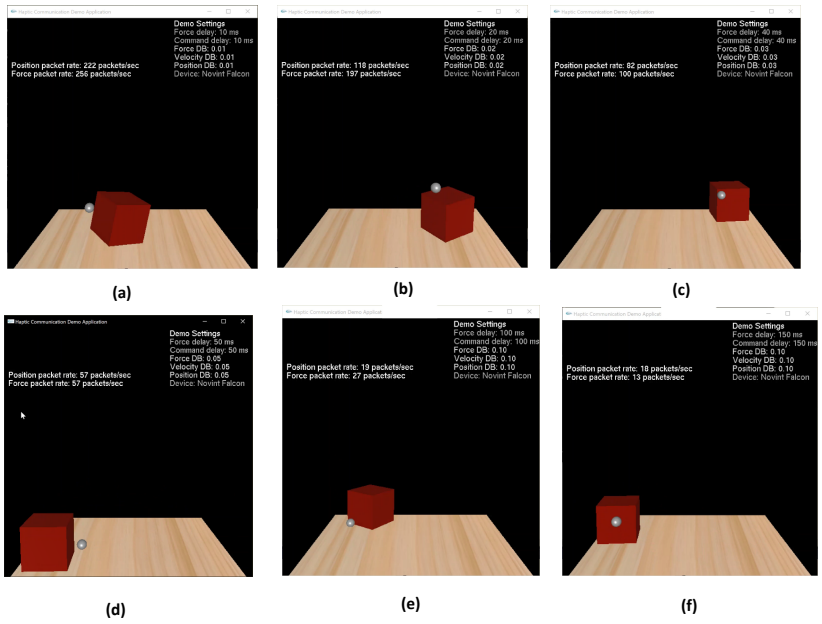
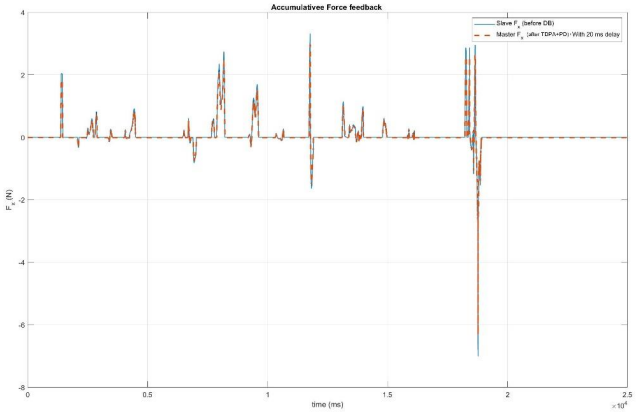


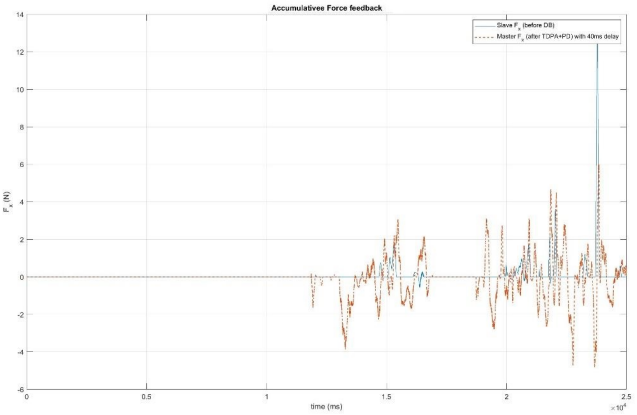
Figure 5.6: Control scenarios parameters. *a*) TDPA+PD with 20 ms latency and DB of 0.01. *b*) TDPA+PD with 40 ms latency and DB of 0.02. *c*) TDPA+PD with 80 ms latency and DB of 0.03. *d*) MMT-PD with 100 ms latency and DB of 0.1. *e*) MMT-PD with 200 ms latency and DB of 0.1. *f*) MMT-PD with 300 ms latency and DB of 0.1.

It is worth noting that for the sake of benchmarking each of the control schemes individually, we let the user try both the TDPA+PD and TDPA+PD under all latency and date rate conditions. The user interacted with the virtual environment by moving

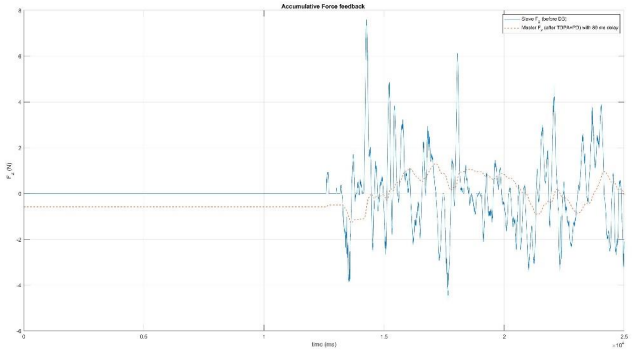
the ball (virtual tool) towards the remote cube (virtual model) and gradually changing the force applied to move it accordingly. Users were asked to rate the interaction quality between each control of the six scenarios and the reference scenario using the DCR scale. The rating has 5 grading where level 5 denotes "imperceptible difference", 4 denotes "perceptible", but not disturbing", 3 denotes "slightly disturbing", 2" denoting disturbing", and level 1 represents "strongly disturbing and unacceptable experience". Since the reference scenario was deemed to give a score of 5 i.e, best experience, the user can recall it any time through the experiment session. Users were asked to try each control scenario three times as well as we have randomly selected the sequence of the control scenarios to eliminate any bias in the results. It should be noted that when reporting the interaction quality, we have requested from the users to consider all possible perceivable artifacts such as noticeable object's impedance changes, force jumps and, sudden vibration.



a) Master and slave force signals for TPDA-PD with 20ms network latency and data rate of 300 pps.



b) Master and slave force signals for TPDA-PD with 40ms network latency and data rate of 200 pps.

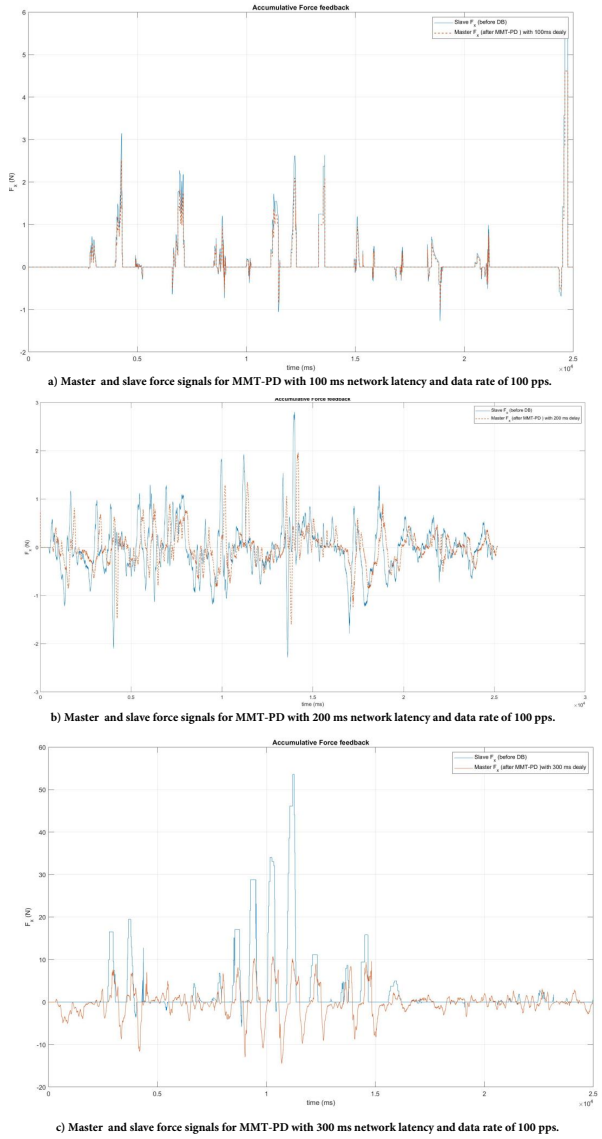


c) Master and slave force signals for TPDA-PD with 80ms network latency and data rate of 100 pps.

Figure 5.7: Measurements of the force signal of the master and salve during TDPA+PD control scenarios

5.1.2 Objective Results

Figure 5.8 illustrates the force signals of the salve and the master during the TDPA+PD control scenarios. The blue line represents the force signal generated by the slave whereas the dashed red line represents the force feedback observed by the master after the data reduction and the given network conditions. As can be seen the TDPA+PD is stable as long as the overall latency is less than 40 ms with data rate maintained around 200 packets-per-seconds (PPS). Otherwise, the master force has much deviation from the slave force as depicted in Figure 5.8.c. This leads to unstable interactions as the users experienced sudden force jumps, demonstrated as unexpected peaks, and overshooting stiffness of the remote object model. On the contrary, during the MMT+PD teleoperation control scenarios, the force of the master was following the salve force with a slight deviation when the latency $\geq$ 300 ms with a data rate of 100 PPS. This reinforces the fact that MMT+PD maintains a stable and transparent QoE even in tight network conditions such as an increased communication latency. This is because the MMT+PD creates a local model of the remote environment and the parameters of the model are only rebuilt when a major model mismatch has occurred, i.e., movement of the cube. It should be noted that the users reported flickering artifacts when the latency of the bilateral channel increased above the threshold of 200 ms compared to the best performance of the TDPA+PD scheme which could be explained by the fact that MMT+PD performance depends on the local model accuracy which will require more time to be effectively updated and rendered to the master in the case of high dynamics model variations with the existence of network latency.



a) Master and slave force signals for MMT-PD with 100 ms network latency and data rate of 100 pps.

b) Master and slave force signals for MMT-PD with 200 ms network latency and data rate of 100 pps.

c) Master and slave force signals for MMT-PD with 300 ms network latency and data rate of 100 pps.

Figure 5.8: Measurements of the force signal of the master and salve during MMT+PD control scenarios.

5.1.3 Subjective Results

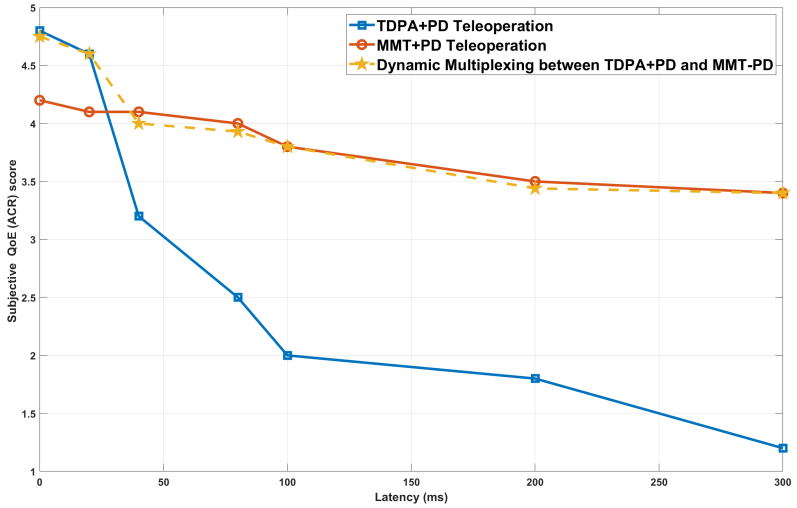


Figure 5.9: QOE reported score for the control teleportation methods and end-end latency.

Figure 5.9 depicts a quantitative assessment of the reported scores (QoE) for the two control methods. When the network latency is low, the TDPA+PD approach scores a higher QoE than the MMT+PD method. The QoE of the TDPA+PD scheme, on the other hand, degrades rapidly with increasing delay. This is because as the latency increases, the participants can feel more vibrations and force jumps, as well as softer environmental impedance. It should be noted that when the network latency is low, the TDPA+PD approach scores a higher QoE than the MMT+PD method. The QoE of the TDPA+PD scheme, on the other hand, degrades rapidly with increasing delay. This is because as the latency increases, the participants can feel more vibrations and force jumps, as well as softer environmental impedance. In contrast, the QoE of the MMT+PD approach is reasonably steady, reinforcing that the QoE of the MMT+PD scheme is primarily determined by model approximation precision rather than latency impairment.

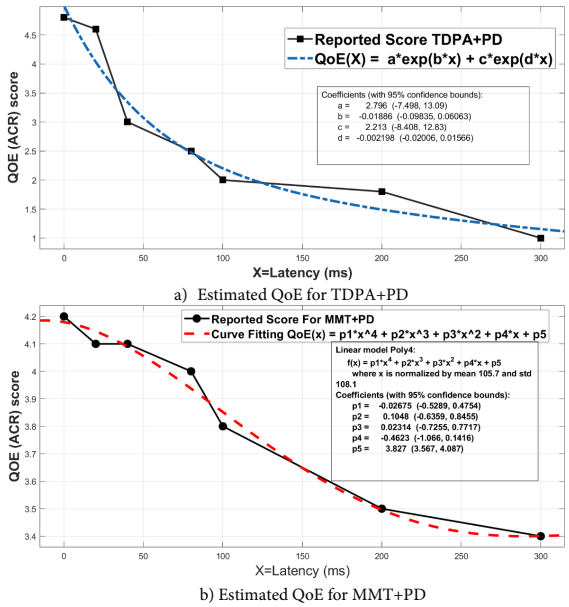


Figure 5.10: Curve fitting to quantify $QoE(delay)$ for both TDPA+PD and MMT+PD

Using the Matlab curve fitting toolbox, we can quantify the QoE for both control schemes. The non-linear regression two-term exponential function was the best fit for TDPA+PD teleportation with the goodness of the curve determined by the coefficient of determination R^2= 0.8955. Hence equation 5.1 can be used to objectively quantify the QoE with respect to latency for the TDPA+PD user experience.

$$QoE = \sum_{i=1}^{2} a_i.e^{-b_i(QoS)} \tag{5.1}$$

where the values of the coefficient (a_i and b_i are calculated based on a 95% confidence interval as depicted in Figure 5.10.a. On the other hand, Figure 5.10.a confirms that the linear model of the quartic polynomial i.e., the fourth-degree polynomial is the best exemplary with (R-square=0.9531) to objectively quantify the QoE for MMT-PD tele-operation with respect of QoS latency increasing as illustrated in equation 5.2.

$$QoE(QoS) = p_1QoS^4 + p_2QoS^3 + p_3QoS^2 + p_3QoS + p_i \tag{5.2}$$

The values of the coefficient p_i of equation 5.2 are already illustrated in Figure 5.10.b calculated with a 95% confidence upper and lower bounds.

In summary, even though the two stabilizing control methods tackle different teleoperation scenarios, the smaller the overall network impairments such as end-end latency, the higher the system transparency and thus the QoE. Combining equations (5.1,5.2) will establish a dynamic multiplexing among control and QoS algorithms according to the giving network conditions, hence the optimal solution to the best QoE teleoperation can be achieved (as illustrated in the orange dashed line in Figure 5.9). The QoS implementation of haptic communication on large scale is discussed in the next coming sections.

5.2 QoS for Immersed CHAVE

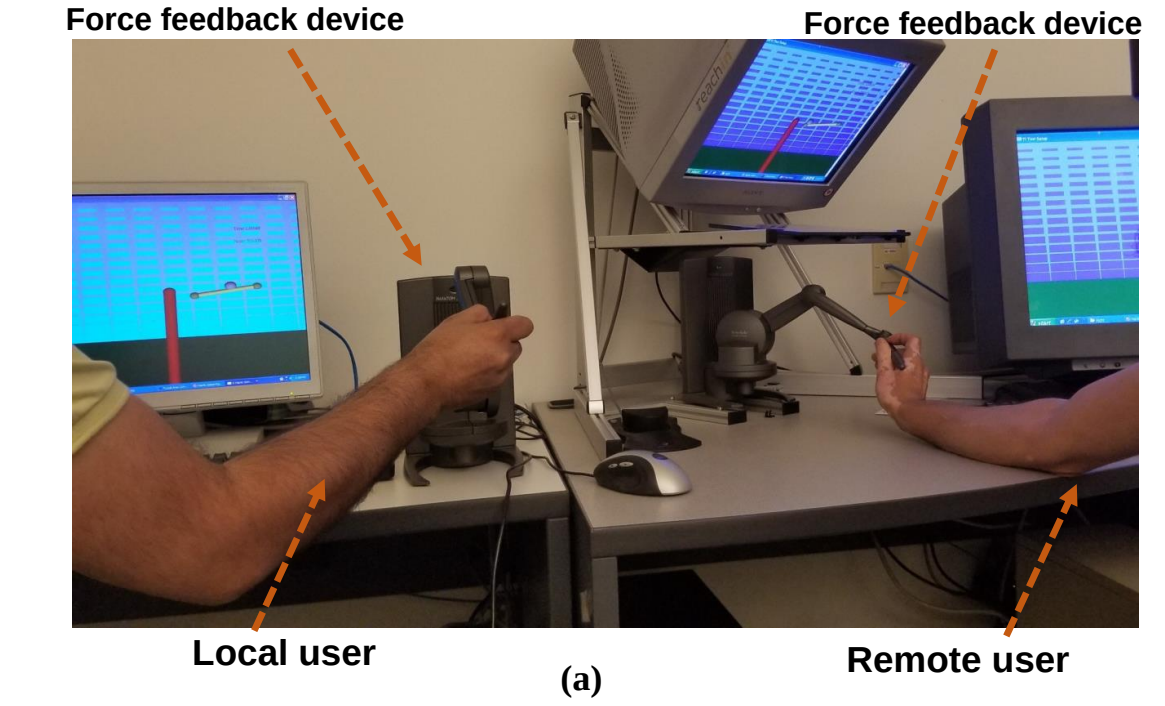
Again, The most common QoS architectures and protocols that are recommended by the IETF include Application Layer QoS, Relative Priority Marking, Service Marking, Integrated Services with Resource Reservation Protocol (RSVP), Differentiated Services (DiffServ), Multiprotocol Label Switching (MPLS), and Software-defined networking (SDN). So far, these approaches have been deployed to support the transmission of traditional modalities e.g, video, audio, and graphics. The provision of a specific QoS implementation for multimodal haptic traffic as well as human-computer interaction has got minimal research attention. Therefore and as indicated in [163], one of the most alarming challenges in the industrialization of 5G is to provide the QoE/QoS implementation needed to support and enhance the performance of haptic communication over the Internet.

According to [42] which evaluated a number of solutions to transmit haptic teleoperation-telepresence data, ALPHAN was found one of the best haptic communication protocols. Unfortunately, ALPHAN was evaluated neither at the Internet where network resources become shared, limited and changing nor on large scale. Therefore, The main contribu-

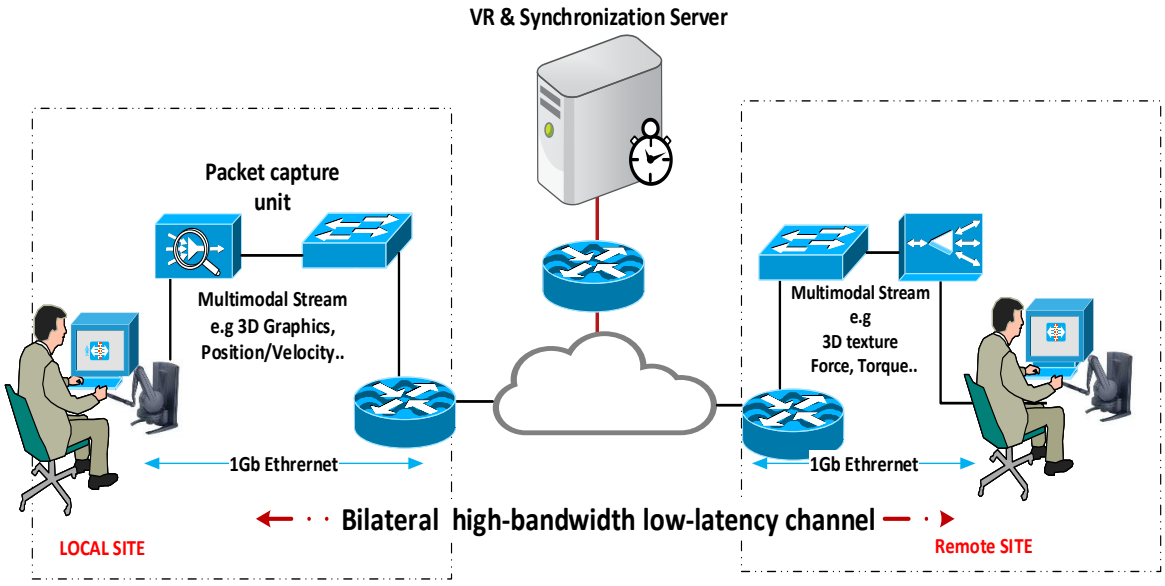
tion of this work is to provide a framework capable of minimizing the overall latency, jitter, and packet loss of large-scale haptic transmission sent over the Internet. Consequently, we conducted a set of experiments and analytical studies to model the Alphan haptic traffic so we can benchmark it on a large scale measuring the impact of several IP-QoS mechanisms on the haptic system performance.

5.3 Haptic Networked Traffic Modeling

To model the haptic network transmission we used an application that is immersed with 3D audiovisual as well as bilateral tactile and kinesthetic codecs. This application is is the Balance Ball game[164], shown in Figure5.11.b. As per haptic rendering, the application was implemented in C++ using the OpenHaptics API. The experiment was carried out on the MCR-Lab at the University of Ottawa. The collaborative application ran on two workstations. The workstations were running Windows platforms on a 2x Intel Xeon 3.8GHz with 4GB of RAM and an Nvidia GTX 1050TI video card. A Touch X haptic device [25] was attached to each workstation. The Touch X haptic device is six degrees of freedom (DOF) positioning and sensing haptic device developed and marketed by Geometry Touch, Inc. It has a compact design and provides three DOF force output capabilities. The technical specification for the Touch X haptic interface is provided in Figure 5.12. A ball is placed on a long wooden board that is held by two players from each side. A player holding one end of the board would feel many forces that mimic the real-life force experience. The game involves the two users collaborating in maintaining the balance of a virtual ball on a board using local and remote haptic devices. Each player holds one end of the board with his/her haptic device and raises it slowly over a virtual pole to a predefined threshold. The challenge is to collaborate in a trail to keep the board horizontally balanced as much as possible from the initial location to the destination. Any displacement in the horizontal balance will cause the ball to move away thus punishing both players. The players should remedy that by using the force



(a)



(b)

Figure 5.11: Experimental Setup. (a) represents A screenshot of the Haptic balance ball game. Two users are interacting haptically with the board, (b) represents schematic model of a multimodal collaborated haptic system over an ultra-low latency Communication Channel.

SPECIFICATIONS	TOUCH X™
Workspace	~6.4 W x 4.8 H x 4.8 D in > 160 W x 120 H x 120 D mm
Range of motion	Hand movement pivoting at wrist
Nominal position resolution	> 1100 dpi ~0.023 mm
Maximum exertable force and torque at nominal position (orthogonal arms)	1.8 lbf/7.9 N
Stiffness	x-axis > 10.8 lb/in (1.86 N/mm) y-axis > 13.6 lb/in (2.35 N/mm) z-axis > 8.6 lb/in (1.48 N/mm)
Force feedback (3 Degrees of Freedom)	x, y, z
Position sensing/input (6 Degrees of Freedom) [Stylus gimbal]	x, y, z (digital encoders) [Roll, pitch, yaw (magnetic absolute position sensor, 14-bit precision)]
Interface	USB 2.0

Figure 5.12: Touch X Technical Specifications

feedback and the 3D graphics to apply their judgment in balancing the board again.

Further, to confirm the authenticity and integrity of our experiments, we extended the testbed used in the ALPHAN networked haptic protocol as shown in Figure 5.11.b. Instead of running the haptic transmission process on a dedicated local area network where minimal background traffic exists, we decided to move the communication process to the Internet where network resources become shared and changing. To emulate and mimic the vision of the Tactile Internet, two 1Gbps ethernet cards were attached to each computer. Two Cisco 3750 switches used to implement the LAN as well as three Cisco 4431 Integrated Services Routers were used to provide very low latency, high-reliability communication channels, high-bandwidth WAN connectivity. In order to eliminate undesirable jitter, the clock of the local and remote site's computers was synchronized with a Network Time Protocol (NTP).

At each side, a Riverbed Application Characterization Environment (ACE) module [165] later named SteelCentral Transaction Analyzer was used to capture haptic traffic sent between local and remote sites. The collected traces of the haptic traffic were used to generate a real-life discrete event flow to simulate the QoS behaviour of haptic

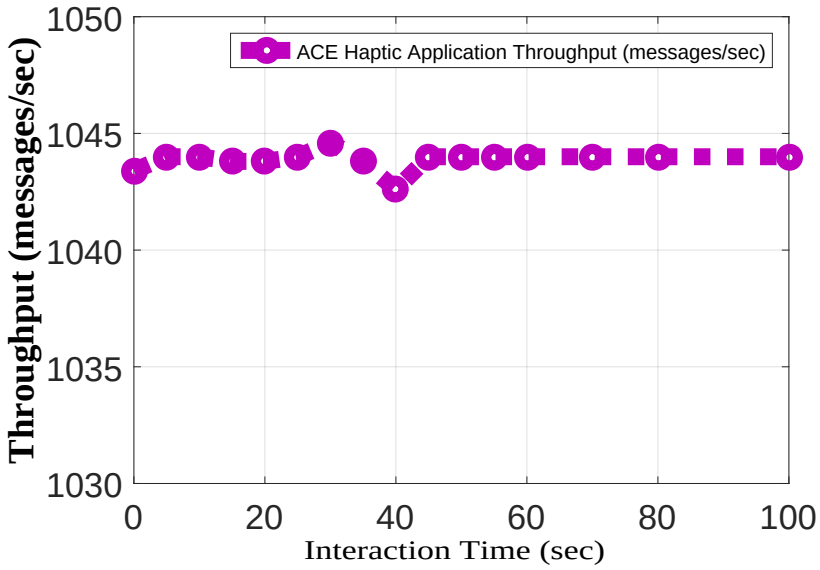


Figure 5.13: Throughput of the Developed ALPHAN-based Haptic Model over a TI Communication Channel

in a large-scale network running ALPHAN as a haptic communication protocol. As shown in Figure 2, the captured haptic model was capable of transmitting around 1044 packets per second. This rate matches ALPHAN experimental rate as ALPHAN sends the static position information at the graphics loop rate i.e., 30 updates per second whereas the dynamic position pieces of information are sent at the haptic loop rate which is 1000 updates per second. Ideally and based on Cisco routers statistics, sending haptic information on a network topology that mimics the vision of the TI infrastructure, haptic communication will enjoy a very low end-to-end latency (less than 0.208 ms), and ultra-low delay variation of 1 μs and a zero-data-loss. This will indeed enhance the overall quality of perception to the haptic users. It should be noted that when the communication channel between the two players is impaired using the network emulator user have reported an abrupted user experience. Briefly, table 5.1 summarizes the impact of impaired QoS KPI on QoE in a collaborated haptic VR system. For instance, when the latency increases, users perceive more force jumps and vibrations of the haptic interface. They might also perceive wrong feedback as the objects in the remote environment get

Table 5.1: Impact of QoS disturbance on over all QoE of DT Modalities.

Undesirable QOS metrics	Impact on QOE
Throughput	Abrupt movement and force jumps
Packet Loss	Inaccurate and abrupt movement of the haptic device
Latency	Reduce user's perception and eliminate collaboration aspect
Jitter	Instability of the system

softer impedance (stiffness).

5.4 Riverbed Simulation Scenarios

Riverbed® Modeler [165], was formerly referred to as OPNET Modeler Suite, is a discrete event (DES) simulator that has been widely used by academic and industrial entities to analyze networks and compare the impact of different technology designs on end-to-end behaviour. Unlike other network simulators, the riverbed modeller was chosen because of its ability to utilize the real-life captured haptic traffic to create a better representation application model that gives more precise emulation results. In addition, Riverbed is capable of simulating in both explicit DES and hybrid simulation modes and supports other simulation features like co-simulation, parallel simulation, high-level architecture, and system-in-the-loop interactive simulations. In a nutshell, the riverbed modeller constructs the models using an object-oriented modelling approach. Network devices such as routers, switches, firewalls, etc, are called node models. Node model comprises of modules connected by packet streams or static wires. Each module is allocated to a process module to attain the required behaviours. Riverbed's process model uses a finite state machine (FSM) approach to support the implementation of protocols, resources, applications, algorithms, and queuing policies. States and transitions graphically define the progress of a process as a result of certain events. Each state of a process model contains embedded C/C++ code, supported by an extensive library of functions

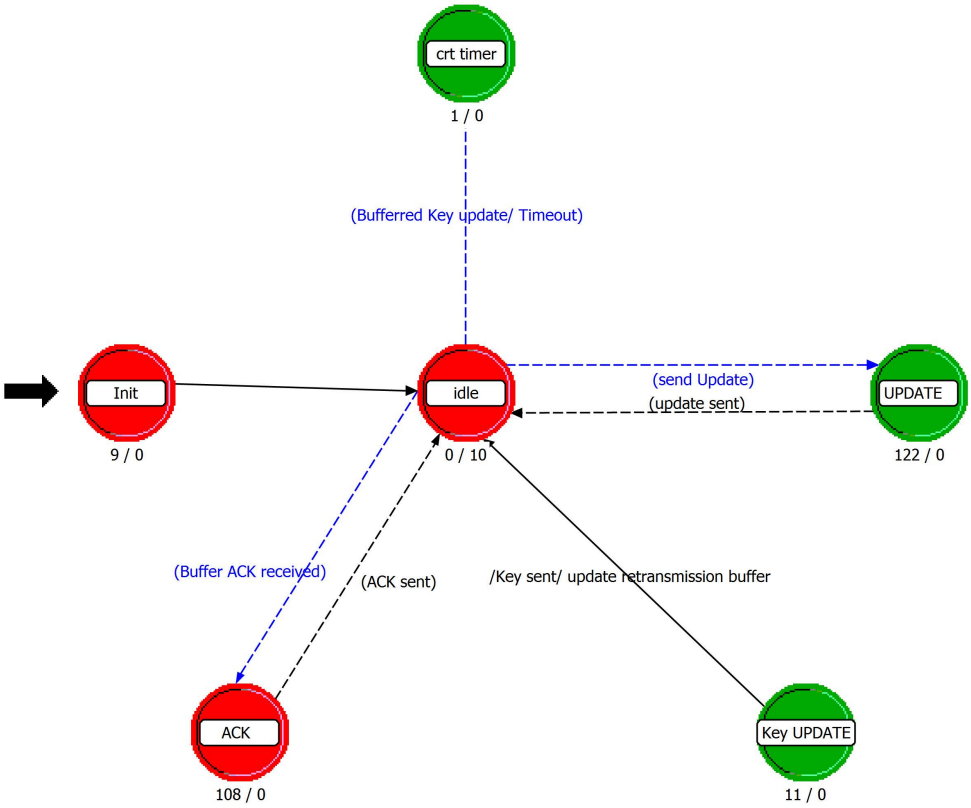


Figure 5.14: Riverbed Process Model of ALPHAN Sending Module.

designed for network programming. Since ALPHAN was implemented in C++, riverbed was feasible in constructing ALPHAN process model as shown in Figure 5.14.

In order to evaluate the performance of the ALPHAN haptic transmission model, we have created a network model using the riverbed's network editor. As shown in Figure 5.15, the human-operator and teleoperator workstations are configured using "Task Configuration Utility" to send and receive data through the process model with a throughput of 1044 messages per second. Discrete event flows that match the characteristics of the real captured haptic traces are defined by the riverbed ACE module and thus deployed as customized applications and profiles. Similarly, 28 other workstations are placed in the local and remote haptic LANs providing a large-scale haptic network. The IP cloud was also implemented with the Tactile Internet results achieved in the previous section i.e, 0

packet discard ratio, and packet latency of 0.1ms. The other three workstations/LANs are configured to use the built-in applications and profiles to send voice-PCM G.711, video H.264, and file transfer FTP Accordingly. Since our main goal is to explicitly optimize haptic communication, heavy FTP was modelled as an implicit, i.e., background traffic.

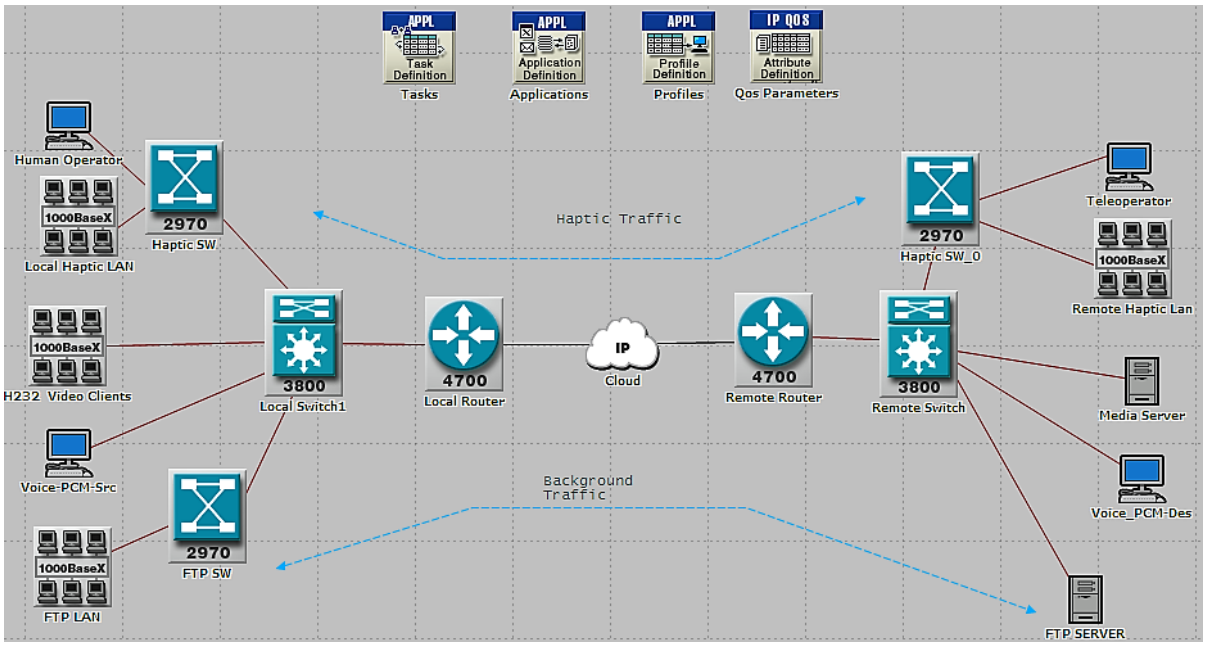


Figure 5.15: Network topology for QoS performance verification of ALPHAN-based Haptic communication Model.

Figure5.15 illustrates the effect of transmitting haptic, audiovisual, and background traffic on the local router's resources when different data rates are selected. By neglecting the intermittent period, we can clearly notice that the router exhausts its resources and reaches to the maximum utilization when the transmission rate is set to be 15Mbps. Consequently, This is the ideal data rate to create a congested scenario that matches the characteristic of the shared Internet. Therefore, The bottleneck link is positioned

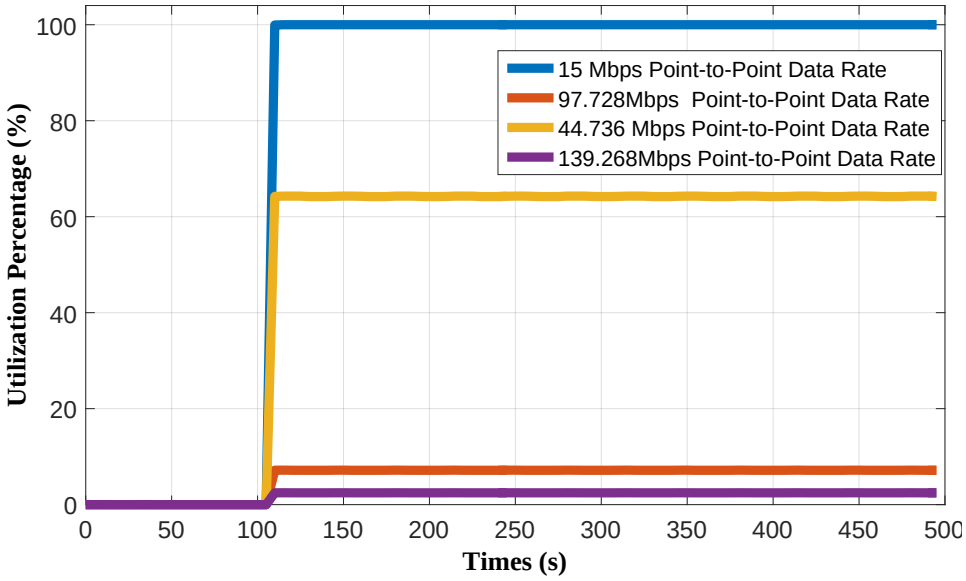


Figure 5.16: Utilization of the local router's resources when different transmission rate is selected.

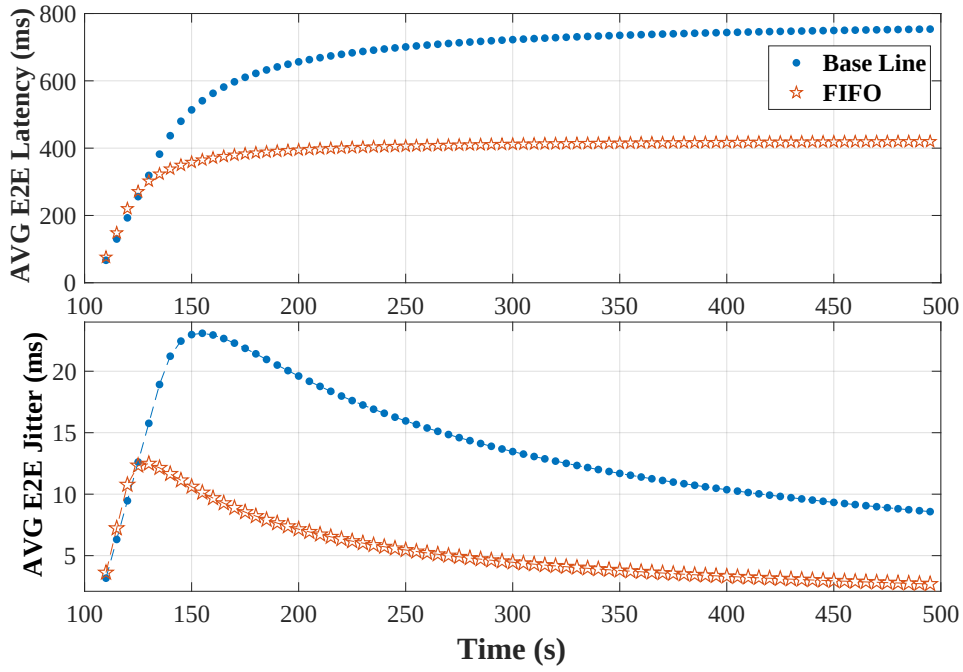


Figure 5.17: End-to-End latency and Jitter of the ALPHAN-based Haptic traffic for Best effort and FIFO.

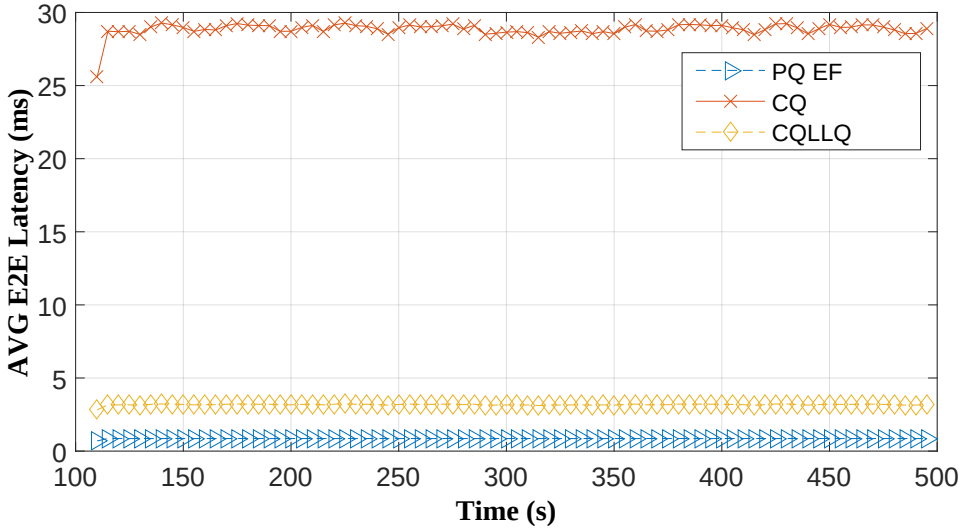


Figure 5.18: Comparison of the impact of CQ, CQ-LLQ and PQ on the End-to-End latency of the ALPHAN-based Haptic traffic.

immediately before the output interface of the local router where different types of queuing algorithms are implemented to optimize the QoS performance of the haptic traffic in terms of latency and jitter. In this context, we are considering the average end-to-end latency and jitter both measured in milliseconds. Unlike [47] which only focused on the provision of WFQ and CBWFQ on the haptic stream, we extend our experiments to cover the impact of FIFO, PQ, CQ, and CQ with LLQ on the ALPHAN haptic transmission model.

As highlighted in Figure 5.17, sending haptic data over ALPHAN in a congested network using the best effort approach i.e, no IP-QoS algorithm i.e. is deployed, will result in high latency and jitter values. When FIFO with large queue size is applied on the bottlenecked router interface, the performance slightly improved as the latency of the haptic stream decreased from 800 ms to 450 ms and this reflects a minimal enhancement of the stream's jitter values. Even though these results still do not meet the minimum threshold to maintain a stable haptic rendering experience.

Figure 5.18 illustrates the overall latency of the haptic stream when PQ, CQ, and

CQ with LLQ are applied on the bottlenecked router interface. For CQ, results were obtained after assigning the haptic queue with the highest byte count among other types of services (TOS) such as voice and video. Given that ALPHAN haptic packet has a size of 74 bytes (28 byte payload and header + 8 byte UDP header+ 20 byte IP header + 18 byte Ethernet header), we used equations 5.3 and to guarantee 60% of the 15 Mbps congested bandwidth to ALPHAN haptic stream. The remaining 40% were shared equally between voice and video TOS queues. Remember that CQ is based on round-robin transmission, therefore if we are looking for more guarantees, the haptic TOS queue can be assigned with the highest priority possible using LLQ. As can be seen from Figure 7 both CQ-LLQ and PQ with expedited forwarding PHB reduced the overall latency to 2.54 ms and 1.36 ms respectively. It is worth noting that the jitter results for the aforementioned algorithms were very minimal in μ seconds and can be neglected.

$$CQ_{BYTE_{Count}}(Q_i) = \frac{A * T_r * C}{8 * p} \tag{5.3}$$

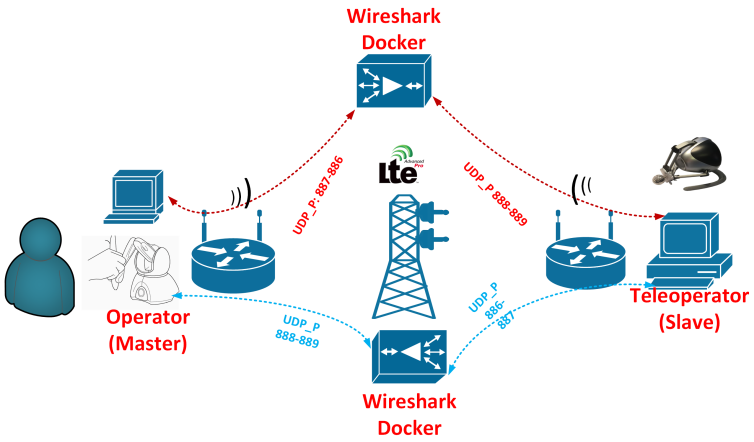
$$\text{BW CQ}_{BYTE_{Count}} = \frac{CQ_{BYTE_{Count}}(Q_i)}{\sum_0^N CQ_{BYTE_{Count}}(N)} * BW\,interface \tag{5.4}$$

Where: A denotes The total packet size of of application i in byte, T_r: Transmission rate bps, C: the number of end-to-end connections, N:the total number of CQ queues, P: The proportional weight of the packet size of i compared to packets' sizes from all other N queues, $BW\,interface$: the bandwidth capacity of the congested router's interface.

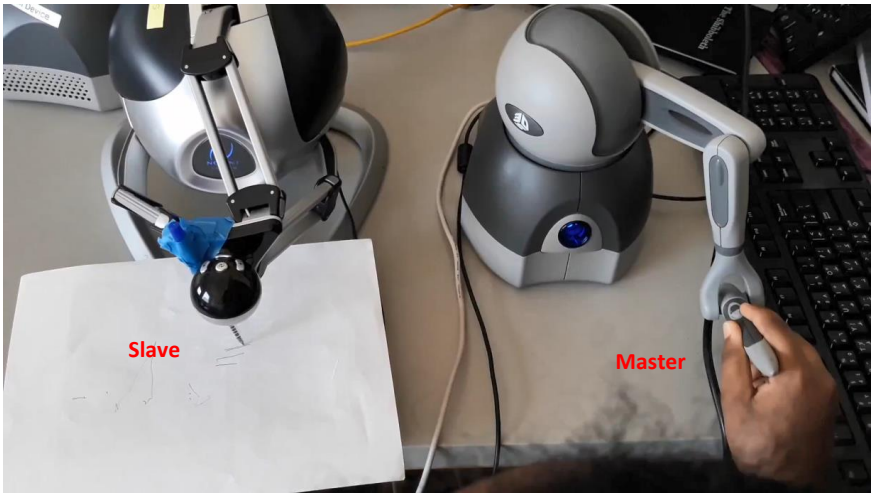
5.5 Qualitative Relationship between QoE and QoS meterics for telehaptic over LTE-A with SDN and MEC

Instead of audio-visual data delivery, TI and DT offer skill-set delivery services and applications. More specifically, content-delivery communication can be utilized to only exchange text messages, voice content, video streaming, and emails, while skill-set delivery networks will take the internet one step ahead by conveying physical touch experiences through haptic devices. However, the TI technology is only in its early stages and its cornerstone, 5G, is not available anywhere. Further, it has been reported that more than 50 billion devices are already connected to the internet in the year 2020 [166]. This certainly creates a huge burden on the current network infrastructures such as the 4G as it can not be able to meet tremendous bandwidth and low latency requirements such as a huge number of devices. Not to forget that the Covid-19 pandemic is not ending soon, hence adding another burden on the internet as most of the daily work is now taking place online. Fired by this motivation, we have endeavoured to assess the correlation between QoE to QoS KPIs created by the incorporating of SDN and mobile edge computing (MEC) technologies to the current wide implemented Long-Term Evolution-Advanced (LTE-A) to realize the deployment of DT skill-set delivery as well as Industry 4.0 in a large-scale. For that purpose, we have created an experimental test-bed to first capture skill-set delivery content between a two-port teleoperation network model, as illustrated in Figure 5.19.

The test-bed consists of a master represented by the user controlling the 3D Touch haptic device and a salve represented by the Omege.3 device which is used to enable the telewriting skill-set delivery. The correspondent workstations of the master and slave are connected via the mobile routers which in turn are configured to use the LTE-A communication channel. The haptic data at the forward and backward channels



a) Logical representation of the experiment test-bed.



b) A real demonstration of the experiment test-bed.

Figure 5.19: Testbed configuration: a) represents the logical scheme of the MST. B) shows a real demonstration of a skill-set delivery (tele-writing) over LTE.

use the non-proprietary UDP transport protocol (ports: 886-889) while delineating a specific route between them. In other words, all incoming and coming traffic have to pass through the stations running Wireshark open-source packet analyzer to be captured before transmitting through the LTE-A routers. This configuration allows us to capture the packet being interchanged between the master and slave for further interpolation

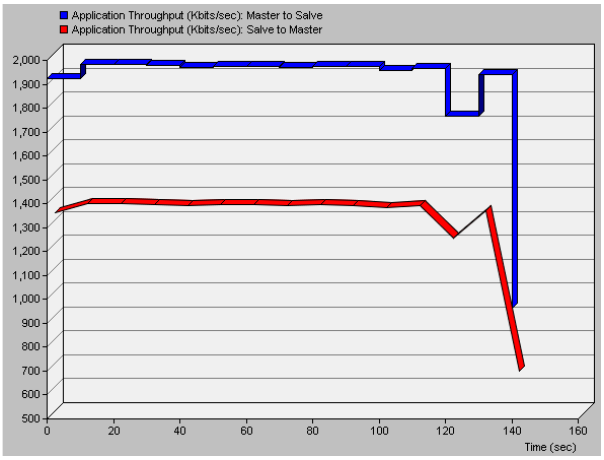


Figure 5.20: Application throughput of two port network MST over LTE.

via the Riverbed Modeler. The application throughput for the MST DT tale writing skill-set delivery is shown in Figure 5.20. It is clear that we have deployed the DP with 10% reduction rate as the traffic from the master to the slave is margined around 2000 Kbps while the traffic from the salve to the master (which deploys 0.1 DB for the force feedbacks) is approximated around 1,300 Kbps.

5.5.1 LTE-A Network Simulation with SDN and MEC Integration, for Haptic Audio/Video (HAV)-Large scale.

The haptic traffic generated from the aforementioned test-bed was used as a user application profile in the riverbed modeller environment. According to Figure 5.21, we have deployed a 7 cells LTE-A based network with randomly distributed devices that move within the cells. Each cell consists of an Evolved-Node-B eNodeB base station along with five randomly distributed devices (User Equipment) that can move on a defined vector-based trajectory. The devices that were used in simulation topology are described in Table 5.2. In this work, the LTE-A simulated by the Riverbed Modeler is release 13. The latter uses OFDMA (or Orthogonal Frequency Division Multiple Access) with MIMO (or Multi-Input Multi-Output): OFDMA is a multi-user variant of the famous

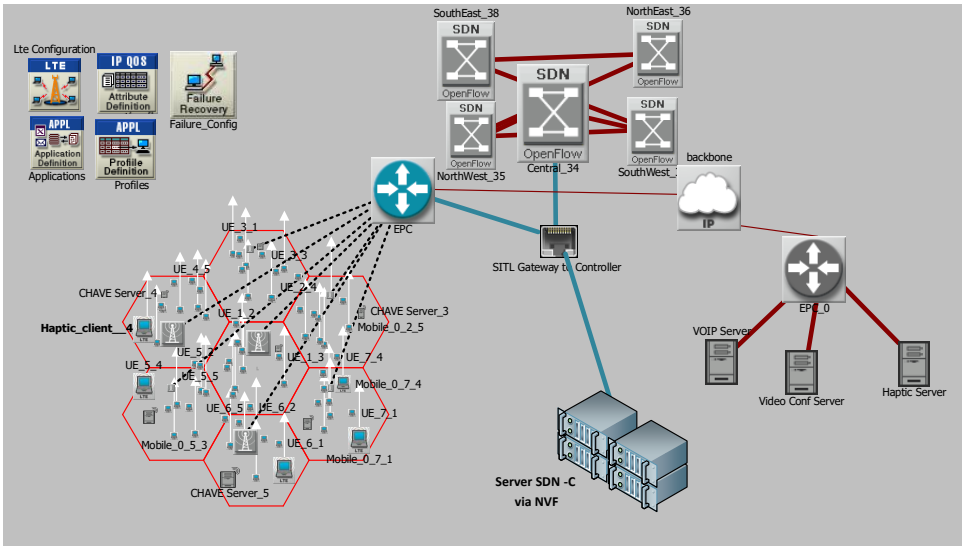


Figure 5.21: Riverbed simulation test-bed .

Orthogonal Frequency Division Multiplexing (or OFDM) digital modulation technique. OFDMA is considered as profoundly reasonable for broadband mobile communications because of its benefits which range from scalability, adaptability, and utilization of different user-friendly multiple antennas (MIMO), as well as its capacity to exploit channel frequency selectivity. As can be seen, we have considered the MEC approach by bringing the computational and storage resources for the haptic audiovisual media closer to the edge of the network. In other words, we have deployed the HAV server for each cell. Consequently, the end-to-end global latency should be reduced significantly. In addition, the QoS orchestration is done using an SDN controller. In this work, the OpenDaylight SDNC is a Java Virtual Machine (JVM) software. It is also recognized as the OpenDaylight platform and can be run from any operating system or any hardware since it supports Java API. The OpenDaylight Controller features open northward APIs, used by the applications. These applications: (1) use the controller to gather information about the network, (2) run algorithms to operate analytics, and afterwards, (3) exploit the Controller to build new rules all over the network. It should be noted that we have configured the controller to dynamically assign more priority to the haptic traffic in case

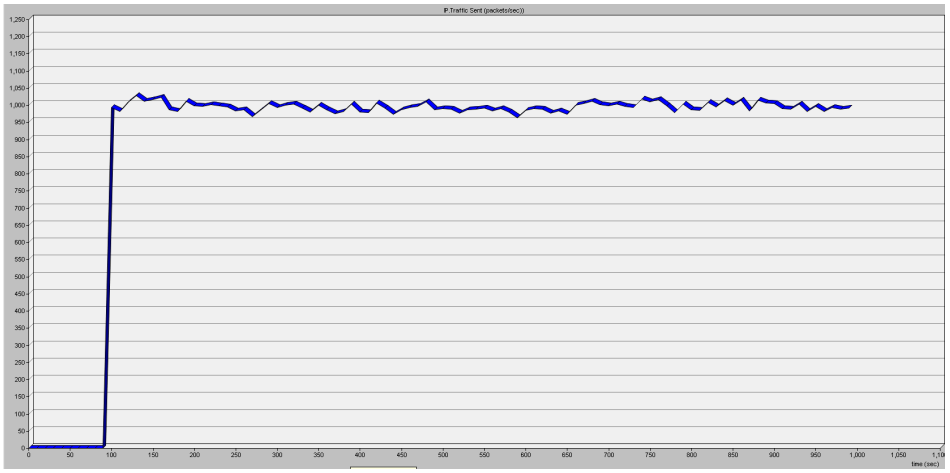
Table 5.2: SDN-MEC scenario parameters.

Component	**Riverbed Modeler Entity**
Base Station	lte_enodeb_4ethernet_4atm_4slip_adv
IP Backhaul	router_slip64_dc
Gateway	lte_access_gw_atm8_ethernet8_slip8_adv
Number of Cells	7
Cell Radius	1KM
Number of UEs per Cell	5 LTE UE (User Equipment)
UE Speed in Movement	10 Kmph
Standard	LTE-A Rel.13
Bandwidth	20 MHZ (FDD)
Modulation type	OFDM
Multimedia Server	ethernet_server
Workstation Application	lte_wkstn_adv
LTE Server	Lte_server_adv
OpenFlow Switch	of_switch_eth16_adv
SDN Controller	OpenDaylight
System in the loop	sitl_virtual_gateway_to_real_world

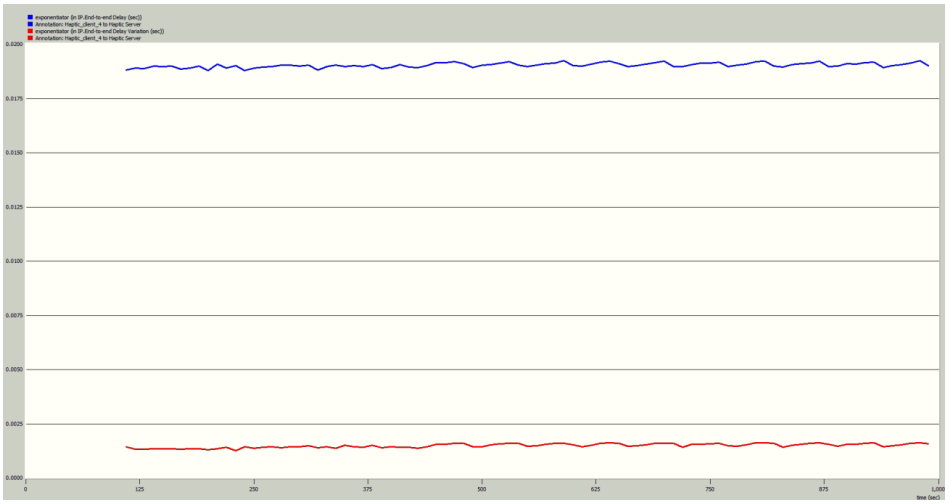
of communicating with the remote haptic server or in the case of MEC system fail over.

The SDN controller has a synopsis of the total network and is in charge of the settlement to be taken, whereas the hardware (switches, routers, etc.) is merely in charge of expediting packets to destination using a set of packet-handling rules. The controller is installed on a virtual machine to deploy Network Function Virtualization (NFV) i.e., network functions can be hosted on virtual machines, in order to provide network scalability, which corresponds to the possibility of making functional additions without the need to install or acquire specialized equipment. Thanks to the STIL Riverbed Mod-

eler's feature. SITL bridge serves as a foreign device whereby the simulation transfers the packets from e 'OpenDaylight' SDN controller (which is installed on an external machine) to the simulation process on Riverbed. In such a simulation manner, physical hardware and simulation can interact as a unified system.



a) Data rate for UE Mobile HAV



b) Overall Latency and Jitter UE Mobile HAV

Figure 5.22: Evaluation of KPIs of LTE-A network with MEC-SDN.

Figure 5.22 shows the QoS KPIs for the exchanging of haptic audiovisual streams over the LTE-A network. By bringing the tactile applications' network functions at the edge of the mobile access network, i.e., close to the UEs, the data rate for transmission

of the multimodal stream is maintained at around 1000 PPS. As can be seen in 5.22.b the delay variation is maintained around 2.5 ms (red line) while the end-end delay is approximated by 18 ms (blue line). We believe that the is the best performance that can be achieved in this simulation topology as the tactile internet demands that the physical transmission must have very small packets to enable a one-way physical layer transmission of 100 μs [167]. However, the duration for the one orthogonal frequency division multiplexing (OFDM) symbol alone is close to 70μ long under current LTE-A cellular systems.

The haptic audiovisual traffic was fed to the ACE Analyst which is part of the riverbed modeller that can analyze and reshape applications' performance over WAN accelerators. We have utilized the ACE module to study the correlations of QoE in term of application response time and the impact of available QoS parameters namely bandwidth and latency for skill-delivery communication. To estimate the parameters of each QoE-QoS relationship, the curve estimation and non-linear regression methods are deployed. Goodness-of-fit tests are used to determine which relationship and function type best explain the correlation of QoE and QoS.

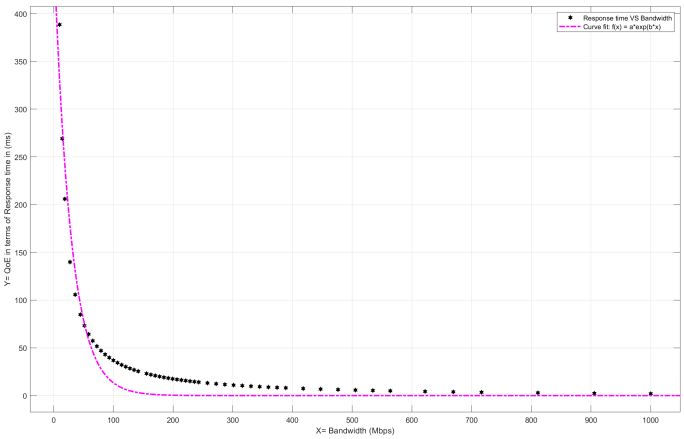


Figure 5.23: Available bandwidth influences on response time.

Figure 5.23 illustrates the impact of the available communication bandwidth on the

HAV application. As can be seen, the one-degree exponent function is used to assess the interdependence of the response time (QoE) over the availability of the communication bandwidth. Based Figure 5.23, the QoE-QoS relationship can be described as:

$$QoE(QoS) = a.e^{b*QoS} \tag{5.5}$$

We have used the Matlab curve fitting toolbox to calculate the coefficients (with 95% confidence bounds): a =470.2 with (lower bound= 415.4, and upper bound = 525.1), b =-0.03556 with (-0.0406, -0.03052) to represent the lower and upper bounds respectively. The goodness of fit for the model function gives a coefficient of determination R^2 equal to 0.9268.

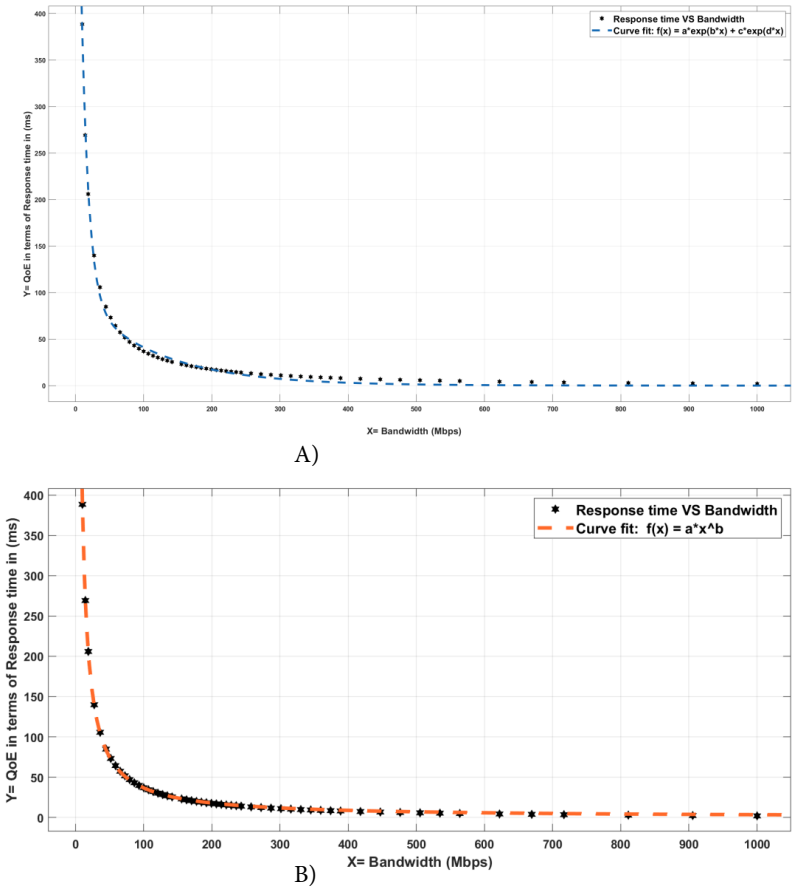


Figure 5.24: Mapping Curves for $QoE(Bandwidth)$.

Figure 5.25 shows more optimal fitting functions to quantify the impact of bandwidth on the HAV application response time. Obviously, the two-terms exponent function outperforms the previous one-term exponent function with R^2= 0.9961. Hence, the QoE-QoS interdependencies can be denoted by equation 5.6:

$$QoE(QoS) = a.e^{b*QoS} + c.e^{d*QoS} \tag{5.6}$$

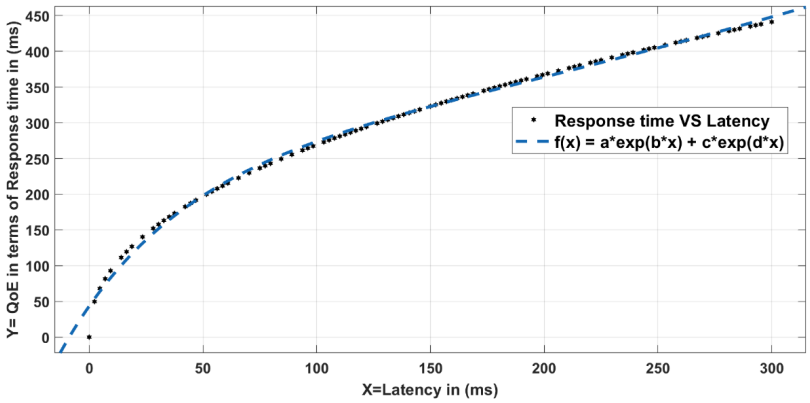
Using the optimization toolbox of Matlab, the values of the coefficients are calculated (with 95% confidence bounds) as the following:

- a =746.1 $\leftrightarrow$(688, 804.2)
- b =-0.09409 $\leftrightarrow$ (-0.1018, -0.08634)
- c =97.89 $\leftrightarrow$ (87.09, 108.7)
- d =-0.008678 $\leftrightarrow$ (-0.009641, -0.007714)

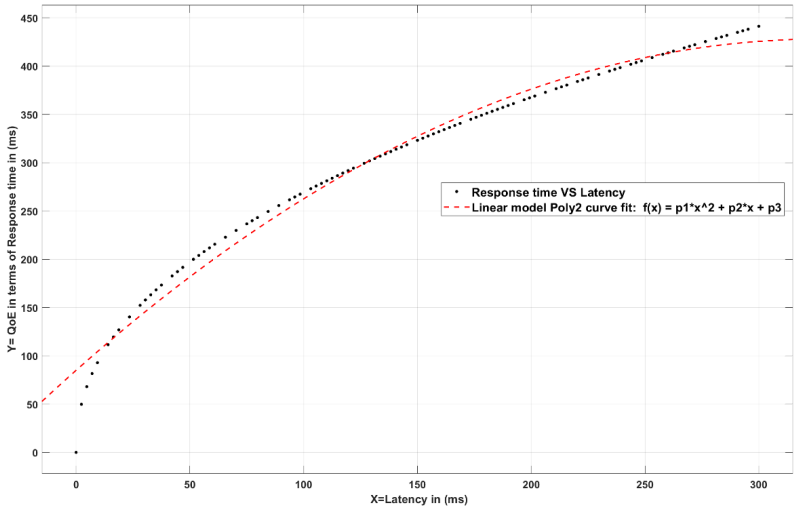
Note that the power mapping curve fit is the ideal function to describe the readership of the application response time with a variation of the throughput, 5.25.b. It yields a coefficient of determination R^2= 0.9998.As such, to best describe QoE in terms of available bandwidth, we should use equation 5.7:

$$QoE(QoS) = a.QoS^b \tag{5.7}$$

The values of a and b are estimated as 4145 and -1.027 with 95% confidence interval. With regards to studying the impact of communication latency on the networked HAV application response time, Figure5.25 shows measurement of QoE with increasing network delay. With R-square approaching to 1, It is crystal clear that the two-terms exponential function form can adequately explain the perception-centric relationship between QoE and QoS in this regard. Hence QoE in terms of delays can be also quantified using equation 5.6. The values of coefficients to realize function 5.6 (with 95% confidence interval) are as follow: $a = 247.1$, $b = 0.001988$, $c = -203.6$, and $d = -0.01999$.



a) two terms exponent curve fitting



b) Second degree polynomial curve fitting

Figure 5.25: Mapping Curves for $QoE(latency)$.

It should be noted that a linear model Poly2 also performs well, with R-square: 0.9835, in describing the influence of delay on the HAV response time. Consequently, we can quantify QoE in terms of varying delays using :

$$QoE(QoS) = p_1 QoS^2 + p_2 QoS + p_3 \tag{5.8}$$

where the values of p_1,p_2,p_3 are equal to -0.003202, 2.096, and 84.95 respectively with 95% confidence interval-calculation .

5.6 General Remarks:

Based on the dynamics of the remote environment i.e, where the DT is located, the QoS requirements and KPIs of the teleoperation DT systems considerably differ. For the case of a highly dynamic remote environment, an example of that is a tele-basketball where the user interacts remotely with his DT in VR/AR basketball match as shown in Figure5.26), to maintain stable operations with quick-moving VR objects), the sending and receiving of haptic stimuli are tremendously time-critical, with an end-end delay has to be margined between 1-10 ms.

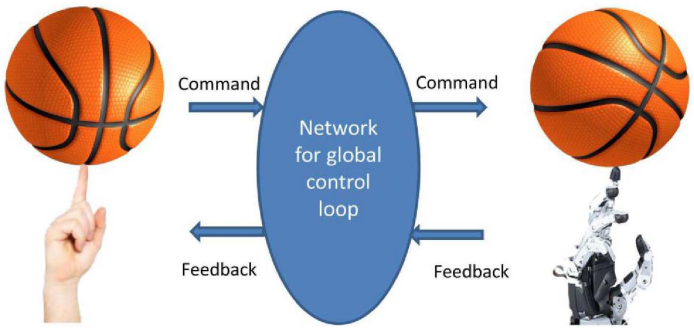


Figure 5.26: Telebasketball over TI adapted from [8]

For the case of a medium dynamic teleoperation scenario (such as tele-rehabilitation), the end-end delay for the exchange of haptic data can be relaxed to 10-100ms, since the teleoperator can cope with deformable objects and react and move with lowered velocity. In the case of quasi-static or static environments such as tele-maintenances, the end-end delay KPI can be more extended by 100 ms to 1s. it is worth noting that for the last two cases, control architectures have to be implemented as well as cooperatively constructed with the QoS communication architectures to stabilize the remote interaction. Examples of stabilized control architecture are 1) Teleoperation Time-domain Passivity Approach with Perceptual Deadband based haptic data reduction (TDPA+PD) [159] as well as 2)Model-mediated Teleoperation in cooperated with Perceptual data reduction (MMT-PD) [168]. The former is suitable for short-range tele-haptic applications that run at the

network edge and support high dynamic haptic exchange (motion, deformation, forces, etc) between the master and the slave, while the latter is suitable for medium to long-range teleoperation scenarios which are described by a low level of haptic updates between the master and the salve. It is worth noting that TPDA-A outperforms (MMT+PD) in terms of stability however it is more susceptible to communication delay. Tables 5.3 and 5.4 illustrate the KPI Requirements and traffic characteristics for the DT teleoperation systems over TI as well as summarize the Higher-layer requirements and capabilities in different teleoperation scenarios.

Table 5.3: KPI Requirements and Traffic Attributes for DT Teleoperation over TI

Traffic characteristics description		Master ➔ Slave Local User ➔Remote Participant	Exchanged Stimuli			Slave ➔Master Remote Participant➔ Local User
Traffic types +		Haptics (position, velocity, angular velocity)	Video	Audio	VR/AR	Haptic feedback (forces, torques, vibrotactile signals)
Burst size		1 DoF: 2-8 B 3 DoFs: 6-24 B 6 DoFs: 12-74 B	MTU~1.5 KB	~50 B	MTU~1.5 KB	- Kinesthetic signals 1 DoF: 2-8 B 3 DoFs: 6-24 B 6 DoFs: 12-74 B - Tactile signals 1 DoF: 2-8 B 10 DoFs: 20-80 B 100 DoFs: 200-800 B
Reliability ([%])		99.9 (w/o compression) 99.999 (w/ compression)	99.999	99.9	99.9	99.9 (w/o compression) 99.999 (w/ compression)
Latency([ms])	High-dynamic environment	1-10	10-20 *		*1-10	1-10
	Medium-dynamic environment	10-100	30-40 *		*<100	10-100
	Static or Quasi-static environment	100-1000	50-150 *		*<300	100-1000
Average data rate		-1000-4000 packets/s (w/o compression) -100-500 packets/s, (w/ compression)	1-100 [Mbps]	5-512 [Kbps]	600 Mbps	-1000-4000 packets/s (w/o compression) -100-500 packets/s, (w/ compression)
Arrival model		-Periodic (w/o compression) -Gilbert-Elliot (2- state discrete Markov chain) (w/ compression)	Periodic	Periodic	Aperiodic	-Periodic (w/o compression) -Gilbert-Elliot (2- state discrete Markov chain) (w/ compression)

+ For position-force bilateral teleoperation architecture (with force signals feedback from the teleoperator to the user).
* For the synchrony of video/audio and haptic feedback (with haptic feedback goes first).

Table 5.4: Higher-layer requirements and capabilities in different DT Tele-operations

Environments / Requirements & Metrics	Teleoperation is a high-dynamic environment	Teleoperation is a medium-dynamic environment	Teleoperation is a static or quasi-static environment
Stability	Guaranteed control and QoS architecture)	Guaranteed (control architecture and QoS required)	Guaranteed (control architecture control and QoS required)
Task performance	High	High/Medium	High
User experience	High	High/Medium	Medium/Low
System transparency as a function of control algorithm and communication delay	High	High	Low
Asynchrony of video and haptic modality	Low	Medium/Low	High/Medium

5.7 Summary

The Tactile Internet where haptic is appended to the traditional multimedia communication is one of the main technologies used in the development of future network infrastructures such as the fifth generation of mobile networks (5G). When it comes to network resources, haptic is very challenging in terms of latency, jitter and packet loss. We have benchmarked the performance of the two well-known stabilizing control teleoperations (TPDA+PD) and MMT+PD. With the facility of the riverbed modeller, we have modelled the networked haptic flow sent by haptic based-ALPHAN application over the Internet. The model was used to evaluate the performance of haptic in a large-scale environment. Based on our experimental study, the provision of the DiffServ priority queuing or custom queuing with LLQ dramatically enhanced the performance of the ALPHAN haptic transmission over a congested IP network. We have also studied the impact of utilizing the advanced (SDN) QoS solution along with MEC on haptic-audiovisual transmission over LTE-A networks. The outcome of this study allows us to correlate QoE to QoS in terms of bandwidth and latency KPIs.

Chapter 6

QoE Use Case 2: Predicting and Quantifying Fatigue for Haptic-Visual System

As has been discussed throughout this thesis, appending and augmenting haptic to everyday life activity is the main objectives of the new technologies such as tactile internet, internet of skills along with their applications e.g. Health 4.0 and Industry 4.0. Since haptic interactions imply continuous arm movements, fatigue is an important physiological aspect that strongly correlates with the QoE of haptic-based application. In some circumstances, fatigue could be the desired effect, e.g, virtual reality fitness application, where the goal is to improve the user's endurance and physical exercising. Cumulative muscle fatigue may cause musculoskeletal injuries and disorders [169]. Therefore, there is a demand to measure and detect fatigue for haptic-based applications especially when Tactile Internet becomes the most prominent networking paradigm.

Fatigue is defined as a reduction in the physical movement activities associated with energy dissipation in the muscle caused by internal and external forces [170]. This definition is related to physical fatigue. On the other hand, there is another form of fatigue, referred to as mental fatigue, which is a result of prolonged periods of cognitive ac-

tivity [171]. However, since mental fatigue is not strongly connected to the muscular interactions with the haptic interface, it is kept out of the main scope of this chapter.

Fatigue can be analyzed subjectively through surveys and observations by users who interact with the haptic interface. The involvement of the users in the assessment process implies that this approach is very expensive, time-consuming, and lacks repeatability. It also cannot be applied in a real-time fashion. In this chapter, we show our framework towards objectively and simultaneously detect the user's arm fatigue when haptically interacting with a multimedia application. In essence, we look into the surface electromyogram signals and their power spectrum's attribute [172] generated from the typical five-arm muscles looking for signs of fatigue. The trends of the median and mean frequency spectrum serve as indicators that can be used by the developers of the haptic application to steer the muscle interaction contextually to increase the QoE of the users. The rest of the chapter is organized as follows; In section 6.1, we review the related work concerning the analysis of fatigue in the multimedia domain. Section 6.2 and 6.3 describe the experimental setup and the collected data. In Section 6.5 and 6.6, we present the results and analyze them from three different angles (raw EMG analysis, MNF/MDF analysis, subjective analysis). We conclude the chapter in section 6.7.

6.1 QoE Fatigue Measurements in Multimedia

In the literature, there are two main objectively approaches to measure physical fatigue: simulation-based methods and experiment-based methods. The first approach is adopted by Kahol et al. [173] where a haptic enabled virtual reality simulator is used to measure the fatigue of physicians by evaluating their task completion time, psychomotor and cognitive skill assessment. Besides, authors in [174] used Hill's model to create a framework capable of calculating the effect of muscles' fatigue on the force-time relationship of skinny muscle by using the PAK program. Ma et al. [175] proposed a framework to evaluate fatigue in maintenance and assembly operations by using a digital human sim-

ulation in a virtual environment. In which, the muscles motion was parameterized into a set of virtual human aspects. The fatigue was evaluated by measuring the endurance time as well as the reduction of joint strength. The virtual framework was evaluated through an avatar carrying out a hole-drilling task. Regarding the experiment-based methods, the most prominent approach is the use of a surface electromyogram (sEMG) measurement system [172]. It is a wearable arrangement that can be used in a wireless mode with a semi-invasive technique i.e., does not need punctures, and thus it maximize the subject's comfortability. Kotani and Horii [176] measured the comfortability of using a pen-tablet input system compared to a typical mouse-keyboard system. For that purpose, the authors captured the EMG signals from different arm's positions namely biceps brachii, flexor digitorum superficialis, extensor digitorum, and trapezius in order to compare the amplitude of EMG muscles signals while using both inputting systems and concluded that the skill process for the pen-tablet system was very short and the subjects felt comfortable to use this inputting system from the beginning. The authors in [177] developed a prototype of a wearable fatigue-tracking system using the experiment-based methods to monitor the overall fatigue status of the human body movement based on the mean frequency of the sEMG signal. The system was proven subjectively by conducting two different experiments with 17 users in which each fatigue level was found to be negative. However, this framework was designed to measure the whole body fatigue under the extreme physical exercising condition and not tested for a special case, e.g, the user repetitively manipulates the haptic device.

6.2 System Framework

The framework for this case study is illustrated in figure 6.1 includes a surface electromyogram station, sEMG sensors with triaxle accelerometers from Delsys Inc [172], Geomagic Touch [25], and a central workstation, which is running a wireless-trigno sensor acquisition system. The geometry touch device is selected due to its six-degree-of-freedom

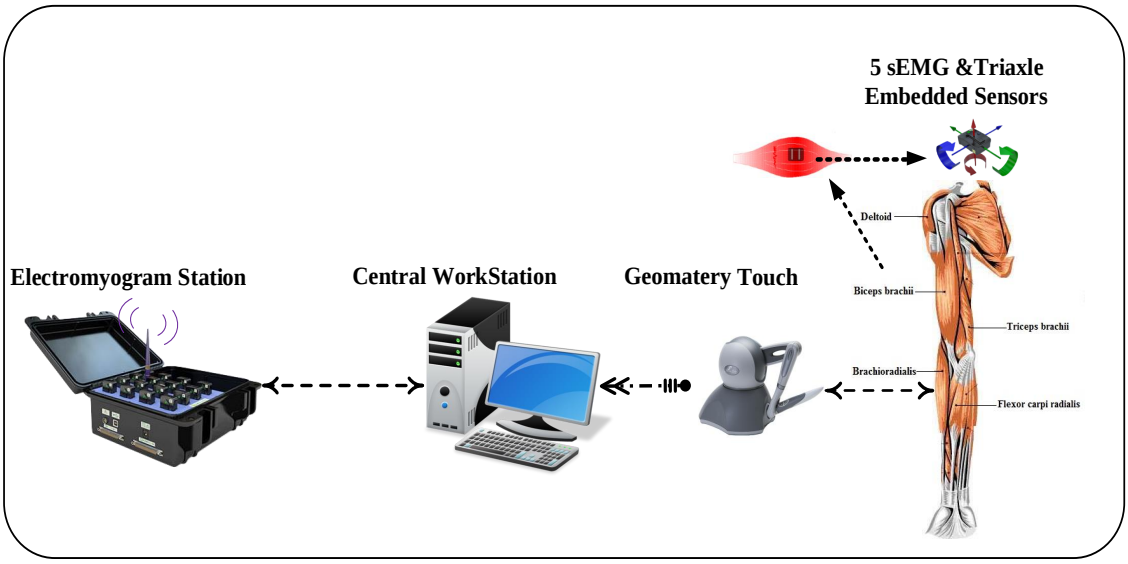


Figure 6.1: System architecture of the fatigue-deducing system for haptic-based applications.

positional sensing, three degrees of freedom of force feedback, ease of use, and automatic workspace calibration. The haptic writing application was hosted on a Unity-based VR server and connected to the user via a bilateral ultra-high bandwidth with a lag-free communication channel. The environment, shown in figure 6.2, is built using C Sharp programming language in which a Haptic Plug-in for Unity [178] is used to give the subjects a realistic writing experience while using the phantom interface.

While conducting repetitive haptic writing tasks, five muscles, named: flexor carpi radialis, brachioradialis, triceps brachii, biceps brachii, and deltoid, are mainly responsible to constantly move the subject's arm. Therefore, the sEMG sensors are placed at the center of the aforementioned muscles. These sensors are connected wirelessly with the main workstation that simultaneously analyzes the online received data from the mobile electromyogram station. It is worth noting that triaxle accelerometers can record dynamic muscle movement and interact simultaneously with the sEMG recordings.

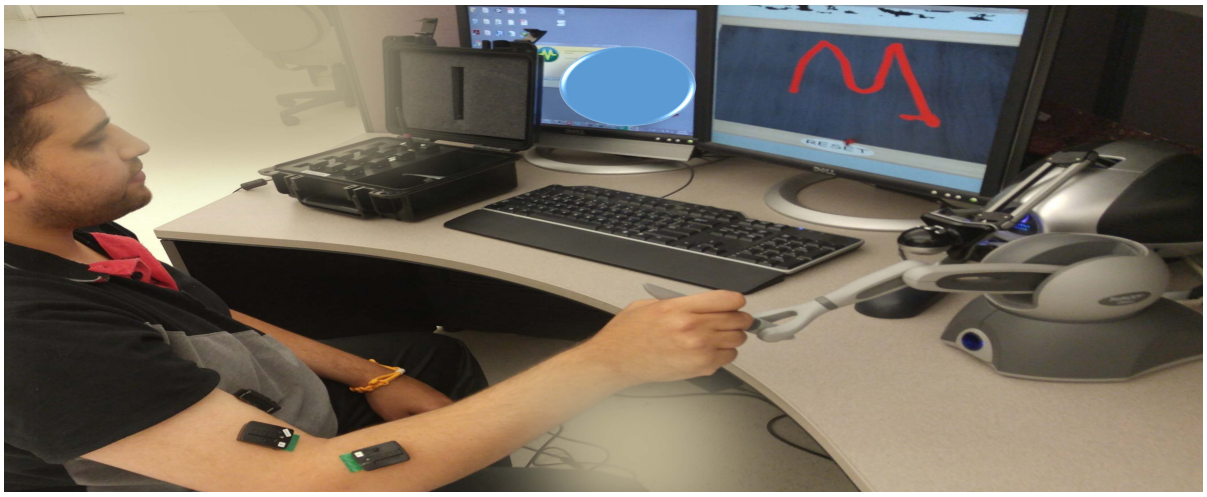


Figure 6.2: A screen shot of the Haptic Learining Tool. The subject was asked to write the illustrated character during the fatigue experiment.

6.3 Experimental Setup

Twenty volunteers, including twelve males and eight females, with different ages (25-59) and different haptic VR backgrounds have participated in the experiments. Before initiating the haptic tasks, we briefly went over the experiment's goal and steps. The volunteers were advised to wipe the skin of their five muscles with rubbing alcohol, then guided to attach the sEMG sensors at the center of each muscle via adhesive tapes. To confirm the accuracy of our results, the experiment did not begin until the subjects felt comfortable with the setup and tried the haptic writing task twice. After that, each subject was asked to re-write a given Arabic alphabetical character and perform a haptic signature repetitively. The duration of the trial was 400 seconds. During that period, the sEMG and acceleration signals were captured via EMGworks 4.7.3 Acquisition software. At the end of the experiment, the subject was asked to submit his/her feedback using a short questionnaire regarding his/her arm fatigue level. Most questions have answers from 1 to 7 with two main anchors (not at all and completely) to reflect the best subject's experience.

6.4 Position and Force profiles results

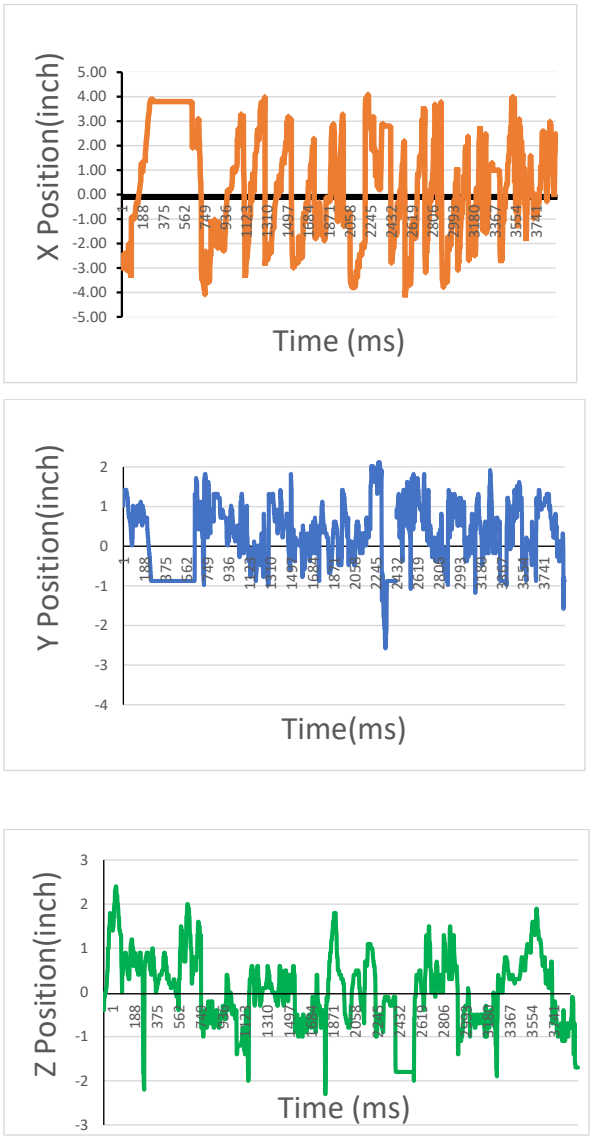


Figure 6.3: Position's profiles for Subject 1.

Figure 6.3 shows the position profiles for subject1 extracted from almost 4000 time-stamps. As can be seen, the subject stretches his arm to the three possible orientations (X,Y,Z) while conducting the experiment. This confirms that the task given to the subjects demands a continuous movement from the subject's arm.

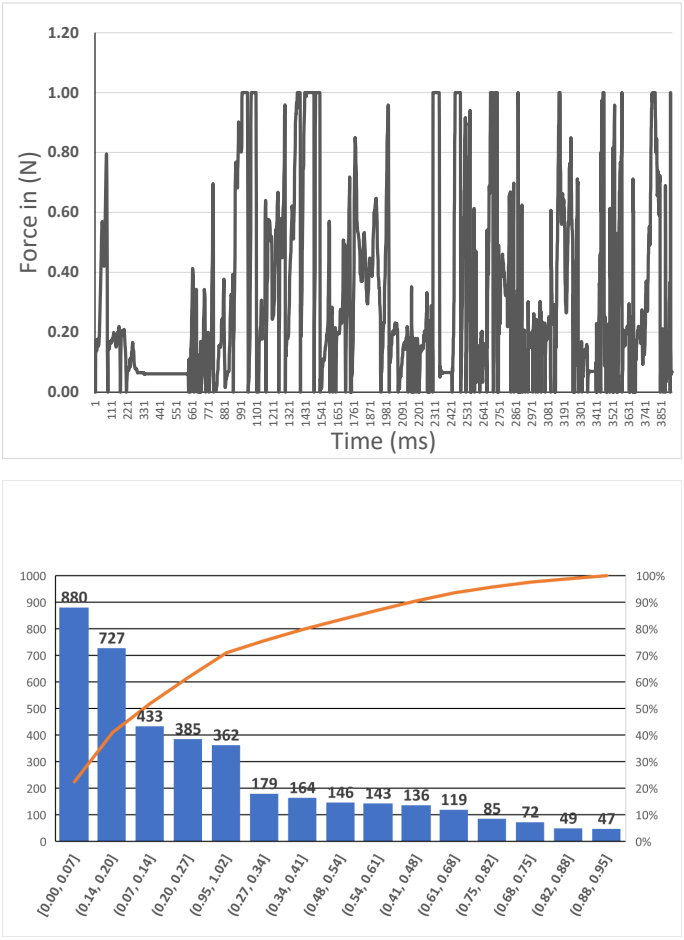


Figure 6.4: Normalized cumulative Force Profile for Subject1 with Pareto analysis.

Figure 6.4 illustrates the normalized cumulative force profile of subject1 while experimenting with the given exercises. The force profile is a representation of the force at all timestamps from all trials. The cumulative force is computed by $F_c = \sqrt[2]{f_x + f_y + f_z}$. where f_x,f_y,and f_y are the force vectors in the x,y, and z directions. It can be noticed that the haptic interactions require a heavy and frequent workload on the subject's muscles as the user reaches the maximum force i.e., 1, 362 times. Another interesting observation is shown by the Pareto histogram which indicates that the subject got tired over the time of the experiment as a force bucket of 0.0 to 0.27 is the most frequent generated force i.e., 80% of the forces needed to finish the experiment.

6.5 sEMG Signal Analysis

For the sake of simplicity, we start this section by discussing the results from the surface electromyogram signals for subject 1, then we present the comprehensive outcomes for the muscles of all subjects and we finally link them to feedbacks from the subjective questionnaire. Figure 6.5.a and 6.5.b illustrate the filtered sEMG signal from the biceps brachii and the flexor carpi radialis respectively. The two signals are filtered by the Bessel filter to provide monotonic and smooth magnitude responses by eliminating ringing issues caused by sharp inputs. The sampling frequency of the sEMG signals was set at 1 kHz. The Y axes in the aforementioned figures is measured in Volts, whereas the X axes is measured in Second. Even though the Biceps Brachii gives a steady voltage of approximately 38mV, flexor carpi radialis was the heavy loaded muscle with 80mV.

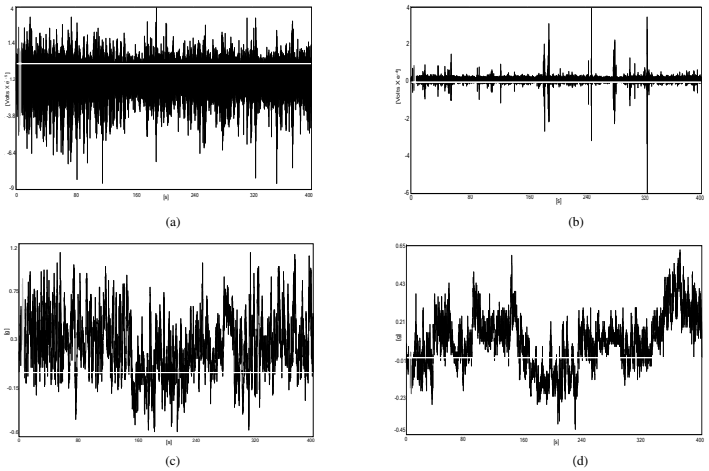


Figure 6.5: EMG Results and Analysis for Subject 1. (a) Filtered sEMG signals for Biceps, (b) Filtered sEMG signals for Flexor Carpi Radialis, (c) Acceleration signal from Brachioradialis (ACCy), (d) Acceleration signal from deltoid (ACCy).

Figures 6.5.c and 6.5.d demonstrate respectively the correspondent re-sampled rotational acceleration signals (ACCy) generated from the Brachioradialis and Deltoid. The resolution of the acceleration signal is computed with 8 bits over the full 3.3V dynamic range, covering accelerometer ideal maximum outputs of ± 1.5g. The ACCy signals

are measured with respect to the standard gravity g, which is equal to $32.174 ft/s^2$ or $9.8 m/s^2$. Notice that deltoid ACCy signal has a periodic pattern while the Brachioradialis ACCy is a non-periodic signal. In addition, the highest magnitude value of the Brachioradialis ACCy, i.e., 1.1g, indicates that the movement of that muscle was slightly higher than the movement of the deltoid with a peak value of 0.65g.

6.6 Frequency Analysis

Generally, The analysis of sEMG and acceleration signals is commonly used as muscle movement /pattern /force detection index whereas its performance and accuracy to detect muscle fatigue are questionable[179]. On the contrary, frequency attributes[180] mean and median frequencies were proven to be efficient to detect muscles fatigue and that's what we will explore in the upcoming section.

6.6.1 Mean Frequency Analysis (MNF)

MNF is the average frequency which can be computed by taking the sum of the product of the EMG power spectrum of a given muscle and the frequency divided by the total sum of the power spectrum. Therefore, and based on the aforementioned definition, MNF after a given interaction time t can be calculated using:

$$MNF = \frac{\int_0^\infty \omega Po_{sEMG}\ (\omega)\, d\omega}{\int_0^\infty Po_{sEMG}\ (\omega)\, d\omega} \tag{6.1}$$

and can be simply expressed by

$$MNF = \sum_{j=1}^{M} f_j Po_j / \sum_{j=1}^{M} Po_j \tag{6.2}$$

Where: M is the length of frequency bin, Po_{sEMG} at frequency ω is equal to Po_j which is the EMG power spectrum at the frequency bin j, and can be computed via fast Fourier transform (FFT):

$$Po_{sEMG}(\omega) = FFT_{sEMG}(\omega)\, FFT^{*}_{sEMG}(\omega)$$
$$= \mid FFT_{sEMG}(\omega) \mid^{2}$$

Where $FFT_{sEMG}(\omega)$ indicates the fast Fourier transform of a given muscle sEMG signal. $FFT^{*}_{sEMG}(\omega)$ is the complex conjugate of $FFT_{sEMG}(\omega)$. According to a previous analysis [177], the mean frequency decreases linearly with muscular interaction time as the fatigue gradually increases. In other word, the mean frequency degrades with the increase of the fatigue intensity. As such, the decrease of mean frequency satisfies a linear correlation with haptic-interaction time of a muscle under fatigue.

$$MNF_t = \lambda_t \cdot t + \Delta_t \tag{6.3}$$

where λ_t is slope parameter of the linear regression fit model.Δ_t is remaining haptic task time. t is working time under fatigue status of the muscle. Hence, the mean frequency fatige index for a given muscle m_i after performing a haptic task in time t can be relatively calculated using:

$$FI_{MNF_{m_i}} \simeq \mid \frac{MNF_t}{MNF_0} - 1 \mid \tag{6.4}$$

where MNF_0 and MNF_t, indicate the mean frequency of sEMG signal of a given muscle m_i at the start and end time of the haptic task respectively. Using equation 6.5, The overall MDF fatigue index represents the average percentages of the relative decrease of the mean frequency of flexor carpi radialis, brachioradialis, triceps brachii, biceps brachii, and deltoid.

$$Overall - FI_{MNF} = \sum_{1}^{m} \left[\frac{FI_{MNF_{m_i}}}{m} \right] \tag{6.5}$$

Please note that the FI_{MNF}statistics for all subjects are shown in detail for all subjects in (Table 6.1).

6.6.2 Median Frequency Analysis (MDF)

Compared to MNF, the analysis of Median Frequency (MDF) of the muscular sEMG siginal, is proven to give better results, because MDF is not susceptible to extremes in

the frequency range as well as it is more immune against random noises, and therefore gives more sensitive and precise measures[181, 180]. The Median Frequency (MDF) is the frequency that divides the sEMG power spectrum into two areas with the same amount of power. MDF can be given using the following formulas:

$$\int_{0}^{MDF} Po_{sEMG}(\omega)d\omega = \int_{MDF}^{\infty} Po_{sEMG}(\omega)d\omega \tag{6.6}$$

$$\sum_{i=1}^{MDF} Po_i = \sum_{i=MDF}^{M} Po_i \tag{6.7}$$

$$\sum_{i=MDF}^{M} Po_i = \frac{1}{2}\sum_{i=1}^{M} Po_i \tag{6.8}$$

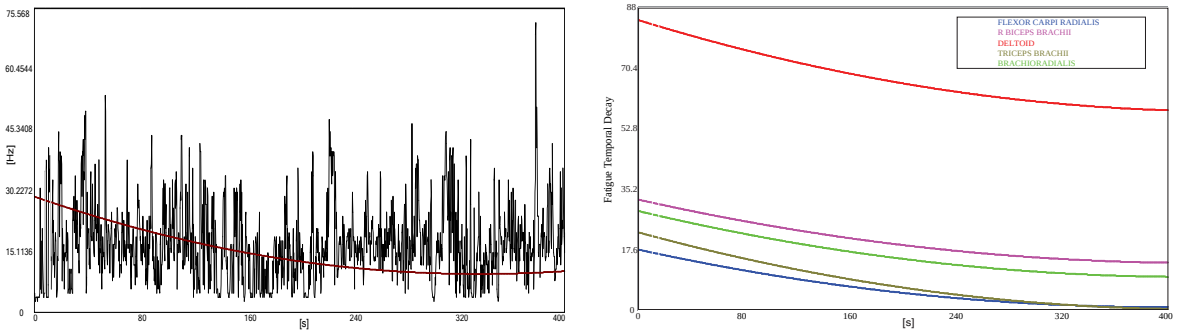


Figure 6.6: MDF Signal Analysis for Subject 1. (a) Curve fit of Brachioradialis MDF signal, (b) Temporal decay for subject 1 arm's muscles.

In the same domain, figure 6.6.a shows the median frequency pattern of the brachioradialis muscle. The procedures to extract the MDF from the sEmg muscle power spectrum is similar to MNF. Hence, The specifications of the moving window used to generate the MDF signal are: the window length is equal to 0.125s, i.e, M=8, while the window overlap is equal to 0.0625s [182]. The red subplot in figure 6 represents a second-order polynomial curve fit that was chosen due to its ideal fit for the median frequency

spectrum's expected decay pattern.The same trend was consistent among all subjects. From figure 6.6.b , It is obvious to observe that the median frequency patterns of the flexor carpi radialis, brachioradialis, triceps brachii, biceps brachii, and deltoid muscles decrease over time, which is an indication of muscle fatigue. We used the slope of the decreasing pattern of the second polynomial regression model that fits the muscle sEMG MDF signal as an adequate predictor to quantify the fatigue level of that corespondent's muscle.

Algorithm 1 MDF Fatigue measurement

```
1:  procedure SUBJECT'S MUSCLE
2:      for each subject do
3:          for each muscle do
4:              GET EMG signal
5:              GET Accy signal
6:              do FFT_sEMG(ω)
7:              Calculate MDF
8:              V= second polynomial regression
                   fit of MDF
9:          end for
10:         FI_MDF = d²V/dt²
11:     end for
12: end procedure
```

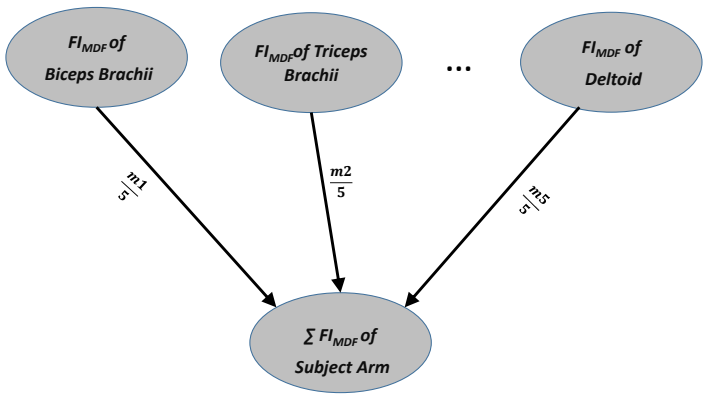


Figure 6.7: Overall MDF Muscular Fatigue Index Calculation.

As such, the procedures of overall MDF fatigue level computation are summarized in

Algorithm 1 and used to compute FI_{MDF} which denotes the cumulative average percentage of the subject's arm fatigue based on the MDF analysis for the five aforementioned sEMG muscles.

6.6.3 Questionnaire Results

To evaluate and validate our results from a user's experience perspective, we asked the subjects to fill out and complete a subjective Likert questionnaire after finishing the experiment to build the fatigue ground truth. The questions of the survey are shown in table 6.3. The Likert scale ranges from one to seven (seven denotes a high score). Reported scores of the arm level fatigue questionnaire were translated into a percentage quantity for each subject based on equation (6.9). As illustrated in table 6.2, a detailed statistical analysis was conducted to obtain descriptive statistics for each individual Frequency Fatigue estimators FI_{MNF} and FI_{MDF} and as well as to quantify the forecast KPIs (Mean Absolute Deviation (MAD) the Mean Square Error (MSE), The Root Mean Square Error (RMSE), and the Mean Absolute Percentage Error (MAPE)) between sEMG fatigue indications (MDF and MNF) and the subjective fatigue reported by our participants. Statically, We noticed that there is a slight difference in quantifying the user's fatigue level of the haptic VR application. For example, FI_{MNF} has an average of 57.76% with a 95% confidence range between 50.76 and 64.36, a median of 51.62, a mode of 42.1, a standard deviation of 14.53, and a standard error of the mean(SE_M) equal to 0.325; theFI_{MNF} has an average of 50.76% with a 95% confidence interval of 42.39 and 58.68, a median of 46.88, a mode of 35, a standard deviation of 17.6, and lastly a standard error of the meanSE_M=0.395. In the context of the forecast KPIs, RMSE (6.13) calculation is an excellent candidate to be used in indicating the average model prediction error between actual and predicted fatigue levels. The usage of error square is important to statistically weigh any irregular values that might affect the results. Compared to traditional error calculations such as Error Rate (ER) and Mean Absolute Deviation (MAD), RMSE performs better in terms of error prediction

Table 6.1: Average sEMG Muscular Fatigue Statistics calculated using MNF and MDF respectively.

Subject #	**Overall FI_{MNF} in (%)**	**Overall FI_{MDF} in (%)**
M1	43.22%	35.00%
M2	50.21%	37.00%
M3	49.12%	22.00%
M4	42.10%	59.00%
M5	70.10%	75.30%
M6	40.10%	45.66%
M7	51.74%	21.00%
M8	42.10%	35.00%
M9	53.20%	39.00%
M10	67.10%	59.00%
M11	45.22%	50.00%
M12	51.50%	42.30%
F1	60.30%	45.60%
F2	70.90%	66.60%
F3	49.40%	39.80%
F4	69.91%	65.90%
F5	80.45%	73.10%
F6	50.79%	48.10%
F7	70.49%	71.10%
F8	93.18%	81.87%

Table 6.2: Statistics and Forecast KPIs for Mean and Median Frequency Fatigue Estimators

Fatigue Index	**Mean**	**95% Confidence Interval**		**Median**	**Mode**	**Standard Deviation**	**Standard Error**	**Pearson Correlation**	**R Square**	**MAD**	**MSE**	**RMSE**	**MAPE**	**P(T<=t) two-tail Test**
		Lower Bound	Upper Bound											
FI_{MNF}	*57.56%*	*50.76%*	*64.36%*	*51.62%*	*42.10%*	*14.53%*	*3.25%*	*78.50%*	*68.86%*	*6.69%*	*1.12%*	*10.60%*	*15.84*	*0.92*
FI_{MDF}	*50.62%*	*42.39%*	*58.68%*	*46.88%*	*35.00%*	*17.60%*	*3.95%*	*95.40%*	*91.01%*	*4.15%*	*0.52%*	*7.21%*	*8.34*	*0.34*

accuracy as well as monotonicity since it has the advantage of penalizing large errors in predicting observations according to variables of previous results. As can be observed from the aforementioned tables, the FI_{MDF} index has an advantage over the FI_{MNF} as its estimated result will only deviate 7.21 points on a scale of 100 from the ground truth value i.e, the actual reported fatigue value. The same conclusion has been reinforced by looking at the coefficient of determination R^2 metric, the (FI_{MDF} model has a close to perfect match to the user's reported fatigue with R^2=91.01 whereas FI_{MNF} model has a R^2=68.86.

$$\frac{(rating-1)}{(number\ of\ response\ categories\ -1)}\ *\ 100 \tag{6.9}$$

$$MAD = \frac{\sum_{i=1}^{n} |Subjective(FI)_i - sEMG(FI)_i|}{n} \tag{6.10}$$

$$MSE = \frac{\sum_{i=1}^{n} (Subjective(FI)_i - sEMG(FI)_i)^2}{n} \tag{6.11}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Subjective(FI)_i - sEMG(FI)_i}{Subjective(FI)_i}\right| \tag{6.12}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Subjective(FI)_i - sEMG(FI)_i)^2}{n}} \tag{6.13}$$

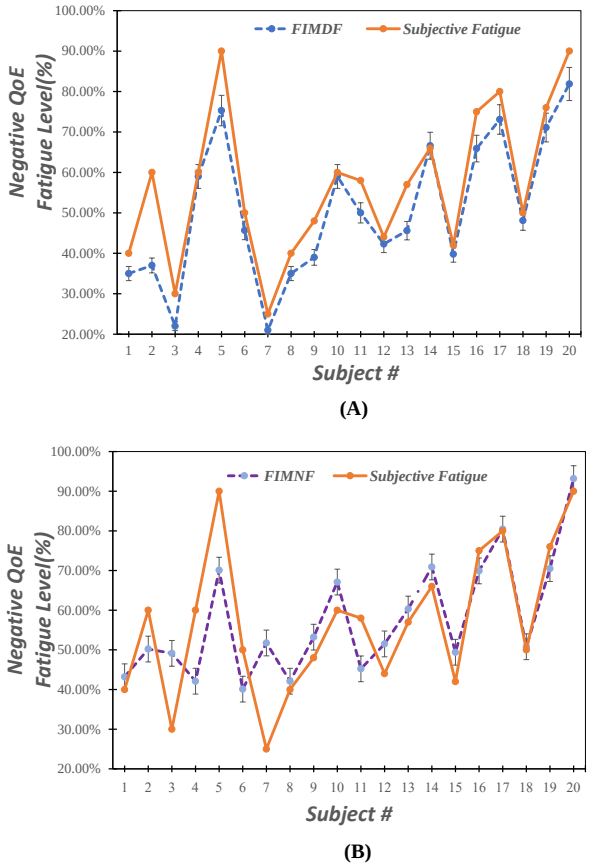


Figure 6.8: Overall Fatigue rating scores by subjects and their FI_{MNF} and FI_{MNF} estimator models.

Figure 6.8 depicts the arm's fatigue level based on our MNF, MDF and subjective analysis. The vertical axis represents the arm's of fatigue percentage whereas the horizontal axis shows the subject index. We also conducted a statistical analysis to calculate the correlation between overall FI_{MNF}, FI_{MDF} and the subjects' self-evaluated results. The t-Test (Paired Two Sample for Means test), was used to determine if the calculated fatigues from MDF and MNF are a significant relationship with the correspondent self-reported one or not. To do so and at the 95% confidence level (i.e., 5% significance level), we propose the following hypotheses (6.14) and (6.15) :

$$H_0 : \mu_1 = \mu_2 \tag{6.14}$$

(denoting the null hypothesis which represents that the paired population means are equal. In another word, MDF and MNF are suitable estimators and predictors of the arm fatigue level associated with the haptic interaction)

$$H_1 : \mu_1 \neq \mu_2 \tag{6.15}$$

(denoting that the paired population means are not equal hence MDF and MNF are not useful predictors of the arm fatigue level associated with the haptic interaction) To verify which hypothesis is right, we should calculate the t and p values. The t score is can be calculated using formula 6.16:

$$t = \frac{M_1 - M_2}{SE_p} \tag{6.16}$$

Where M_1 and M_1 denotes the means of the two paired samples e.g, subjective fatigue mean and FI_{MDF} and SE_p denotes the standard error of the difference between these two samples which can be given as equation6.17:

$$SE_p = \sqrt{(SE_{M1})^2 + (SE_{M2})^2 - 2r(SE_{M2})(SE_{M2})} \tag{6.17}$$

$$SE_M = \frac{STDEV(M)}{\sqrt{Count(M)}} \tag{6.18}$$

Where $STDEV(M)$ is the standard deviation of the sample M, $Count(M)$ is the number of items per sample M, SE_{M1} and SE_{M2} represent the standard error of the means of the self-reported fatigue experience and $sEMG$ MDF or MNF fatigue indicator samples, and r is the correspondent Pearson correlation coefficient for each fatigue index. Using table 6.1 The Pearson correlation coefficient for the MNF fatigue index is equal to 0.74. On the other hand, the MDF fatigue index outperformed MNF with a Pearson correlation coefficient = 0.908. Therefore for MNF and MDF sEMG fatigue estimator models, we can compute the values of t using equation 6.16 as 1.729 and 1.685 respectively.

With the obtained t values and the degree of freedom (df = 19, calculated from the number of users-1), we deployed the t distribution table to obtain the two-tailed

Table 6.3: Fatigue Level Questionnaire

Q#	N/A 1 2 3 4 5 6 COMPLETELY 7
1	To what extent, if any, did using the haptic device cause fatigue?
2	Please indicate your current arm fatigue level?
3	To what extent do you think the haptic playback was realistic?
4	To what extent do you think the haptic feedback was useful?
5	To what extent did using the haptic device gives a better experience?
6	Please indicate your current frustration and boredom level?

$p - values$ for FI_{MNF} and FI_{MDF} 0.92 and 0.34 respectively. Since both ($p - values > \alpha = 0.05$ i.e., the significance level, then the null hypothesis H_0 holds. Consequently, there is a significant positive relationship between the objective calculated $sEMG$ fatigue indicators and the ground truth subjective questionnaire fatigue results with advantages to the median frequency over the mean frequency.

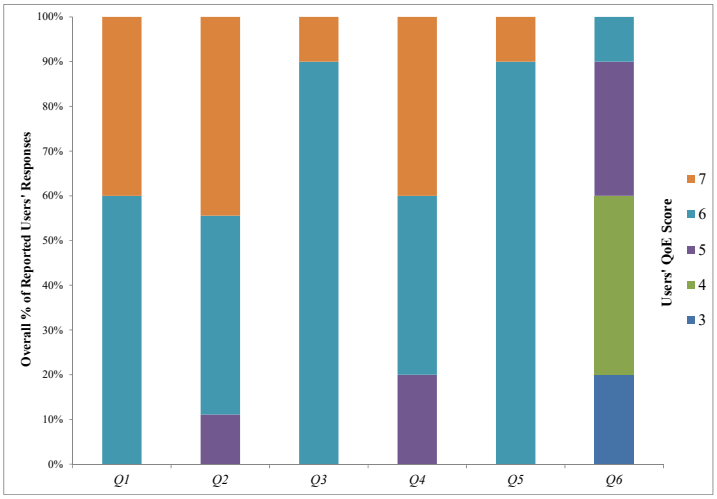


Figure 6.9: Questionnaire Results

From Figure 6.9, we can also notice that almost all the subjects indicated that they felt fatigued from using the haptic phantom device. 40% of the subjects indicated high fatigue whereas 60% of the responses are in the moderate range. In addition, and from a user's experience perspective, all the subjects found the haptic playback application more realistic and useful when compared to a traditional drawing system using Microsoft Styles. There are diverse feedbacks from the subjects regarding their boredom and frustration levels which perhaps are related to the diversity of the subject's haptic experience and background.

6.7 Summary

In this chapter, we propose and deploy a framework to dynamically and objectively detecting, tracing, and quantifying the induced arm fatigue experience among subjects who repetitively interact with a haptic-enabled Tactile Internet application. For that goal, we used the trends of the MDF and MNF models generated from the electromyography power spectrum as adequate estimators to quantify fatigue. Both MDF and MNF can be taken into consideration when developing a new haptic application to steer the muscles' movement in a contextual way that improves the QoE of the application as Our outcomes reveal a significant relationship between users' perceived fatigue experience and sEMG analysis. The results from our experiments reinforce our expectations as MDF outperformed MNF when both compared to the haptic users' perception of fatigue.

Chapter 7

Conclusion and Future Work

Undoubtedly, the world is presently facing a pandemic that is reshaping daily lives and business activities. With the special focus on maintaining physical distancing norms for curbing the expeditious spread of the disease, many institutions, individuals, and industries rely on communications networks or telecoms for ensuring service consistency so as to avoid complete termination of their business operations and other activities. As a result, there is a huge demand to move traditional health and manufacturing actions to what are referred to as Health 4.0 and Industry 4.0 over the Digital twin architecture. Thanks to the Tactile Internet where new modalities such as haptic are appended to traditional multimedia communication. This is one of the main technologies used in the development of future network infrastructures such as the fifth generation of mobile networks (5G). This certainly paved the road for the emergence of new technologies in particular the digital twin system and Internet of Skills.

Given that multimedia is no longer limited to audiovisual application, as the realization of the digital twin demands the spatial, temporal, and contextual combination of new HD modalities e.g., olfaction, haptic and gustatory, we proposed a new quality-of-experience (QoE) taxonomy for evaluating digital twin systems, through referencing objective, implicit biomedical, and subjective parameters, such as QoS, psychophysiological metrics, user's perception, and usability factors. The taxonomy was evaluated

through the usage of machine learning algorithms to automatically predict the QoE for the said system at all times hence reducing the need for the time-consuming user subjective feedbacks. Thus, the developers, manufacturer, and network providers of the DT systems can enhance their products, devices, and services. In this thesis, we also review the most in-depth listing of the technical challenges required to realize the Digital twin technologies. We show that the QoS requirements differ based on the Digital twin modality type and nature. We also designed a QoE-QoS evaluation studies to optimize the dissemination of new media, haptic and VR in particular, on a large scale environment using up-to-date and modren networking technologies such as SDN and mobile edge computing. In addition, We comment on the up-to-date studies related to QoE frameworks and measurements in the area of new immersive and interactive multimodal applications.

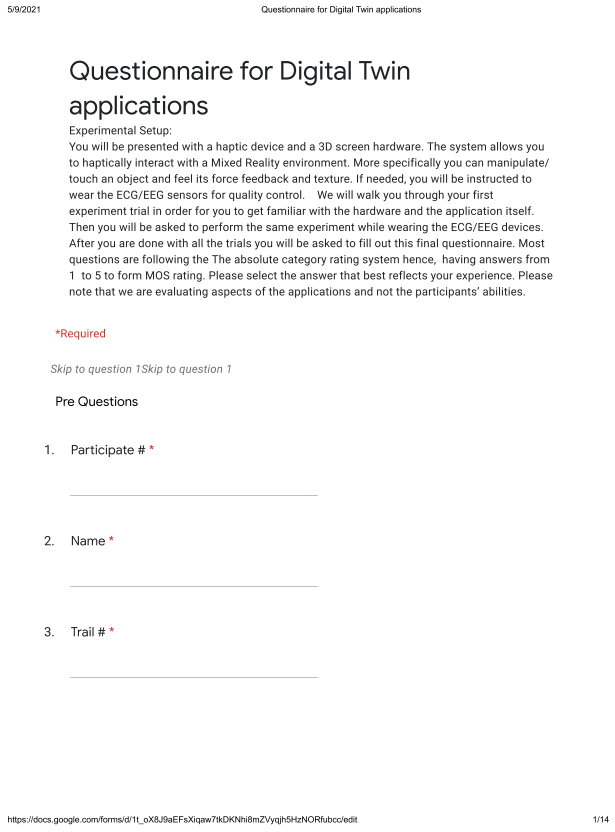
As the DT system implies that remote haptic will be embedded in our daily media interaction and consumption, studying the way users perceive fatigue is a proactive step in enhancing the overall QoE of the said system. As such, we proposed a framework to dynamically and objectively deduce and predict the arm's fatigue experience among subjects who repetitively interact with a haptic-enabled interface. The outcomes of our research confirmed that our MDF and MNF fatigue estimator can be taken into consideration when developing Industry 4.0 and Health 4.0 systems in a seamless way i.e., without the involvement of the user, to steer the muscles' movement in a contextual way that improves the QoE of the application.

QoE for DT is a relatively novel area with new contributions emerging persistently. We have argued that the user behaviour and perception are not linear thus can be described as huge, fuzzy, and complex. We attempt to objectify underlying perceptual and cognitive behaviour by using implicit bio-metrics, equations, and machine learning algorithms in a traditional and specific situations. Undeniably, there is consistently something that can be done better thus we suggest a room for enhancement. As future work, we plan

to incorporate AR devices, olfactory display and digital taste in our use case as well as considering more bio-metrics such as eye-tracking, lactic acid, and eye blinking rate in our QoE framework. Accordingly and after enlarging our dataset, we plan to employ powerful AI such as deep learning algorithms to build QoE prediction models. This would be done in stages, similar to how we investigated fatigue in this dissertation. The new media data can also be used to standardise and detect other QoE factors that capture any variation in the user's emotional state during and after the experience of the DT system. Furthermore, we plan to benchmark more stabilizing control algorithms for DT-based teleoperation through a real 5G and beyond networks that use millimetre-wave and terahertz communication. Lastly, effective techniques to objectively mitigate fatigue resulting from haptic interfaces are our prospective research.

Chapter 8

Appendix



5/9/2021 Questionnaire for Digital Twin applications

Questionnaire for Digital Twin applications

Experimental Setup:
You will be presented with a haptic device and a 3D screen hardware. The system allows you to haptically interact with a Mixed Reality environment. More specifically you can manipulate/ touch an object and feel its force feedback and texture. If needed, you will be instructed to wear the ECG/EEG sensors for quality control. We will walk you through your first experiment trial in order for you to get familiar with the hardware and the application itself. Then you will be asked to perform the same experiment while wearing the ECG/EEG devices. After you are done with all the trials you will be asked to fill out this final questionnaire. Most questions are following the The absolute category rating system hence, having answers from 1 to 5 to form MOS rating. Please select the answer that best reflects your experience. Please note that we are evaluating aspects of the applications and not the participants' abilities.

*Required

Skip to question 1Skip to question 1

Pre Questions

1. Participate # *
2. Name *
3. Trail # *

https://docs.google.com/forms/d/1t_oX8J9aEFsXiqaw7tkDKNhi8mZVyqjh5HzNORfubcc/edit 1/14

Figure 8.1: Questionnaire for DT QoE Evaluation Page 1.

4. Gender *

Mark only one oval.

- ◯ Male
- ◯ Female
- ◯ Other: _____

5. Your age range *

Mark only one oval.

- ◯ 10-19
- ◯ 20-29
- ◯ 40-49
- ◯ 50+
- ◯ Other: _____

6. Hight *

_____

7. Weight *

_____

8. Left or right arm

Mark only one oval.

- ◯ L
- ◯ R

Figure 8.2: Questionnaire for DT QoE Evaluation Page 2.

9. Please describe your current Physical State *

Mark only one oval.

- ◯ Fatigued
- ◯ Natural
- ◯ Conformable

10. Please describe your current Mental State *

Mark only one oval.

- ◯ Positive
- ◯ Natural
- ◯ Stressed

11. How familiar are you with haptic devices? *

Mark only one oval.

- ◯ I don't know what a haptic device is
- ◯ Familiar with them but never used them
- ◯ Have used them before but only once or twice
- ◯ Use them regularly
- ◯ Have programming and/or engineering knowledge of them

Figure 8.3: Questionnaire for DT QoE Evaluation Page 3.

12. How familiar are you with Nvidia 3D glasses and Mixed reality applications? (Please circle one) *

Mark only one oval.

- I don't know what Nvidia 3D glasses or VR reality are
- Familiar with them but never used them
- Have used them before but only once or twice
- Use them regularly
- Have programming and/or engineering knowledge of them

Skip to question 13

Content Quality

All questions have answers from 1 to 5 where 1 indicates the lowest rating while 5 indicate the highest rating. Please select the answer that best reflects your experience for each question.

13. 1. To what extent do you think haptic provides realistic feedbacks and stimuli? where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.4: Questionnaire for DT QoE Evaluation Page 4.

14. 2. To what extent do you think the haptic feedback was useful and natural? where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

15. 3. To what extent did the application provide realistic visual information? where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

16. 4. Did enhancing the 3D visual resolution increase the realism of the application? where 1= strongly disagree and 5= strongly agree *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.5: Questionnaire for DT QoE Evaluation Page 5.

17. 5. Based on your ratings for Q1-4, please give a general rating for the content quality? where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Skip to question 18

Hardware Quality

All questions have answers from 1 to 5 where 1 indicates the lowest rating while 5 indicate the highest rating. Please select the answer that best reflects your experience for each question.

18. 6. To what extent did using the haptic device make for a better experience than using the stylus? (compare it with using the styles) where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.6: Questionnaire for DT QoE Evaluation Page 6.

19. 7. To what extent did using the haptic device make for a better experience than using the mouse? (compare it to the mouse with MS paint) where 1 indicates (Not at all) and 5 indicates completely. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

20. 8. How would you rank your experience using the haptic device? where 1 indicates (very poor) and 5 indicates Excellent. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

21. 9. How would you rank your experience using Nvidia 3D glasses? where 1 indicates (very poor) and 5 indicates Excellent. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Figure 8.7: Questionnaire for DT QoE Evaluation Page 7.

22. 10. To what extent do you feel comfortable wearing the Nvidia 3D/ VR HMD/ ECG/EEG devices? where 1 indicates (Not comfortable) and 5 indicates Very comfortably. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

23. 11. Please indicate if you are feeling cybersickness e.g sweating, nausea, drowsiness? *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

24. 12) To what extent do you feel the testing environment was safe? where 1 indicates (Not Hazardous) and 5 indicates safe. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Figure 8.8: Questionnaire for DT QoE Evaluation Page 8.

25. 13. Based on your ratings for Q6-12 please give a general rating for the hardware quality? where 1 indicates (very poor) and 5 indicates Excellent. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Skip to question 26

Network Quality

All questions have answers from 1 to 5 where 1 indicates the lowest rating while 5 indicate the highest rating. Please select the answer that best reflects your experience for each question.

26. 14. How would you rank the synchronization between graphics and haptic feedback? where 1 indicates (very poor) and 5 indicates Excellent. *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.9: Questionnaire for DT QoE Evaluation Page 8.

27. 15. To what extent do you think the system responds precisely (jitter) to your hand movement? 1= very poor and 5=Completely precise *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

28. 16. To what extent do you think the system responds quickly (latency) to your hand movement? 1= very poor and 5=very swift *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

29. 17. To what extent did the collaboration increased your experience ?

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.10: Questionnaire for DT QoE Evaluation Page 10.

30. 18. Based on your ratings for Q14-17, please give a general rating for the network quality? *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Skip to question 31

User Experience

All questions have answers from 1 to 5 where 1 indicates the lowest rating while 5 indicate the highest rating. Please select the answer that best reflects your experience for each question.

31. 19) To what extent, if any, did use the haptic device and VR increase interaction and presence ? *

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Figure 8.11: Questionnaire for DT QoE Evaluation Page 11.

32. 20. To what extent, if any, did use haptic device cause arm fatigue? *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

33. 21. To what extent, if any, did use 3D glasses cause eye fatigue? *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

34. 22. To what extent did the meaning of immersed content fit the real environment? *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Figure 8.12: Questionnaire for DT QoE Evaluation Page 12.

35. 23. Please rank your user experience and degree of perception

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Skip to question 36

Usability

All questions have answers from 1 to 5 where 1 indicates the lowest rating while 5 indicate the highest rating. Please select the answer that best reflects your experience for each question.

36. 24. To what extent you found the system is easy to use. *

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

37. 25. To what extended did you find the instruction on the screen was useful?

Mark only one oval.

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5

Figure 8.13: Questionnaire for DT QoE Evaluation Page 13.

38. 26. What is your Usability score ?

Mark only one oval.

- 1
- 2
- 3
- 4
- 5

Skip to question 39

Final QoE rating

39. Any constrictive comments about your experience:

40. Please give a grade, over 100, for the overall quality of this haptic VR experience: ___/100 *

Figure 8.14: Questionnaire for DT QoE Evaluation Page 14.

References

[1] Abdulmotaleb El Saddik. Digital twins: The convergence of multimedia technologies. *IEEE MultiMedia*, 25(2):87–92, 2018.

[2] Eckehard Steinbach, Matti Strese, Mohamad Eid, Xun Liu, Amit Bhardwaj, Qian Liu, Mohammad Al-Ja'afreh, Toktam Mahmoodi, Rania Hassen, Abdulmotaleb El Saddik, et al. Haptic codecs for the tactile internet. *Proceedings of the IEEE*, 107(2):447–470, 2018.

[3] Ryo Arima, Mya Sithu, Yutaka Ishibashi, et al. Qoe assessment of fairness between players in networked virtual 3d objects identification game using haptic, olfactory, and auditory senses. *International Journal of Communications, Network and System Sciences*, 10(07):129, 2017.

[4] Takuji Narumi, Takashi Kajinami, Shinya Nishizaka, Tomohiro Tanikawa, and Michitaka Hirose. Pseudo-gustatory display system based on cross-modal integration of vision, olfaction and gustation. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 127–130. IEEE, 2011.

[5] Rania Hassen and Eckehard Steinbach. Hssim: An objective haptic quality assessment measure for force-feedback signals. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, 2018.

[6] 3D systems. Touch haptic device, 2021. Online accessed; Thursday, May 6, 2021.

[7] Mohamad Hoda, Basim Hafidh, and Abdulmotaleb El Saddik. Haptic glove for finger rehabilitation. In *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, pages 1–6. IEEE, 2015.

[8] Oliver Holland, Eckehard Steinbach, R Venkatesha Prasad, Qian Liu, Zaher Dawy, Adnan Aijaz, Nikolaos Pappas, Kishor Chandra, Vijay S Rao, Sharief Oteafy, et al. The ieee 1918.1 "tactile internet" standards working group and its standards. *Proceedings of the IEEE*, 107(2):256–279, 2019.

[9] ITU-T Study Group 13. Recommendation itu-t y.2060, overview of the internet of things., 2012.

[10] Siddhartha Kumar Khaitan and James D McCalley. Design techniques and applications of cyberphysical systems: A survey. *IEEE Systems Journal*, 9(2):350–365, 2014.

[11] Gerhard P Fettweis. The tactile internet: applications and challenges. *IEEE Vehicular Technology Magazine*, 9(1):64–70, 2014.

[12] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Enric Pujol, Ingmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapiador, Narseo Vallina-Rodriguez, et al. The lockdown effect: Implications of the covid-19 pandemic on internet traffic. In *Proceedings of the ACM Internet Measurement Conference*, pages 1–18, 2020.

[13] Internatinal Telecommunication Union ITU-T. The tactile internet. *ITU-T Technology Watch Report*, 2014.

[14] Markus Rank, Zhuanghua Shi, Hermann J. Müller, and Sandra Hirche. Perception of delay in haptic telepresence systems. *Presence: teleoperators and virtual environments*, 19(5):389–399, 2010.

[15] Barry W Ache and Janet M Young. Olfaction: diverse species, conserved principles. *Neuron*, 48(3):417–430, 2005.

[16] Abdulmotaleb El Saddik, Mauricio Orozco, Mohamad Eid, and Jongeun Cha. *Haptics Technologies: Bringing Touch to Multimedia*. Springer Series on Touch and Haptic Systems. Springer Berlin Heidelberg, 2011.

[17] Zhenhui Yuan, Gheorghita Ghinea, and Gabriel-Miro Muntean. Quality of experience study for multiple sensorial media delivery. In *Wireless Communications and Mobile Computing Conference (IWCMC), 2014 International*, pages 1142–1146. IEEE, 2014.

[18] Zhenhui Yuan, Shengyang Chen, Gheorghita Ghinea, and Gabriel-Miro Muntean. User quality of experience of mulsemedia applications. *ACM Trans. Multimedia Comput. Commun. Appl.*, 11(1s):15:1–15:19, October 2014.

[19] Mohamad Eid and Abdulmotaleb El Saddik. Admux communication protrocol for real-time multimodal intreaction. In *Proceedings of the 2012 IEEE/ACM 16th International Symposium on Distributed Simulation and Real Time Applications*, DS-RT '12, pages 118–123, Washington, DC, USA, 2012. IEEE Computer Society.

[20] Gabriel Robles-De-La-Torre. The importance of the sense of touch in virtual and real environments. *IEEE Multimedia*, 13(3):24–30, 2006.

[21] Kyle B Reed, Michael Peshkin, Mitra J Hartmann, James Patton, Peter M Vishton, and Marcia Grabowecky. Haptic cooperation between people, and between people and machines. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 2109–2114. IEEE, 2006.

[22] Eckehard Steinbach, Sandra Hirche, Marc Ernst, Fernanda Brandi, Rahul Chaudhari, Julius Kammerl, and Iason Vittorias. Haptic communications. *Proceedings of the IEEE*, 100(4):937–956, 2012.

[23] Blake Hannaford and Jee-Hwan Ryu. Time-domain passivity control of haptic interfaces. *Robotics and Automation, IEEE Transactions on*, 18(1):1–10, 2002.

[24] Kenneth Salisbury, Francois Conti, and Federico Barbagli. Haptic rendering: introductory concepts. *Computer Graphics and Applications, IEEE*, 24(2):24–32, 2004.

[25] GNeomagic Touch. Geomagic touch® haptic device, 2020. Online accessed; Sunday, Dec 30, 2020.

[26] Joseph "Jofish" Kaye. Making scents: Aromatic output for hci. *interactions*, 11(1):48–61, January 2004.

[27] K. Toko. *Biochemical Sensors: Mimicking Gustatory and Olfactory Senses*. Biochemical Sensors: Mimicking Gustatory and Olfactory Senses. Pan Stanford, 2013.

[28] J. Gardner and P.N. Bartlett. *Sensors and Sensory Systems for an Electronic Nose*. Nato Science Series E:. Springer Netherlands, 2013.

[29] Eckhard W. Heymann. The neglected sense-olfaction in primate behavior, ecology, and evolution. *American Journal of Primatology*, 68(6):519–524, 2006.

[30] Gheorghita Ghinea and Oluwakemi A Ademoye. Olfaction-enhanced multimedia: perspectives and challenges. *Multimedia Tools and Applications*, 55(3):601–626, 2011.

[31] Sosuke Hoshino, Yutaka Ishibashi, Norishige Fukushima, and Shinji Sugawara. Qoe assessment in olfactory and haptic media transmission: Influence of inter-stream synchronization error. In *Communications Quality and Reliability (CQR), 2011 IEEE International Workshop Technical Committee on*, pages 1–6. IEEE, 2011.

[32] M. Sithu, Y. Ishibashi, Pingguo Huang, and N. Fukushima. Ikebana competition in networked virtual environment with haptic and olfactory senses. pages 1–3, Dec 2014.

[33] 5492 ISO. International standard 5492. sensory analysis vocabulary. ref. no. iso 5492:2008 (e). Technical report, 2008.

[34] Marianna Obrist, Rob Comber, Sriram Subramanian, Betina Piqueras-Fiszman, Carlos Velasco, and Charles Spence. Temporal, affective, and embodied characteristics of taste experiences: A framework for design. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2853–2862, New York, NY, USA, 2014. ACM.

[35] Jae-young Lee, Shahram Payandeh, and Ljiljana Trajkovic. Performance evaluation of transport protocols for internet-based teleoperation systems. *Proceedings of OPNETWORK. Washington DC: OPNET Technologies Inc*, pages 1–6, 2010.

[36] Rafael Asorey Cacheda, Daniel Castro Garcia, Antonio Cuevas, Francisco Javier Gonzalez Castano, Javier Herrero Sanchez, Georgios Koltsidas, Vincenzo Mancuso, Jose Ignacio Moreno Novella, Seounghoon Oh, and Antonio Panto. Qos requirements for multimedia services. In *Resource Management in Satellite Networks*, pages 67–94. Springer, 2007.

[37] Boris Grot, Stephen W. Keckler, and Onur Mutlu. Preemptive virtual clock: A flexible, efficient, and cost-effective qos scheme for networks-on-chip. In *Proceedings of the 42Nd Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO 42, pages 268–279, New York, NY, USA, 2009. ACM.

[38] Duminda Wijesekera and Jaideep Srivastava. Quality of service (qos) metrics for continuous media. *Multimedia Tools and Applications*, 3(2):127–166, 1996.

[39] Nalini Venkatasubramanian and Klara Nahrstedt. An integrated metric for video qos. In *Proceedings of the Fifth ACM International Conference on Multimedia*, MULTIMEDIA '97, pages 371–380, New York, NY, USA, 1997. ACM.

[40] Bikash Sabata, Saptarshi Chatterjee, Michael Davis, Jaroslaw J Sydir, and Thomas F Lawrence. Taxonomy for qos specifications. In *Object-Oriented Real-Time Dependable Systems, 1997. Proceedings., Third International Workshop on*, pages 100–107. IEEE, 1997.

[41] Yan Chen, Toni Farley, and Nong Ye. Qos requirements of network applications on the internet. *Information, Knowledge, Systems Management*, 4:55–76, 2004.

[42] G Kokkonis, K Psannis, M Roumeliotis, S Kontogiannis, and Y Ishibashi. Evaluating transport and application layer protocols for haptic applications. In *Proc of. IEEE International Symposium on Haptic Audio-Visual Environments and Games*, pages 66–71, 2012.

[43] Eckehard Steinbach, Sandra Hirche, Julius Kammerl, Iason Vittorias, and Rahul Chaudhari. Haptic data compression and communication. *IEEE Signal Processing Magazine*, 28(1):87–96, 2011.

[44] Kyoung Shin Park and Robert V Kenyon. Effects of network characteristics on human performance in a collaborative virtual environment. In *Virtual Reality, 1999. Proceedings., IEEE*, pages 104–111. IEEE, 1999.

[45] Peter Hinterseer, E Steinbach, and Subhasis Chaudhuri. Perception-based compression of haptic data streams using kalman filters. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–V. IEEE, 2006.

[46] N Sakr, J Zhou, ND Georganas, J Zhao, and X Shen. Prediction-based haptic data reduction and compression in tele-mentoring systems. In *Instrumentation and Measurement Technology Conference Proceedings, 2008. IMTC 2008. IEEE*, pages 1828–1832. IEEE, 2008.

[47] Alan Marshall, Kian Meng Yap, and Wai Yu. Providing qos for networked peers in distributed haptic virtual environments. *Advances in Multimedia*, 2008, 2008.

[48] Abdelwahab Hamam, Mohamad Eid, Abdulmotaleb El Saddik, and Nicolas D Georganas. A quality of experience model for haptic user interfaces. In *Proceedings of the 2008 Ambi-Sys workshop on Haptic user interfaces in ambient media systems*, page 1. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008.

[49] Edgar N Gilbert. Capacity of a burst-noise channel. *Bell system technical journal*, 39(5):1253–1265, 1960.

[50] Z.N. Li, M.S. Drew, and J. Liu. *Fundamentals of Multimedia*. Texts in Computer Science. Springer International Publishing, 2014.

[51] K. Sayood. *Introduction to Data Compression*. The Morgan Kaufmann Series in Multimedia Information and Systems. Elsevier Science, 2012.

[52] Wenwu Zhu, Chong Luo, Jianfeng Wang, and Shipeng Li. Multimedia cloud computing. *Signal Processing Magazine, IEEE*, 28(3):59–69, 2011.

[53] Yin Li, Jian Sun, and Heung-Yeung Shum. Video object cut and paste. *ACM Trans. Graph.*, 24(3):595–600, July 2005.

[54] Touradj Ebrahimi. Quality of multimedia experience: Past, present and future. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM '09, pages 3–4, New York, NY, USA, 2009. ACM.

[55] H King, Blake Hannaford, Julius Kammerl, and Eckehard Steinbach. Establishing multimodal telepresence sessions using the session initiation protocol (sip) and advanced haptic codecs. In *Haptics Symposium, 2010 IEEE*, pages 321–325. IEEE, 2010.

[56] Robert Galambos, Scott Makeig, and Peter J Talmachoff. A 40-hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences*, 78(4):2643–2647, 1981.

[57] Dimitrios Miras, Amela Sadagic, Ben Teitelbaum, Jason Leigh, Magda El Zarki, and Haining Liu. A survey on network qos needs of advanced internet applications. *Internet2 QoS Working Group*, 2002.

[58] Ronald T Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.

[59] Shun-Yun Hu, Jehn-Ruey Jiang, and Bing-Yu Chen. Peer-to-peer 3d streaming. *Internet Computing, IEEE*, 14(2):54–61, 2010.

[60] Mohammad Aljaafreh. An efficient hybrid objects selection protocol for 3d streaming over mobile devices. Master's thesis, University of Ottawa, Canada, 2012.

[61] Michael Englert, Yvonne Jung, Marcel Klomann, Jonas Etzold, Paul Grimmy, and Jinyuan Jia. A streaming framework for instant 3d rendering and interaction. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, VRST '15, pages 192–192, New York, NY, USA, 2015. ACM.

[62] Denis Gracanin, Yunxian Zhou, Luiz DaSilva, et al. Quality of service for networked virtual environments. *Communications Magazine, IEEE*, 42(4):42–48, 2004.

[63] Paul E Keller, Richard T Kouzes, Lars J Kangas, and Sherif Hashem. Transmission of olfactory information for tele-medicine. *Interactive Technology and the New Paradigm for Healthcare*, 18:168–172, 1995.

[64] Fabrizio Davide, Martin Holmberg, and Ingemar Lundstrom. 12 virtual olfactory interfaces: electronic noses and olfactory displays. *Communications Through Virtual Technology: Identity Community and Technology in the Internet Age (Pages. 193-220)*, 2001.

[65] Jozef Babiarz, Kwok Chan, and Fred Baker. Configuration guidelines for diffserv service classes. Technical report, 2006.

[66] Kian Meng Yap, Alan Marshall, Wai Yu, G Dodds, Qiang Gu, and Rima T'faily Souayed. Characterising distributed haptic virtual environment network traffic flows, 2007.

[67] Diego Kreutz, Fernando MV Ramos, Paulo Esteves Verissimo, Christian Esteve Rothenberg, Siamak Azodolmolky, and Steve Uhlig. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1):14–76, 2014.

[68] Samaresh Bera, Sudip Misra, and Niloy Saha. Traffic-aware dynamic controller assignment in sdn. *IEEE Transactions on Communications*, 68(7):4375–4382, 2020.

[69] John William Evans and Clarence Filsfils. *Deploying IP and MPLS QoS for multiservice networks: theory and practice*. Elsevier, 2010.

[70] Mohamad Eid and Abdulmotaleb El Saddik. Admux communication protrocol for real-time multimodal intreaction. In *Proceedings of the 2012 IEEE/ACM 16th International Symposium on Distributed Simulation and Real Time Applications*, pages 118–123. IEEE Computer Society, 2012.

[71] Zhenhui Yuan, Gheorghita Ghinea, and Gabriel-Miro Muntean. Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery. *Multimedia, IEEE Transactions on*, 17(1):104–117, 2015.

[72] Burak Cizmeci, Xiao Xu, Rahul Chaudhari, Christoph Bachhuber, Nicolas Alt, and Eckehard Steinbach. A multiplexing scheme for multimodal teleoperation. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 13(2):21, 2017.

[73] Kalevi Kilkki. Quality of experience in communications ecosystem. *J. UCS*, 14(5):615–624, 2008.

[74] Fred D Davis, Richard P Bagozzi, and Paul R Warshaw. User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8):982–1003, 1989.

[75] G. Calvert, C. Spence, and B.E. Stein. *The Handbook of Multisensory Processes*. A Bradford book. MIT Press, 2004.

[76] Sebastian Möller and Alexander Raake. Quality of experience. *New York, US: Springer*, 2014.

[77] Khalil Ur Rehman Laghari and Kay Connelly. Toward total quality of experience: A qoe model in a communication ecosystem. *Communications Magazine, IEEE*, 50(4):58–65, 2012.

[78] PL Callet, S Möller, and A Perkis. Qualinet white paper on definitions of quality of experience (2012). *Eur. Network Qual. Exp. Multimedia Syst. and Serv.(COST Action IC 1003)*, 2013.

[79] Cristina Hava Muntean, Gabriel-Miro Muntean, Jennifer McManis, and Alexandra I Cristea. Quality of experience-laos: create once, use many, use anywhere. *International Journal of Learning Technology*, 3(3):209–229, 2007.

[80] K Kumar. A marriage made in qoe heaven. *CED Magazine*, pages 37–39, 2005.

[81] Telephone Transmission Quality. Methods for objective and subjective assessment of quality. *ITU-T Recommendation*, page 830, 1996.

[82] R ITU. Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems. *ITU-R Recommendation BS*, 1116, 1994.

[83] B Series. Method for the subjective assessment of intermediate quality level of audio systems. *International Telecommunication Union Radiocommunication Assembly*, 2014.

[84] Telephone Installations and Local Line. Itu-tp. 910. *Subjective video quality assessment methods for multimedia applications, Recommendation ITU-T*, page 910.

[85] IT Union. Methods for the subjective assessment of video quality audio quality and audiovisual quality of internet video and distribution quality television in any environment. *SERIES P: TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS*, 2016.

[86] Sana Aroussi and Abdelhamid Mellouk. Survey on machine learning-based qoe-qos correlation models. In *2014 International Conference on Computing, Management and Telecommunications (ComManTel)*, pages 200–204. IEEE, 2014.

[87] Sajad Khorsandroo, Rafidah Md Noor, and Sayid Khorsandroo. The role of psychophysics laws in quality of experience assessment: a video streaming case study. In *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, pages 446–452, 2012.

[88] Sajad Khorsandroo, Rafidah Md Noor, and Sayid Khorsandroo. A generic quantitative relationship to assess interdependency of qoe and qos. *KSII Transactions on Internet and Information Systems (TIIS)*, 7(2):327–346, 2013.

[89] Peter Reichl, Bruno Tuffin, and Raimund Schatz. Economics of logarithmic quality-of-experience in communication networks. In *2010 9th Conference of Telecommunication, Media and Internet*, pages 1–8. IEEE, 2010.

[90] Peter Reichl, Sebastian Egger, Raimund Schatz, and Alessandro D'Alconzo. The logarithmic nature of qoe and the role of the weber-fechner law in qoe assessment. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.

[91] Markus Fiedler, Tobias Hossfeld, and Phuoc Tran-Gia. A generic quantitative relationship between quality of experience and quality of service. *Network, IEEE*, 24(2):36–41, 2010.

[92] Junaid Shaikh, Markus Fiedler, and Denis Collange. Quality of experience from user and network perspectives. *annals of telecommunications-annales des telecommunications*, 65(1):47–57, 2010.

[93] Stanley Smith Stevens. *Psychophysics: Introduction to its perceptual, neural and social prospects*. Routledge, 2017.

[94] Stanislas Dehaene. The neural basis of the weber–fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, 7(4):145–147, 2003.

[95] Sally Floyd and Vern Paxson. Difficulties in simulating the internet. *IEEE/ACM Trans. Netw.*, 9(4):392–403, August 2001.

[96] Lieven De Marez and Katrien De Moor. The challenge of user-and qoe-centric research and product development in today's ict-environment. *Observatorio (OBS*)*, 1(3), 2007.

[97] Muslim Elkotob, Daniel Grandlund, Karl Andersson, and Christer Ahlund. Multimedia qoe optimized management using prediction and statistical learning. In *IEEE Local Computer Network Conference*, pages 324–327. Ieee, 2010.

[98] Sana Aroussi, Thouraya Bouabana-Tebibel, and Abdelhamid Mellouk. Empirical qoe/qos correlation model based on multiple parameters for vod flows. In *2012 IEEE Global Communications Conference (GLOBECOM)*, pages 1963–1968. IEEE, 2012.

[99] N Sakr, ND Georganas, and J Zhao. A perceptual quality metric for haptic signals. In *2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games*, pages 27–32. IEEE, 2007.

[100] Rahul Chaudhari, Eckehard Steinbach, and Sandra Hirche. Towards an objective quality evaluation framework for haptic data reduction. In *2011 IEEE World Haptics Conference*, pages 539–544. IEEE, 2011.

[101] Stanley S Stevens. On the psychophysical law. *Psychological review*, 64(3):153, 1957.

[102] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[103] Abdelwahab Hamam, El Saddik Abdulmotaleb, and Jihad Alja'am. A quality of experience model for haptic virtual environments. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10(3):28:1–28:23, April 2014.

[104] Gheorghita Ghinea, Christian Timmerer, Weisi Lin, and Stephen R. Gulliver. Mulsemedia: State of the art, perspectives, and challenges. *ACM Trans. Multimedia Comput. Commun. Appl.*, 11(1s):17:1–17:23, October 2014.

[105] Nataša Banović-Ćurguz and Dijana Ilišević. Mapping of qos/qoe in 5g networks. In *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pages 404–408. IEEE, 2019.

[106] Thomas E Whalen, Sylvie Noël, and John Stewart. Measuring the human side of virtual reality. In *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2003. VECIMS'03. 2003*, pages 8–12. IEEE, 2003.

[107] Darragh Egan, Sean Brennan, John Barrett, Yuansong Qiao, Christian Timmerer, and Niall Murray. An evaluation of heart rate and electrodermal activity as an objective qoe evaluation method for immersive virtual reality environments. In *2016*

Eighth International Conference on Quality of Multimedia Experience (QoMEX), pages 1–6. IEEE, 2016.

[108] Conor Keighrey, Ronan Flynn, Siobhan Murray, Sean Brennan, and Niall Murray. Comparing user qoe via physiological and interaction measurements of immersive ar and vr speech and language therapy applications. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, pages 485–492, 2017.

[109] *Evaluating the Impact of Network Delay on user Quality of Experience of an Interactive Virtual Reality Industry 4.0 Application, author=David Concannon, year=2020, publisher=Athlone Institute of Technology.*

[110] Haneen Alsuradi, Chaitali Pawar, Wanjoo Park, and Mohamad Eid. Detection of tactile feedback on touch-screen devices using eeg data. In *2020 IEEE Haptics Symposium (HAPTICS)*, pages 775–780. IEEE, 2020.

[111] Safaa Eldeeb, Jordyn Ting, Deniz Erdogmus, Douglas Weber, and Murat Akcakaya. Eeg-based texture classification during active touch. In *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE, 2019.

[112] He Xu, Fernando Pereira, Christian Timmerer, and Touradj Ebrahimi. Towards quality of sensory experience in multimedia. In Nicolas Demassieux and Mario Campolargo, editors, *Proceedings of 2015 European Conference on Networks and Communications (EUCNC)*, pages 627–628, Brussels, Belgium, June 2015. IEEE.

[113] Gebremariam Mesfin, Nadia Hussain, Elahe Kani-Zabihi, Alexandra Covaci, Estêvão B Saleme, and Gheorghita Ghinea. Qoe of cross-modally mapped mulsemedia: an assessment using eye gaze and heart rate. *Multimedia Tools and Applications*, 79(11):7987–8009, 2020.

[114] Estêvão Bissoli Saleme, Celso AS Santos, and Gheorghita Ghinea. A mulsemedia framework for delivering sensory effects to heterogeneous systems. *Multimedia Systems*, 25(4):421–447, 2019.

[115] Elahe Kani-Zabihi, Nadia Hussain, Gebremariam Mesfin, Alexandra Covaci, and Gheorghita Ghinea. On the influence of individual differences in cross-modal mulsemedia qoe. *Multimedia Tools and Applications*, 80(2):2377–2394, 2021.

[116] Michele A Saad, Alan C Bovik, and Christophe Charrier. Blind image quality assessment: A natural scene statistics approach in the dct domain. *IEEE transactions on Image Processing*, 21(8):3339–3352, 2012.

[117] Anush Krishna Moorthy and Alan Conrad Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE transactions on Image Processing*, 20(12):3350–3364, 2011.

[118] Zahid Akhtar, Kamran Siddique, Ajita Rattani, Syaheerah Lebai Lutfi, and Tiago H Falk. Why is multimedia quality of experience assessment a challenging problem? *IEEE Access*, 7:117897–117915, 2019.

[119] Longyu Zhang, Haiwei Dong, and Abdulmotaleb El Saddik. Towards a qoe model to evaluate holographic augmented reality devices. *IEEE MultiMedia*, 26(2):21–32, 2018.

[120] Sebastian Moller, Klaus-Peter Engelbrecht, Christine Kuhnel, Ina Wechsung, and Benjamin Weiss. A taxonomy of quality of service and quality of experience of multimodal human-machine interaction. In *2009 International Workshop on Quality of Multimedia Experience*, pages 7–12. IEEE, 2009.

[121] Abdelwahab Hamam, Abdulmotaleb El Saddik, and Jihad Alja'Am. A quality of experience model for haptic virtual environments. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 10(3):28, 2014.

[122] Rania Hassen and Eckehard Steinbach. Vibrotactile signal compression based on sparse linear prediction and human tactile sensitivity function. In *2019 IEEE World Haptics Conference (WHC)*, pages 301–306. IEEE, 2019.

[123] Wanmin Wu, Ahsan Arefin, Raoul Rivas, Klara Nahrstedt, Renata Sheppard, and Zhenyu Yang. Quality of experience in distributed interactive multimedia environments: Toward a theoretical framework. In *Proceedings of the 17th ACM International Conference on Multimedia*, MM '09, pages 481–490, New York, NY, USA, 2009. ACM.

[124] Kjell Brunnström, Sergio Ariel Beker, Katrien De Moor, Ann Dooms, Sebastian Egger, Marie-Neige Garcia, Tobias Hossfeld, Satu Jumisko-Pyykkö, Christian Keimel, Mohamed-Chaker Larabi, et al. Qualinet white paper on definitions of quality of experience. 2013.

[125] Imtiaz Parvez, Ali Rahmati, Ismail Guvenc, Arif I Sarwat, and Huaiyu Dai. A survey on low latency towards 5g: Ran, core network and caching solutions. *IEEE Communications Surveys & Tutorials*, 20(4):3098–3130, 2018.

[126] S Poretsky, J Perser, S Erramilli, and S Khurana. Rfc 4689–terminology for benchmarking network-layer traffic control mechanisms. *IETF, October*, 2006.

[127] Rene Serral-Gracia, Eduardo Cerqueira, Marilia Curado, Marcelo Yannuzzi, Edmundo Monteiro, and Xavier Masip-Bruin. An overview of quality of experience measurement challenges for video applications in ip networks. In *Wired/Wireless Internet Communications*, pages 252–263. Springer, 2010.

[128] Ken Iiyoshi, Mahrukh Tauseef, Ruth Gebremedhin, Vineet Gokhale, and Mohamad Eid. Towards standardization of haptic handshake for tactile internet: A webrtc-based implementation. In *2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pages 1–6. IEEE, 2019.

[129] Changdong Liu, Yong Xie, Myung J Lee, and Tarek N Saadawi. Multipoint multimedia teleconference system with adaptive synchronization. *Selected Areas in Communications, IEEE Journal on*, 14(7):1422–1435, 1996.

[130] Tobias Hoßfeld, Andreas Binzenhöfer, Markus Fiedler, and Kurt Tutschku. Measurement and analysis of skype voip traffic in 3g umts systems. *Proc. of IPS-MoMe*, pages 52–61, 2006.

[131] Leida Li, Hao Cai, Yabin Zhang, Weisi Lin, Alex C Kot, and Xingming Sun. Sparse representation-based image quality index with adaptive sub-dictionaries. *IEEE Transactions on Image Processing*, 25(8):3775–3786, 2016.

[132] Alexandre Abraham, Elvis Dohmatob, Bertrand Thirion, Dimitris Samaras, and Gael Varoquaux. Extracting brain regions from rest fmri with total-variation constrained dictionary learning. In *International conference on medical image computing and computer-assisted intervention*, pages 607–615. Springer, 2013.

[133] Probal Mitra and Günter Niemeyer. Model-mediated telemanipulation. *The International Journal of Robotics Research*, 27(2):253–262, 2008.

[134] Xiao Xu, Burak Cizmeci, Anas Al-Nuaimi, and Eckehard Steinbach. Point cloud-based model-mediated teleoperation with dynamic and perception-based model updating. *IEEE Transactions on Instrumentation and Measurement*, 63(11):2558–2569, 2014.

[135] Grigore C Burdea and Philippe Coiffet. *Virtual reality technology*. John Wiley & Sons, 2003.

[136] Patrick Le Callet, Sebastian Möller, Andrew Perkis, et al. Qualinet white paper on definitions of quality of experience. *European network on quality of experience in multimedia systems and services (COST Action IC 1003)*, 3(2012), 2012.

[137] Farid Pazhoohi and Alan Kingstone. The effect of movie frame rate on viewer preference: An eye tracking study. *Augmented Human Research*, 6(1):1–5, 2021.

[138] Andrew Perkis, Christian Timmerer, Sabina Baraković, Jasmina Baraković Husić, Søren Bech, Sebastian Bosse, Jean Botev, Kjell Brunnström, Luis Cruz, Katrien De Moor, et al. Qualinet white paper on definitions of immersive media experience (imex). *arXiv preprint arXiv:2007.07032*, 2020.

[139] Hawazin Faiz Badawi, Fedwa Laamarti, and Abdulmotaleb El Saddik. Devising digital twins dna paradigm for modeling iso-based city services. *Sensors*, 21(4):1047, 2021.

[140] Nathan Weisz and Jonas Obleser. Synchronisation signatures in the listening brain: a perspective from non-invasive neuroelectrophysiology. *Hearing research*, 307:16–28, 2014.

[141] Valay A Shah, Maura Casadio, Robert A Scheidt, and Leigh A Mrotek. Spatial and temporal influences on discrimination of vibrotactile stimuli on the arm. *Experimental brain research*, 237(8):2075–2086, 2019.

[142] 3D vision. nvidia corporation, 2021. Online accessed; Thursday, May 6, 2021.

[143] Facebook Corporation. Oculus rift s, 2021. Online accessed; Thursday, May 6, 2021.

[144] Giulia Cisotto, Anna V Guglielmi, Leonardo Badia, and Andrea Zanella. Classification of grasping tasks based on eeg-emg coherence. In *2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 1–6. IEEE, 2018.

[145] Hye-Geum Kim, Eun-Jin Cheon, Dai-Seg Bai, Young Hwan Lee, and Bon-Hoon Koo. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation*, 15(3):235, 2018.

[146] Imtiaj Ahmed, Ville J Harjunen, Giulio Jacucci, Niklas Ravaja, Tuukka Ruotsalo, and Michiel Spape. Touching virtual humans: Haptic responses reveal the emotional impact of affective agents. *IEEE Transactions on Affective Computing*, 2020.

[147] zephyr Corporation. zephyr belt, 2021. Online accessed; Thursday, May 6, 2021.

[148] Naveen Masood and Humera Farooq. Investigating eeg patterns for dual-stimuli induced human fear emotional state. *Sensors*, 19(3):522, 2019.

[149] Sirawaj Itthipuripat, Jan R Wessel, and Adam R Aron. Frontal theta is a signature of successful working memory manipulation. *Experimental brain research*, 224(2):255–262, 2013.

[150] Elisa Magosso, Giulia Ricci, and Mauro Ursino. Alpha and theta mechanisms operating in internal-external attention competition. *Journal of Integrative Neuroscience*, 20(1):1–19, 2021.

[151] Zheng-yan Jiang. Study on eeg power and coherence in patients with mild cognitive impairment during working memory task. *Journal of Zhejiang University Science B*, 6(12):1213–1219, 2005.

[152] J Craig Henry. Electroencephalography: basic principles, clinical applications, and related fields. *Neurology*, 67(11):2092–2092, 2006.

[153] Extend Reality Ltd. Virtual reality toolkit, 2021.

[154] Photon. Photon, 2021. Online accessed; Sunday, Jan 17, 2021.

[155] Ayon Dey. Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3):1174–1179, 2016.

[156] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N

Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97, 2012.

[157] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[158] Jee-Hwan Ryu, Jordi Artigas, and Carsten Preusche. A passive bilateral control scheme for a teleoperator with time-varying communication delay. *Mechatronics*, 20(7):812–823, 2010.

[159] Xiao Xu, Clemens Schuwerk, Burak Cizmeci, and Eckehard Steinbach. Energy prediction for teleoperation systems that combine the time domain passivity approach with perceptual deadband-based haptic data reduction. *IEEE transactions on haptics*, 9(4):560–573, 2016.

[160] Siwen Liu, Molin Li, Xiao Xu, Eckehard Steinbach, and Qian Liu. Qoe-driven uplink scheduling for haptic communications over 5g enabled tactile internet. In *2018 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pages 1–5. IEEE, 2018.

[161] Amit Bhardwaj, Burak Cizmeci, Eckehard Steinbach, Qian Liu, Mohamad Eid, Jose AraUjo, Abdulmotaleb El Saddik, Ritu Kundu, Xun Liu, Oliver Holland, et al. A candidate hardware and software reference setup for kinesthetic codec standardization. In *2017 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pages 1–6. IEEE, 2017.

[162] Xiao Xu, Qian Liu, and Eckehard Steinbach. Toward qoe-driven dynamic control scheme switching for time-delayed teleoperation systems: A dedicated case study. In *2017 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*, pages 1–6. IEEE, 2017.

[163] Mamta Agiwal, Abhishek Roy, and Navrati Saxena. Next generation 5g wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 18(3):1617–1655, 2016.

[164] Hussein Al Osman, Mohamad Eid, Rosa Iglesias, and Abdulmotaleb El Saddik. Alphan: Application layer protocol for haptic networking. In *Haptic, Audio and Visual Environments and Games, 2007. HAVE 2007. IEEE International Workshop on*, pages 96–101. IEEE, 2007.

[165] Riverbed. Riverbed modeler, 2018. Online accessed; July 1, 2018.

[166] Rajesh Gupta, Sudeep Tanwar, Sudhanshu Tyagi, and Neeraj Kumar. Tactile internet and its applications in 5g era: A comprehensive review. *International Journal of Communication Systems*, 32(14):e3981, 2019.

[167] University Southhampton. Haptics/tactile devices, 2019. Online accessed; Sunday, Dec 30, 2019.

[168] Xiao Xu, Burak Cizmeci, Clemens Schuwerk, and Eckehard Steinbach. Model-mediated teleoperation: toward stable and transparent teleoperation systems. *IEEE Access*, 4:425–449, 2016.

[169] Shrawan Kumar. Theories of musculoskeletal injury causation. *Ergonomics*, 44(1):17–47, 2001.

[170] Keiji Fukuda, Stephen E Straus, Ian Hickie, Michael C Sharpe, James G Dobbins, and Anthony Komaroff. The chronic fatigue syndrome: a comprehensive approach to its definition and study, 1994.

[171] Jan-Niklas Antons, Robert Schleicher, Sebastian Arndt, Sebastian Möller, and Gabriel Curio. Too tired for calling? a physiological measure of fatigue caused by bandwidth limitations. In *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, pages 63–67. IEEE, 2012.

[172] Delsys.com. Emg sensors, 2019. Online accessed; Sunday, Dec 30, 2019.

[173] Kanav Kahol, Mario J Leyba, Mary Deka, Vikram Deka, Stephanie Mayes, Marshall Smith, John J Ferrara, and Sethuraman Panchanathan. Effect of fatigue on psychomotor and cognitive skills. *The American Journal of Surgery*, 195(2):195–204, 2008.

[174] CY Tang, B Stojanovic, CP Tsui, and Milos Kojic. Modeling of muscle fatigue using hill's model. *Bio-medical materials and engineering*, 15(5):341–348, 2005.

[175] Liang Ma, Damien Chablat, Fouad Bennis, Wei Zhang, Bo Hu, and François Guillaume. Fatigue evaluation in maintenance and assembly operations by digital human simulation in virtual environment. *Virtual Reality*, 15(1):55–68, 2011.

[176] Kentaro Kotani and Ken Horii. An analysis of muscular load and performance in using a pen-tablet system. *Journal of physiological anthropology and applied human science*, 22(2):89–95, 2003.

[177] Haiwei Dong, Izaskun Ugaldey, and Abdulmotaleb El Saddik. Development of a fatigue-tracking system for monitoring human body movement. In *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pages 786–791. IEEE, 2014.

[178] M Poyade, M Kargas, and V Portela. Haptic plug-in for unity, 2014. *Digital Design Studio (DDS), Glasgow School of Art, Glasgow, United Kingdom.*

[179] Mario Cifrek, Vladimir Medved, Stanko Tonković, and Saša Ostojić. Surface emg based muscle fatigue evaluation in biomechanics. *Clinical Biomechanics*, 24(4):327–340, 2009.

[180] Angkoon Phinyomark, Chusak Limsakul, Huosheng Hu, Pornchai Phukpattaranont, and Sirinee Thongpanja. *The usefulness of mean and median frequencies in electromyography analysis.* INTECH Open Access Publisher, 2012.

[181] Roberto Merletti, Marco Knaflitz, and CARLO J De Luca. Myoelectric manifestations of fatigue in voluntary and electrically elicited contractions. *Journal of Applied Physiology*, 69(5):1810–1820, 1990.

[182] Delsys. emgscripts-fatigue, 2019. Online accessed; Sunday, Dec 30, 2019.