Doctoral Dissertation

Academic Year 2017

Embodied-Driven Design:

A Framework to Configure Body

Representation & Mapping

Keio University Graduate School of Media Design

MHD Yamen Saraiji

Embodied-Driven Design:
A Framework to Configure Body Representation & Mapping

by

MHD Yamen Saraiji

Masters in Media Design, Keio University. Japan, 2015

B.Sc. CS Damascus University. Syria, 2010

A Doctoral Dissertation

submitted to Keio University Graduate School of Media Design

in partial fulfillment of the requirements for the degree of

Ph.D. in Media Design

Thesis Committee:

Associate Professor Kouta Minamizawa	(Supervisor)
Professor Hideki Sunahara	(Member)
Professor Keiko Okawa	(Member)
Professor Naohito Okude	(Member)
Professor Emeritus Susumu Tachi	(Member)

Embodied-Driven Design:

A Framework to Configure Body Representation & Mapping

Category: Science / Engineering

Abstract

Our bodies are inherently constrained by their physical structure and are limited in the perceptual spectrum which we use to perceive and act through. We use tools and devices to alter or overcome certain physical characteristics and capacities of our bodies temporarily or permanently, enhancing our body functions beyond the biological limits. Augmented Human, a recent research paradigm, was established addressing body augmentation using technological driven approaches, accelerating the process of machine-mediated interaction. However, design approach of such interactions has been problem driven, focusing on solving certain bodily limitations. There have been no systematic approaches for a body driven design that allows the re-usability, or re-configuration of human body interaction through the digital means. How to design intuitive, natural connections to digital means so they can be integrated and embodied into our bodies and actions?

This thesis addresses such a question by proposing an embodied driven methodology for body interaction with digital means. *Embodied-Driven Design (EDD)* is an approach to configure body representation and mapping with the augmentative digital means. EDD explores the plasticity of the human body to adapt to a new schema, and proposes a systematic design methodology to achieve such new body schema. Using the proposed approach, an abstract model of the human body is used to define the flow between body's modalities (body's sensory organs and joints) with a target representation (referring to a digital mean) creating a new body schema. Each of these modalities acts as information channels of the human body as well as of the representation(s). The information flow can be altered, mapped, or combined between the modalities within the meta-model, reflecting the results as embodied actions through the used representation.

Embodied-Driven Design framework was designed and implemented to realize the proposed approach. The framework introduces body schema building blocks that act as transfer functions of body's modalities information (sensory and spatial). Meta-model defines the relationships and the flow between these building

blocks with the target representation(s). To facilitate body schema design process, four meta-modeling categories were driven from the framework: *Direct Body Transfer*, *Representation Alteration*, *Topology Reconfiguration*, and *Modality Expansion* meta-models. The first category focuses on directly mapping body modalities with the representation maintaining their information flow without alteration. The second category addresses body physical limitations and constraints by modifying its physical attributes through the representation. The third category reconfigures body topology and mapping with the representation to overcome a physical disability or limitation. The last category focuses on expanding modalities' information flow into multiple representations, achieving non-linear body mapping.

The use of EDD meta-modeling approaches in body augmentation has been highlighted through addressing five different use-cases to overcome body physical limitations or disabilities. By changing the body mapping through the meta-model, we showed the possibility to: stretch the spatial awareness of the body (discussed in Telexistence based Surveillance System), substitute body with a virtual one to expand its physical limits (discussed in Enforced Mutual Presence), substitute a disabled body modality with another one (discussed in HUG Project), reconfigure body limbs to an gain extra two arms (discussed in MetaLimbs), and to expand modality awareness to multiple locations achieving ubiquitous presence (discussed in Layered Presence). These five projects showed the plasticity of human body and cognition to adapt to new modality configurations.

Embodied-Driven Design can be used in human augmentation, supportive and assistive body extensions, and beyond spatial interaction. Using EDD, it is possible to further explore body plasticity in a systematic manner, and use the existing knowledge of embodiment in designer oriented approach. We hope Embodied-Driven Design can help at radically changing the design approach of human augmentation, and to assist researchers and designers in prototyping and evaluating their ideas.

Keywords:

Human Augmentation, Phenomenology, Embodiment, Perception & Awareness, Body Schema Mapping, Meta-Modeling

Keio University Graduate School of Media Design

MHD Yamen Saraiji

Table of Contents

List of Figures

List of Tables

Acknowledgements

بِسْمِ اللَّه الرَّحمَنِ الرَّحيمِ

The Ph.D. journey I had in the past three years have faced many challenges and obstacles on research-wise, personal wise, and on my overall lifestyle. Without the immense support I had from many kindred minds and colleagues that shared parts of this journey over research lab, meetings and ideation sessions, or coffee, whom to I am heartily thankful.

First and foremost I would like to thank my supervisor and superhuman prof. Kouta Minamizawa for his great support and encouragement, and providing me the needed environment for conducting my research freely.

Also, I owe great gratitude to my thesis co-supervisors: prof. Hideki Sunahara, prof. Keiko Okawa, prof. Naohito Okude, and prof. Susumu Tachi for their great iterative feedback, helping me constructing my thesis thoroughly and logically. Highlighting prof. Okude's continues challenges that, by all means, reflected a great deal improving my writing style. Great gratitude goes to prof. Tachi who involved me in the Telexistence project, which I worked with for the past six years.

Also, I would like to thank external reviewers: prof. Kai Kuzne, Tilman Dingler, prof. Mathew Waldman, and Roshan Peiris who kindly spared some of their time to read, criticize, and giving me feedback of my written work. And for the support I received from prof. Masahiko Inami on both academic and professional level.

I also owe a great debt of gratitude to many friends and colleagues in my laboratory and beyond who supported me while conducting my research and writing my thesis. I would like to highlight few out of many of these amazing people on this list. Emily Kojima was more like our spir-

I

FOUNDATIONS & PHILOSOPHY

"But O alas, so long, so far
Our bodies why do they forbear?
Theyre ours, though they are not we, we are
The intelligencies, they the sphere"
JOHN DONNE, THE ECSTASY

Chapter 1

Introduction

"Before paper, wires, and silicon, the primordial communication medium is the body. At the center of all communication rests the body, the fleshy gateway to the mind."

Frank Biocca, The Cyborg's Dilemma 1997

We, as human beings or Homo sapiens, have evolved through a long adaptation process both our physical and cognitive abilities to fit into our environment and to achieve highly complex tasks using our bodies. Although compared to other biological animals, our bodies still considered quite limited in terms of sensory feedback. For example, we can not see or hear as well as other hunting animals due to the limited visual & auditory spectrum, neither we can fly or dive in the sea same as other mammals can do. What was the edge that not only made us survive but also to be the superior race on earth? It is basically the use of tools and other mediums to overcome our biological limitations and to give us an edge both physically and mentally. Using the tools, our boundaries were not limited anymore to the skin, or the reach of the hand, but rather it created an extension to our physical bodies. Furthermore, when digital tools and means has emerged at our disposal, and integrated into our daily routine, the boundaries of our world has extended beyond the skull and the sight of our eyes. The digital era created radical changes in our cognitive skills, and changed how we manage our work and daily social activities.

1.1 Radical Bodies

The role of tools has effected our adaptation to our environment, and created radical means to our body functions. The use of technological means can be generally divided into two application categories: supportive, and augmentative. The first category, supportive applications domain, is mainly concerned with recovering certain body modality disability (such as a lost limb, poor vision), while the second category, augmentative applications, is used to expand what a healthy and functional body modality can do (such as providing digital information to our vision modality). These two categories are illustrated in Figure 1.1 . Both categories falls under what is referred into as "Radical Bodies", that is the use of tools and devices along our biological body to overcome what can be considered as a limitation, or leveraging certain functions, and thus changing the behaviour of our bodies radically.
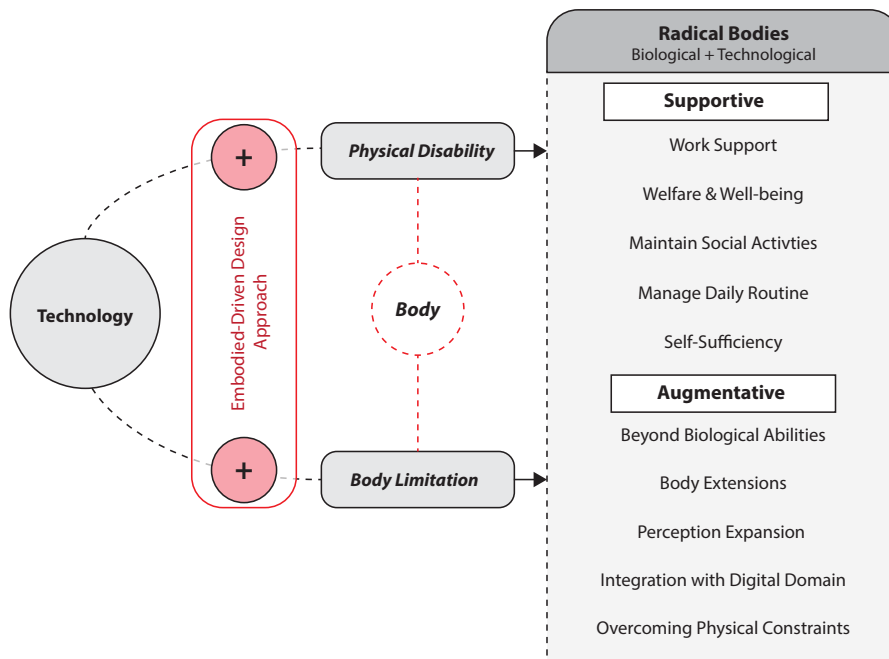
Figure 1.1: Radical Bodies, and the proposed Embodied-Driven Design (EDD) approach position.

Supportive tools help the physically disabled community to regain their lost functions and to assist them in engaging in their daily and social activities in an independent manner. According to the American Community Survey (ACS) [18], an estimation of 12.6% of the overall population in the United States suffers from a physical disability in 2015. The physical disability is varied as a hearing, vision, cognitive, ambulatory, self-care, and independent living. Also, the survey showed that the rates of disability increase with age. From the entire US population, for working class with ages of 18-64, the rate was 10.5% (contributing to 51.1% of the disabled community), while for elder people with ages of 65 and above the rate was 35.4% (contributing to 41.2% of the disabled community). The body inherited functions (non-cognitive) showed to be the highest factors to the disabilities. Tools for this category of users are described as Assistive and Adaptive technologies [1]. Assistive Technology is any object or system that increase, maintain, or improve functional capabilities of individuals with disabilities (such as glasses, hearing aids, ...etc), while an Adaptive Technology is any object or system that is specifically designed for the purpose of increasing or maintaining the capabilities of people with disabilities and seldom to be used by non-disabled persons (such as wheelchairs, white cane, ...etc). That is, assistive technologies are the wider umbrella that adaptive technologies are a subset of. Adaptive tools usually require the disabled person to adapt to its new function and to train his cognitive skills to substitute his lost function (a blind person needs to adapt to the limits and range of his white cane, as it will become his eyes). Thus it can be seen that our bodies (disabled or functional) have the capacity and plasticity to adapt to new representations of its functions. It is, however, the adaptation process is variant depending on the design and interaction with the substitution device or tool. If the user is required to maintain constant awareness toward the device used (through his various perceptual modalities), then the device is considered as a foreign object which does not function as an embodiment of his action.

Augmentative tools on the other hand leverage the functions of the hu-

man body, and allows to extend and expand beyond the inherited biological limitations. Augmentation technologies can be categorized into two classes: invasive and wearable (non-invasive). Each set of these two classes has its own merits as well as its disadvantages physically, socially, technically,..etc. As an example of the earliest trials on direct body implants of electronic devices other than for medical purposes, in 1998 professor Kevin Warwick implanted a radio transmitter into his left arm which communicates with computers to open doors whenever he approaches them. Interestingly, after nine days of using it, he reported the implant felt as being part of his body [104]. Although this experiment can be seen as naive (a wearable device can serve the same purpose), but it was one of the initial trials of radical bodies changes, and has a future of transcending our bodies via the combination of biological and technological means. However, one of the major disadvantages of such augmentation is the high cost of replacement and upgrades both physical wise (need for a surgery, the risk of infections) and money wise (the cost for that surgery). Such augmentations are still immature technological wise, and widely unacceptable socially. The other set of augmentation, non-invasive type, had a better success at proving that our bodies can be altered safely and at rather low costs. This set has existed as far as the glasses were used to enhance poor vision. Neil Harbisson [33] used this type of wearable devices to overcome his color blindness by using a head-mounted camera which translates colors into sound frequency, and after adequate usage of it his perception adapted to this new perception. Current wearable technologies even provide a way to alternate our perception and realities. A **perceptual proxy** device, such as Virtual Reality (VR) Head Mounted Display (HMD), headphones, or tactile actuators, under certain conditions, can deliver sufficient information to sensory organs to convince the brain of the new reality. The idea of altering and augmenting our bodies has been a driving force for various technologies and media we use on a daily basis. This wave produced a constant hype towards human-centered technologies. A visible trend in empowering technologies can be seen by Gartner's review of emerging technologies hype cycle of the year
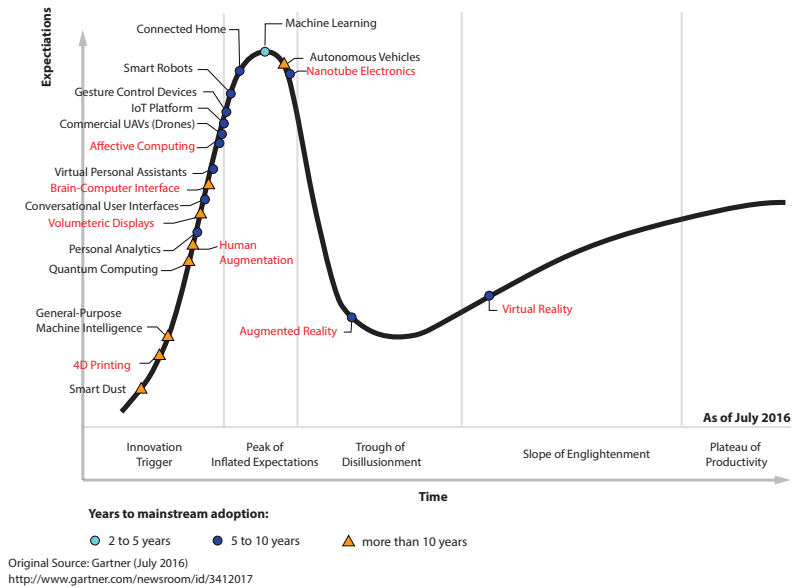
Figure 1.2: Expectations of human centered technologies. Gartner's Hype Cycle for Emerging Technologies 2016 (Modified from source for formatting purpose).

2016 [27] in Figure 1.2 . Virtual Reality and Augmented Reality have already become mature enough for the end user's application due to the long history of trials and applications [81]. However, Human Augmentation is still an emerging topic with high expectations for enhancing and empowering our bodies instead of just being a supportive hardware or for rehabilitation purposes. This topic is still in its infancy and requires further philosophical approaches and scientific experiments to evaluate its effectiveness.

> *Radical Bodies are the result of the technological integration with our bodies, leveraging our physical constraints and enabling wider set of embodied functions.*

These tools and technological means have influenced radical changes to our perception and levitated our physical/mental abilities exponentially. The radical changes do not have to refer to the physical alteration of the human body (or mutations), but rather to how we act and communicate with the environment surrounding us. For example, today's advancement

in the digital communication enabled us to overcome the spatial and temporal constraints. Internet-connected smartphones which are being used on a daily basis, have created radical changes in our social activities. A user communicating through it with a remote person is referred to be "here" and "there" at the same time. "Here" is where his physical body is located at which uses the smartphone to type and watch, while "there" is where a part of his cognition and awareness is actually at. Furthermore, virtual reality tools can carry the entire sense of presence and awareness into a totally different space or reality, in which the human body is no longer constrained by any physical limitations. The physical body is no longer defining the action, rather is being the medium between the internal intention and where the action is being performed. The more these technologies are integrated to our daily routine and appliances, the further the radical changes will be reflected on our bodies, and the more our biological bodies will be transparent in the process (mediums of actions). What if its possible to think of the body as being just a medium? Is possible to redefine this medium or "rewire" methodically? How to maintain this newly defined medium as being part of the body, or being the new body? In this thesis I address these questions by proposing the idea of "Embodied-Driven Design (EDD)", the use of body functions and modalities abstraction to allow us to redesign the functionality and the mapping of the body while maintaining the sense of ownership toward it. Figure 1.1 shows the position of Embodied-Driven Design as an approach to establish a binding between our biological bodies and technologies in order to over come a physical disability or a limitation inherited within our bodies. EDD establishes an abstract modeling approach in order to reconfigure our bodies with an alternative representations, thus altering the embodied actions and cognition using mediated technology.

1.2 Embodied-Driven Design

In this section, an overview of the thesis is described, the motivation and concept progression, along with the proposed approach addressing the thesis statement.

1.2.1 Thesis Statement

In this thesis, I discuss the possibility to re-configure our bodies in a way that can alter our cognitive abilities or overcome our innate limitations and physical disabilities using mediated representations. With such re-configurations and alterations of our body schema, we can look at the use of the representations (for example technological tools) as being part of our bodies, embodied in a way we interact through them instead of using them. This sort of embodiment can establish a new form of interaction loop between our bodies and the environment (locally or in a different location), that our cognition is shifted into the active performance of intended actions instead of how we use the tools to achieve such actions. The relationship between body and action is mediated through embodied means of interactions and representations. This shift of perspective in using tools and technology would create an impact on how we perceive through and act within the environment in a way our bodies become actively part of it through our mediated and embodied actions.

1.2.2 Motivation & Approach

The capability of remapping our exteroceptive modalities (visual, auditory, tactile, taste, and smell), and proprioceptive modalities such as body and joints posture with an alternative body representation can help to levitate the upper bounds of our physical limits. For example, for a true ubiquitous presence a person should be capable of receiving inputs from multiple locations simultaneously, thus we can say this person is ubiquitous as being in everywhere. His body and sensory modalities are not "physically" there,

but his perception of these modalities is linked to them, and body sensory flow has became altered and remapped. Other topics can also be achieved, such as expanding the localization of body parts from being tightly coupled in one location (on the body) into being loosely coupled and distributed in multiple locations. Adding to that, body representation does not have to be tangible as long as it serves the application intended to, for example when the body is needed to represent the visual appearance of the user or the state the user is at, then it's sufficient to use a non-physical representation for that purpose. In these terms, the concept of body stretches from a tight physical representation which its parts are coupled spatially and tangibly, into loose physical/virtual representations that can be restructured and remapped. This expansion would provide new forms of awareness alterations and perception enhancements. Indeed there are issues that might arise when new forms of perceptual augmentation occur which we are not adapted to, such as perception overload (when several information sources are provided into the same perceptual modality, such as eyes or ears), or non-intuitive interaction is being used (not correlated with body motion, and requires training to learn it).

In this thesis, I argue that the human body schema can be altered and adapt to new forms of body structures as long as the actions mapped between the body and its new representation are correlated. Embodied-Driven Design conceptualizes the idea of using digital media and mediated technologies to alter human body properties on a high and abstract level, achieving new ways to transcend our perception and body capabilities. This approach addresses both local and remote body mapping, expanding the human spatial reach and reducing the physical constraints. Embodied-Driven Design derives from the school of philosophical thought named as "phenomenology", which is concerned with how we perceive, experience, and act in the environment surrounding us using our body. Further details regarding the philosophical background of this research will be discussed in Chapter 2 .

1.2.3 Concept Development

The human body is the interface of the mind, the medium to the tangible world and the embodiment of our thoughts. Its where the action is performed at and its the bridge which connects our world too, what we may say our self. My research and background in Telexistence area had inspired me to focus more on the possibilities to use mediated technology to expand our body instead of just replicating it. How to create a model of the human body that enables us to extend, enhance, and modulate our embodied functions was and still the driving force of Embodied-Driven Design . I believe that in order to come up with such a model, an abstraction of the human body is needed that provides a systematic approach to embodied modeling.

This thesis investigates toward an abstraction model of the human body that enables us to redefine the actions and affordabilities which our cognition can represent. Our bodies constitute physical limitations, while the inner mental model of the body is far more adaptable to variations and reconfigurations. This work shows the possibility to rewire (or reprogram) the body in order to overcome an existing physical disability or to leverage the capacity of human body perceptual & postural limitations. These arguments are supported by philosophical background research and empirical studies that build up the concept of Embodied-Driven Design .

The idea behind Embodied-Driven Design was developed across several projects and iterations which helped me to generalize the idea of meta-modeling, and the design of body schema transfer systems. Along these iterations, four different meta-modeling approaches were developed and used to address body limitations and physical disabilities. These meta-modeling approaches were used in five projects. These iterations along with the projects that involved in each of them are illustrated in Figure 1.3 with the corresponding projects that address.

The first design iteration *Direct Body Transfer* was initiated based on my involvement in Telexistence area of research since 2013. This approach develops the concept of mapping our body's functions and modalities di-
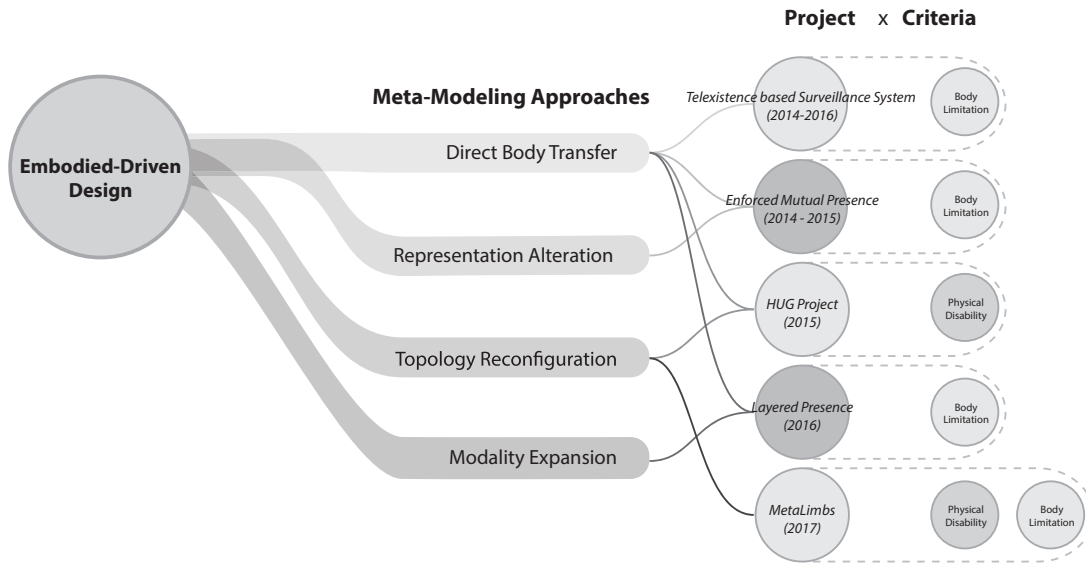
Figure 1.3: The use of Embodied-Driven Design Meta-Modeling approaches to address five projects, highlights the need of such design approach to resolve body limitations and disabilities.

rectly into a corresponding representation without the need of alteration or modifications. Such mapping pattern is useful to transfer body parts directly from one location to another. Telexistence type of systems uses such pattern to copy user's body motion into a remotely connected avatar system, and replicate the feedback back to user's body. "Telexistence based Surveillance System" project in the period of 2014-2016 (Section 5.1) mainly uses this design approach to overcome body spatial limitations.

The second design iteration *Representation Alteration* aims to overcome the physical limits constituted by our bodies. It combines a hybrid approach for altering the representation of our bodies using virtual and physical means. Thus the body is mediated through tangible and non-tangible representation overcoming certain physical limitations. "Enforced Mutual Presence" achieved in 2014-2015 (Section 5.2), explores the use of virtual representation along a physically teleoperated system to achieve a full sense

11

of body presence despite the fact the system does not have a full physical body. Virtual hands substitute user's arms were projected locally and remotely at the teleoperated system. This substitution not only helped at providing the full awareness of the body, but also the use of non-physical hands helped to overcome the tangible constraints of arm's length limits. Thus the operator is capable to stretch his body across spatial space. With such alternative body representations, it is possible to design new body forms that overcome our bodies' physical limits (such as extending arm reach and pointing for mobility impaired people).

The third design iteration *Topology Reconfiguration* explored the plasticity of human body operation mapping, mainly the postural mapping. Two projects helped in formulating this design approach: the first is "HUG Project" in 2015 (Section 5.3) that used eye gaze input to control the head rotation of a remote system supporting a physically disabled person to operate it. The second project is "MetaLimbs" in 2017 (Section 5.5) that explored remapping body legs with artificial arms mounted on user body, thus expanding the number of independent arms. Both projects showed the immediate adaptation the body can achieve to a different mapping schema. This iteration established the design of body postural topology that is used as one of the foundations of this approach.

Last design approach explored is *Modality Expansion* that addresses the possibility to expand the perceptual modalities of our body and the capacity to remap our sensory and postural topology to multiple representations simultaneously. "Layered Presence" in 2016 (Section 5.4) studied this body schema and proposed a modeling approach to achieve such expanded spatial awareness. This project helped to shape the perceptual modalities mapping and an interaction model which helps in developing this approach.

These four iterations motivated the design of a meta-modeling approach for human body schema, and defined the foundations of the proposed modeling framework. Embodied-Driven Design and its realization as a framework are aimed to assist designers and researchers in this field to prototype and deploy new forms of body schema.

1.2.4   Research & Design Topics:

Table 1.1   summarizes an overview of the topics that constructs the concept of Embodied-Driven Design . Research and design topics addressed in this thesis outline body schema modeling approach.

The first part addressed is focused on the design approach for Embodied-Driven Design and is discussed in Chapter 3 . A framework is developed based on a systematic model of human body schema in which its possible to remap and restructure it hierarchically. It outlines the methodology to form a meta-modeling approach that realizes the proposed concept. EDD is constructed and designed based on the following topics:

T1. **Loosely-Coupled Modalities (LCM):**

Possibility to decouple body modalities into a set of "input/output channels", and to re-couple them. The loosely representation helps to break the coupling of the physical modalities and constraints, and thus allows to achieve a higher level of mapping of the human body to its representation(s). LCM defines two topologies for the human body: Postural and Sensory. These topologies serve as generalization and abstraction for the postural modalities of the body (connection between the various body joints), and the sensory modalities. Further details are discussed in Section 3.2 .

T2. **Body Schema Meta-Modeling:**

Ability to remap our bodies modalities into new body schema. This topic argues that our body structure can be reconfigured or reprogrammed. The meta-models define the structure of the sensory and postural channels mapping between the human body and the corresponding representation(s)' modalities. In these models, transfer blocks are used to characterize the function and data flow between the modalities used. Further details are discussed in Section 3.3 .

T3. **Meta-Modeling Approaches:** Four meta-modeling approaches are proposed that summarizes the general usability methods of EDD

	Category	Reference
Topic	**Embodied-Driven Design**	Chapter 3
T1.	Loosely-Coupled Modalities: Recon-structure body schema based on a hierarchical topology	Section 3.2
T2.	Body Schema Meta-modeling: Abstraction of body modalities interaction and mapping	Section 3.3
T3.	Meta-Modeling Approaches: Four design patterns for body schema editing	Section 3.4
Project	**Meta-Modeling Evaluation**	Chapter 5
P1.	Telexistence base Surveillance System: Addressing spatial limitations	Section 5.1
P2.	Enforced Mutual Presence: Addressing body physical limitations	Section 5.2
P3.	HUG Project: Addressing body plasticity to overcome physical disability	Section 5.3
P4.	Layered Presence: Addressing perceptual plasticity	Section 5.4
P5.	MetaLimbs: Addressing body plasticity and limbs limitations	Section 5.5

Table 1.1: Topics and projects addressed in this thesis to realize
Embodied-Driven Design .

within this thesis, and which can be generalized for applications beyond what is discussed. The purpose of these approaches is to facilitate the design process, and to highlight how to address body limitations and disabilities using meta-modeling. Further details are discussed in Section 3.4 .

The second part of this thesis in Chapter 5 , evaluates meta-modeling approaches through five different scenarios addressing human body limitations and/or physical disabilities. An overview of the main points addressed within these projects as follows:

1. Spatial limitations: Expanding body spatial accessibility and awareness through mediated representations, which is mainly addressed through P1. "Telexistence base Surveillance System" in Section 5.1 .

2. Body physical limitations and disabilities: Possibility to use different body representation to overcome physical constraints, addressed through P2. "Enforced Mutual Presence" in Section 5.2 , P3. "HUG Project" in Section 5.3 ,and P5. "MetaLimbs" in Section 5.5 .

3. Perceptual plasticity: Use of EDD to alter body modalities to expand the amount of perceptual feedback. P4. "Layered Presence" in Sec-

tion 5.4 illustrates visual perception plasticity into multiple remote locations.

1.2.5 Summary of Contribution

The main contributions of this thesis can be categorized as follow:

- Conceptual Contribution:

  This thesis provides the building blocks for expanding and enriching the body augmentation area of research by connecting the theoretical studies and various questions in the philosophy of body ownership and metacognition with systematic methods for testing these theories. By providing a framework for body schema design and modalities alteration, it is possible to design and evaluate various models of body schema with an experience-oriented approach.

- Design and Implementation:

  This thesis presents the design strategy of body schema meta-modeling based on four modeling categories: direct body transfer, representation alteration, topology reconfiguration, and modality expansion. These meta-models summarize the general design approaches which the body schema can be designed through. Each of these categories is addressed separately.

- Social and Industrial Contributions:

  During the course of this thesis, two projects in which the Embodied Driven Design (EDD) framework was used into. The first is "HUG project" in which EDD used as a social contribution, and "Telexistence based Surveillance System" as an industrial contribution.

- Technical Contribution:

  Along the development of this work, various system optimization and technical details that assist the developers in telepresence/telexistence

related applications were described in this thesis, such as *Foveated Streaming*. These details are further discussed in Chapter 4 .

- Physical Embodiment Toolkit:

  As an outcome of this work, an embodiment toolkit has been designed and implemented. This toolkit provides essential functions for remote presence applications, and its functions have been integrated into the EDD framework as a set of representation modalities which can be accessed and mapped with user's body. This toolkit allows designers to use rapid prototyping and testing of their ideas along with the meta-editor.

1.3 Thesis Outline

This research builds on the findings related to human body plasticity found in the areas of philosophy of body and perception, cognition science, and body augmentation. Also, it introduces the idea of Embodied-Driven Design : altering body schema and sense of presence using a systematic and at a meta-level design. This thesis highlights the idea of digitized modalities -converting the biological/analog sensory and motor modalities into a digital representation- in order to overcome the inherited physical constraints and limitations.

1.3.1 Organization

The core of this thesis is divided mainly into four parts and seven chapters organized as follows:

- **Part I : Foundations & Philosophy**

  Chapter 1 . provided the motivation behind this work, and briefly introduced the proposed concept: Embodied Driven Design.

  Chapter 2 . explores the literature and philosophical background which Embodied-Driven Design is established on top of. Discusses topics related to Embodiment and Cognition that are regarded as the fundamentals of this research, and in which the idea of Embodied-Driven Design is placed within. It also provides several related work in human augmentation, presence alteration, and body representation which this research concept derives from, highlighting the contributions of this thesis.

- **Part II : Embodied-Driven Design**

  Chapter 3 . discusses in further details the proposed concept of embodiment design and the overall design considerations for an EDD framework, including the concept of loosely-coupled representation

17

for human body. Four types of meta-modeling categories are described which summarize the design approaches of EDD: Direct body transfer, Representation alteration, Topology reconfiguration, and Modality expansion.

Chapter 4 . provides the implementation process of EDD framework and meta-modeling editor environment, and describes in details the various building blocks for modeling body schema. The development of a physical toolkit is also discussed along EDD framework for general purpose presence embodiment.

Chapter 5 . discusses five different projects and use-cases that highlights the role of EDD meta-modeling to resolve body limitations and physical disabilities. These study cases show the possibilities to modify our body schema to achieve new cognitive capabilities, such as expanding visual awareness and altering body limbs.

- **Part III : Future Direction & Conclusion**

  Chapter 6 . concludes with a summary and reflections of this work, main contributions and outcomes, limitations, and finally presents an outlook for this research future direction.

1.4 Glossary

The following keywords are described in the context of this thesis:

- **Agency**

  Motor skill control, including the subjective experience of control, intention, sensory-motor selection and the conscious experience of will. Further details about the topic [86].

- **Body Schema**

  The mental model or self-image of the body, a subjective experience obtained through the cycle of operation and sensory feedback to create the global structure of one's perceived body.

- **Embodiment**

  The subjective experience of 'having and using' a body which interacts with the external surroundings, and reflects the inner intentions in correlation with cross-modal perception. The act of embodiment alters the mental image of body schema.

- **Loosely-Coupled Modalities**

  The possibility to decouple body modalities into a set of "input/output channels", and to re-couple them. The loosely representations help to break the rigid constraints of the physical modalities, and map body modalities separately.

- **Meta-Modeling**

  The process of defining an abstract model for the relationships between the various body modalities provided with an artificial representation's modalities.

- **Modality**

  In this thesis, a modality refers to a functionality derived from the human body or an artificial representation. The functionality can

be described as an input if its related to the sensory feedback of the body, and an output if its related to an action can be performed by the body, such as joints motion.

- **Ownership**

  A reflected experience of identification a body as being owned by one's self (phenomenally experienced as 'mine-ness').

- **Perceptual Proxy**

  A device used to deliver/receive information to sensory/from actuating organs to induce a new sense of reality or to capture the structure of the body. Such as virtual reality equipment, haptic feedback interfaces, postural tracking devices.

- **Radical Bodies**

  The concept proposed in this thesis reflecting the radical changes of technology progression on human bodies in terms of interaction and perception within an individual or social contexts.

- **Representation**

  In this thesis, representation refers to the physical or virtual form of the modalities which can act as an extension or replacement of certain human body modalities.

Chapter 2

Literature Review

"Before paper, wires, and silicon, the primordial communication medium is the body. At the center of all communication rests the body, the fleshy gateway to the mind"

Frank Biocca, The Cyborg's Dilemma 1997

Embodied-Driven Design approach has been proposed based on the work and philosophy done in embodiment area of research, and derives mainly from Presence Augmentation and alteration, Body Representation, and related disciplines. This chapter explores theories in philosophy regarding embodiment and body schema theories, also discusses related work that focuses on altering body interaction with the environment for augmentation or substitution purposes.

2.1   Body and Embodiment

"I think, therefore I am" Ren Descartes

Several theories of mind and body have been formulated on the idea that our bodies and thinking are two separate worlds. *Computationalism* theories that were derived from this perspective of mind, has been widely adopted as in various fields of robotic design to mimic the sensing behavior of the body (the one that interacts with the physical domain), and the processing behavior of the mind (the one which uses algorithmic computations). These perspectives isolate the role of the body in shaping cognition, and thus they disembody the brain from the body.
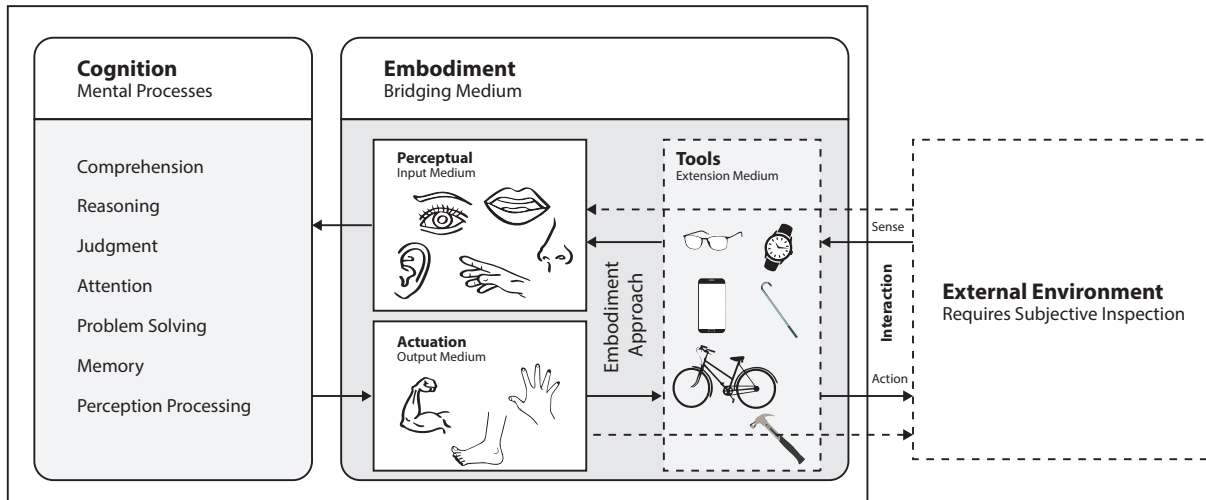
Figure 2.1: Embodiment of tools to extend our cognition and actions into the external environment.

In philosophy and literature, embodiment has taken various definitions and concepts. MacLachlan [55] has defined Embodiment as: *the identification of an abstract idea with a physical entity*, however this base definition is rather too broad and the concept of the *idea* can be self, emotion, intelligence, ...etc. Whenever this *idea* is represented in a tangible matter, it can be declared as being *incarnated*. Although Embodiment falls into different categories (psychology, sociology, computer science, ...etc), however, all are mainly centered around the human being, the entity with cognition and awareness. In literature and philosophy, Embodiment focuses on the relationship between our bodies and the environment surrounding it, how do we conceive information, and links to them inside-out/outside-in our mental model. Furthermore, how human cognition is shaped based on what we perceive, that is, we are what we perceive(d) and experience(d). As Shapiro describes this theme of embodiment as Conceptualization [93], our body structure defines our comprehension of surrounding world. Thus it can be viewed that human body is the interface of the mind, the channels which communicate with the outside world, and as the tangible representation of an internal idea, a thought, or an opinion when being expressed in

a physical form. The expression can be a form of an act (moving a muscle group), verbal sounds (speaking), or a mediated action (through a tool, an interface, or written symbols). Likewise, perception is done through one of the five sensory organs we have: visual, auditory, tactile, smell, and taste organs.

> "The body is our general means of having a world. Sometimes it restricts itself to gestures necessary for conservation of life, and correlatively it posits a biological world around us. Sometimes, playing upon these first gestures and passing from their literal to their figurative sense, it brings forth a new core of significance through them - this is the case of new motor habits, such as dance." (Maurice Merleau-Ponty Phenomenology of Perception 1945 p.147)

An ecological theory of perception was presented by Gibson [28] which suggests that brain resonates to information presented in the environment surrounding it. That is, a structured information stimulating the sensory organs, and a brain resonating to these sensory inputs. Adding to that, Gibson's view of information acquisition insists on the fact that sensory organs should hunt for them by active involvement in the environment.

> "The perceptual systems, as it turns out, correspond to the organs of active attention with which organism is equipped... Head movements, ear movements, hand movements, nose and mouth movements, and eye movements are part and parcel of the perceptual systems they serve. These adjustments consti- tute modes of attention... They serve to explore the information available in sound, mechanical contact, chemical contact, and light." (Gibson 1966 p.58)

Modern philosophers looked into the expansion of our mind using the tools that we use, representing an extension of our bodies [58, 93]. De- scribing the body extension phenomenon we experience when using tools

Figure 2.1 . As an example, visually impaired and blind people commonly use a white cane to extend their tactile perception, and to feel objects surrounding them. Another example for entire body expansion is a bus driver, what distinguishes a professional driver from a novice one is the capability to have complete awareness of the bus he is driving, and feel it as being an extension of his body. To use the mirrors as an extension of his visual sense, driving wheel as an expansion of his motor cortex functions, and the horn as a new way of communication. The process of adapting to the tool also shape our embodiment into it, until we learn how to use it without thinking of the procedure which its operated with, as an example, when driving a bicycle Figure 2.2 we do not think of the handlebar while steering (embodied), actually when we think of how we are using it then we might lose balance and fall off (became disembodied).

Figure 2.2: The use of bicycle morph our bodies, and embodies our actions through it.

Building on top of these principles of embodiment and cognition, psychologists and researchers in this field [9, 26, 62, 77] began to formulate a better understanding of the body and its connection with our comprehensions of self and environment surrounding us. The cycle of interactions as input and outputs using our available modalities with the external world helps to shape a schema of our body with a mental model of it.

2.1.1 Embodied Self

As has been discussed before, the type of representation used to reflect the inner cognitive thoughts or the medium of interaction plays an important role in communicating with the external world. The representation itself has attributes and functions that enable our cognitive functions to use and bridge with the outside world. These attributes define the level of interaction and awareness in the external environment.

In teleoperated environments, commonly operator's body is represented by a robotic form that focuses on the required operational purpose [102], in which the considerations for designing the slave system is driven from the requirement of the scenario it will be used for. The design of the slave system reflects the control mechanism and the type of behavior the operator should follow to drive and operate the slave body, for example, to use a joystick to control the direction or a data glove to match slave manipulator position with hand's posture. For social and communication type of scenarios, the robotic representation is not always a necessity to be considered in the design of the system. For example, in mutual body representation, a visual representation is sufficient to transfer the state of the person located remotely [38], or for a group of people sharing the same space and collaborating remotely [48] in which its possible to transfer important social aspects of communication such as eye gaze and hands gestures. Hybrid representations which both tangible and imagery representation to solve the body and social aspects as well as the physicality to manipulate remote objects were proposed. Physical Telepresence [50] or as commonly known as inFORM [25] uses shape displays for interaction mediation. This type of body representation is mediating the action of the body rather than the actual spatial and physical state of it. From the user point of view, these systems do not provide a full sense of immersion, but rather as being looking through a window or medium which they perform their actions within rather than being the medium.

Virtual representation of the body has enabled to overcome the physical

constraints of the representation. Ogawa et al. [71, 72] proposed a user interface which extends hands interaction by deforming the visual feedback of user's arm, such as extending its length to reach far objects. Asai et al. [3] used this type of interaction for people with a mobility impairment to easily point to physically far surfaces. The virtual representation has also enabled social researchers to investigate the effect of changing body skin color or race to a different one to decrease the racial bias against dark-skinned people [76, 87]. The study highlights the importance of multisensory correlations to induce the illusion of body ownership toward the virtual body, which relates to the level of embodiment the participant's experience. Within the same context, [57] showed that experiencing a body with superpowers (as being a superhero in sci-fi) to fly in a virtual environment based on embodied actions, like acting as Superman to fly, would lead the participants to engage in a pro-social activity to help others in need after finishing the experience. Further investigations on reproducing the illusion of virtual body ownership for non-humanoid creatures has been studied in [53]. Probably the virtual bodies are the true alchemy of embodiment.

2.2   Body Schema & Modalities

Body Schema involves aspects of both brain processing and body peripheral (Proprioception signals, sensory signals), and it is considered as a collection of processes that stores the posture of the body and relationship of body and its representation. When the body state became in motion, the body schema updates unconsciously according to that motion [36]. A mental image maintains a constant representation of the body while a task is being carried on Figure 2.3 . The observations/decision loop which is unconsciously updating our mental representation relies mainly on the information channels that interact with the body. In some sense, its possible to represents the body as being a medium between the internal mental image, and the surrounding environment in which the body is located at, and at which the task is being performed in.

The idea of the body schema and the mental image of the body has been derived from philosophical and empirical perspectives. Aristotle in his book (On The Soul, II) referred to the soul as being the form of the physical body, which perishes together with it at the death. And according to Spinoza (The Ethics, II) said the soul is the *idea* that the body develops of itself. In a modern term, the *idea* is simply a mental representation, more precisely, a self-representation. Gestalt properties, such as body shape, are rather global properties of the entity, and thus the body schema or the self-representation is being a neural mechanism to represent exactly such global properties. In a modern philosophy, Matzinger [66] has presented that the body and self can be a plastic representation and highly context sensitive. He referred to the self-representation with the concept of "Self-model", which is a process driven by a series of cross-modal interactions from the several sensory channels in order to define the image of self and body. For example, in the rubber-hand illusion (RHI) [11], the sensation of being stroked with a prob is integrated with visual and kinesthetic feedback in order for the brain to constructs a matching proprioceptive map between the
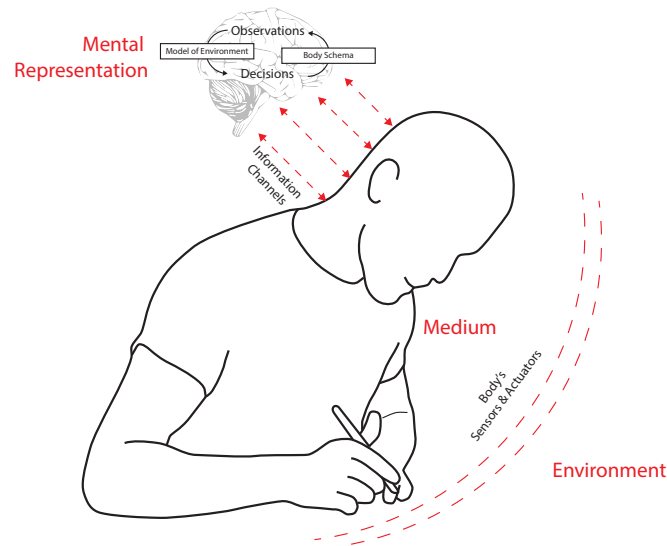
Figure 2.3: Body schema is a constant internal process representing the medium of interaction (body).

subject's mental image of his body and what he perceived from the sensory channels. As a result, the subject's mental image of the body ownership is redefined and would then believe the rubber hand is part of his body. This illusion is extended to modulate the body schema to perceive a third arm as being a part of the body [30].

By considering *the body as a medium for interaction*, then its possible to create a proxy that maps the sensors and actuators of the body into a different representation. In Telexistence, previous research showed the possibility to transfer body awareness using a remotely operated avatar matching operator's body schema [23]. A visual-kinesthetic transfer model is defined to create the sense of bodily presence and ownership toward the remote representation. And fundamental elements for presence, such as vision matching, real-time control and operation, and tactile feedback, are important elements in creating such a schema transfer model. By maintaining such baselines for presence, and defining a top-level model to remap the other modalities, it would become possible to redefine body schema while preserving the sense of ownership toward the alternative representation.

Within the scope of this thesis, Embodiment emphasizes the role of the body in defining and shaping the mind and its cognition through the sensory-motor skills being performed, and the sensory-perceptual inputs being perceived. That is, the body is being the medium and the interface of the mind.

2.2.1 Isolated Modalities

A modality can be considered as a single independent channel of sensory input/output between human and the surrounding environment. As in our daily interaction, we rely on cross-modalities to have a clear comprehension of the tasks we involved in. Human body physical structure imposes tight coupling of these modalities, and our mental representation is mapped to this tight coupling. However this tight coupling can introduce an overlapping information, and some of the sensory inputs are not necessary for the

Figure 2.4: Enhancing perceptual modalities using wearable devices.
Eidos Audio (left), and Eidos Vision (right).

Photo © Eidos (2012).

task being involved with or would result in sensory distraction between the various sensory modalities.

Isolating the involved modalities for a specific task would enhance the operation and achieve a higher sense of *attention* by reducing non-relative sensory and perceptual information (distraction).

> "Everyone knows what attention is. It is taking possession of the mind, in clear and vivid form, of one out of what seems several simultaneously possible objects or trains of thought. Focalization, concentration of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatter-brained state which in French is called distraction, and Zerstreutheit in German." (William James - The Principles of Psychology Vol. 1 1980 p.404)

In his definition of *Attention*, James highlights the role of isolation of a sensory stimulus through the concentration towards it in order to effectively engage in or observe from a certain action.

Abstracting the individual modalities, based on their function and type of information they carry, would assist the design and use of digital tools and media to enhance them individually, and then provide their contents back to the user through feedback devices. Eidos [12] shown in Figure 2.4 ,

illustrates the possibility of controlling our individual senses and adjusting them separately reducing other distracting information. Isolated modalities based design helps enhancing individual modalities functions, or even achieve supernatural capabilities such as augmenting visual perception [47].

2.2.2 Body Ownership

In order to establish that our perception of body and self can be altered via modifications of our sensory modalities and embodied representation, psychological studies and empirical evidence are required to support this claim. Previously we discussed the phenomena of rubber hand illusion which can create a strong sense of ownership to an artificial body based on the cross-modality experience of vision and haptic feedback. That example showed how the perception of body part ownership can be altered through external stimuli.

In a contrasting example of altering the sense of body ownership: the phantom limb sensation [64]. In this phenomena, a person perceives a missing arm or leg as still existing regardless of the fact that it's being not physically present. That is, their mental image of their body still poses the perception of the capability of moving the limb, the proprioception feeling of its posture, but without visual or tactile feedback from the missing limb. This inconsistency of perceptual feedback with the internal mental image of the body causes what is called phantom limb pain. In one of the previous studies on the phantom limb pain treatment, the visual illusion of missing limb was 'resurrected' back to the patient using a mirror setup [80] in which the patent sees his functioning limb in the location of the missing limb. An illusion of kinesthetic sensation emerged during the trials, even when the experimental used his hand instead of the patent hand to induce this sensation.

2.2.3 Modalities Substitution

Brain plasticity and adaptability to modulate sensory feedback and restructure the perceptual awareness has been widely investigated in psychology and neurology. A subset of this area of research is modalities or sensory substitution. This research involves the study of the effective sensory modalities that can be used to replace another one. Thus the research involves the study of meaningful conversion of the information from one modality to another, and the representation used to deliver the newly converted information.

In neurology area of study in sensory substitution systems, a person who lost a certain sensory due to peripheral damage still has the capacity to sense within his brain structure. For example, for a blind person, the peripheral sensory system (retina) is not functional which results the fact of blindness is still has the capacity to "see" from neurological perspective unless if the visual cortex of the brain is damaged. Similarly for those with positional orientation and balance disorders (the vestibular apparatus), or deaf due to sound transduction damage in the ear (the cochlea). Previous work [4, 5] found that the input from a sensory substitution system can stimulate the brain regions which lost its sensory peripheral.
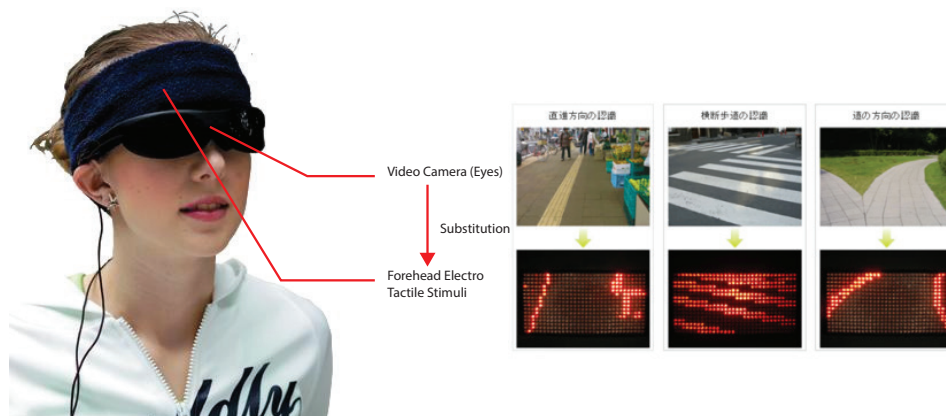
Figure 2.5: Visual sensory substitution using tactile feedback

Photo © AuxDeco, Inc (modified from original)

Sensory substitution devices (SSDs) have been commonly used as means of non-invasive methods to restore a certain sense of the surroundings. SSDs are generally divided into two main variations: visual-to-tactile substitution devices that translates optical information into tactile stimuli, and visual-to-auditory substitution devices that convert sound into an audio wave. Widely used is the tactile based SSDs to restore the visual modality for blind people, and as one of the earliest examples of such tools is the white cane which still being used until now by blind people (recently proposed EyeCane [56] augmenting the traditional white cane). Braille is another commonly used SSD to provide tactile amount of information of written symbols. White et.al [108] developed a tactile SSD that converts the optical images captured by a camera into an input to an array of solenoid vibrators located on the back of the user to induce visual information through tactile sensory channels. Commercial products that use electro-tactile stimuli to substitute the visual modality has been also introduced previously, AuxDeco by EyePlus-Plus Co. Inc [37] uses forehead as input surface for the visual information, while BrainPort by Wicab Inc. [109] uses the tongue. The other spectrum of SSDs used auditory feedback. As one of the very first SSDs is "Ultrasonic Eyeglasses for the Blind" proposed by Leslie Kay [44] which is a biomimicry glasses that uses ultrasonic sensors to measure the distance and the geometry of the objects, and produces auditory tones that vary in frequency depending on the distance. More recently, the vOICe by MetaModel Inc. [65] uses the audio modality as a substitution to the vision by converting visual stimuli into musical notes. Other variations of vision-to-auditory approaches to transform the image data have been widely investigated [14,32,63] which are not only used for visually impaired people, but also as approaches to transform visual information to meaningful auditory information.

2.2.4 Body Morphology

To overcome the rigidity in physical constraints our body schema constitutes, first a digitization of our body's sensory and actuating modalities is needed. The conversion from analog/physical representation into a digital form would provide the required flexibility to remap, augment, and alternate what we observe and how we operate. The process of digitization of the body can be done using sensors to track the different types of actuation body presents, and perceptual proxies devices which produce different types of sensory feedback to the body. As seen in the previous section, the sensory feedback can be modulated or substituted before providing it to the user's body in order to overcome a certain disability or to augment a modality. Such process will be referred to as modality transfer.

Modality transfer can be seen in two types: Local schema transfer when the medium of interaction is spatially located along the body so its modalities are cooperating with the alternative modalities used to shape the body schema. And remote schema transfer when the medium of interaction extends spatially into a different location, thus the cognition relies completely
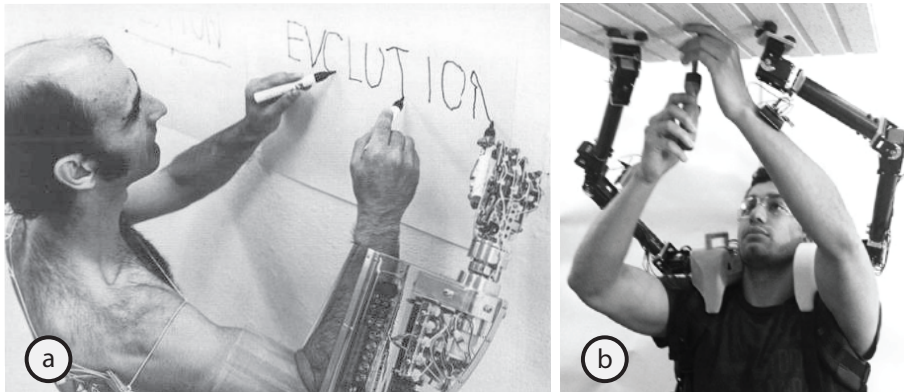
Figure 2.6: Body physical morphology using additional limbs. (a) Stelarc Third Hand motion is driven from body muscles, where (b) MIT's Supernumerary Robotic Limbs (SRL) action are application driven.

Photo © (a) Stelarc, and (b) MIT Federico Parietti.

on the used representation to communicate with the environment.

In local schema transfer, it is possible to induce new body schema when we use tools that, for example, shift the function of the modality in a constant manner. For example, when using a pen for writing, the pen becomes an embodiment of an object with the functionality to point, write, and touch. When we adapt to the use of the tool, it becomes transparent and a part of our new body schema. Robotics has been used also to reshape our functions in an embodied or assistive manner Figure 2.6 . Stelarc Third Hand [40] was one of the first "successful" trials to embody an extra robotic prosthesis to his biological body. The prosthesis was physically strapped into his right arm, but had functions to grasp, pinch, and rotate its wrist. The arm is considered as being embodied because its motion was driven from his bodily functions, that were EMG[1] signals from his legs and abdomen. The EMG signals acted as a real-time digitization of some functions of his body which served to be a transparent channel communicating with the prostheses. This bodily driven control allowed him after many years of use to adapt this arm to his body schema, and to operate it as an extension of his body [97]. On the other hand, other types of tools are considered as assistive rather than being embodied with the human body, such an example is the Supernumerary Robotics (SR) [10, 51, 74, 110]. Such type of SR systems are designed to work side by side with a human operator in a coordinated manner, it has an independent behavior when performing the tasks to optimize the operation which the human is performing through independent observations and reactions. Thus its motion is not driven by direct bodily actions, but rather from the task's context, which means the arms could be considered as a separate entity disembodied from the human operator.

For remote schema transfer, the process of body schema adaptation requires the modalities to be transferred from one location to another by digitizing their functions so its possible to capture and reproduce them from one

---

[1]Electromyography which records muscle activities using electrodes placed on a relevant skin surface.
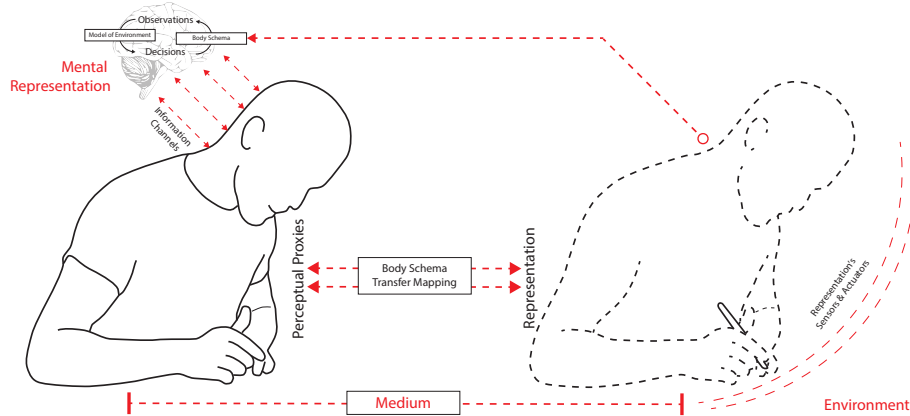
Figure 2.7: An overview of the process of digitizing human body perception, and extending the medium of interaction (human body) into an alternated representation. Body modalities are transferred into an alternative representation.

side to another. Figure 2.7 provides an overview of the process of digitizing body schema, and transferring it from an alternative representation into the body via perceptual proxies. Body schema transfer mapping block acts as a definition for the process of real-time updating the representation with the human body using the digitized sensory channels. Work in telepresence and telexistence derives similar approach for creating a sense of presence in a remote environment for teleoperation purpose. These approaches, as will be discussed in the next section, transfer the body modalities from one location to another to stimulate the sense of remote presence.

In conclusion, in order to alternate body schema, there should be a consistency of sensory inputs regardless of the physicality of the observed body or the representation, thus maintaining the sense of ownership. Also, to preserve the sense of agency towards the body, the observations from the body (e.g. movement) should be correlated with the actions of the user.

2.3 Presence Augmentation & alteration

In this section, discussion about related work in body presence augmentation and alteration. Work in Telepresence, Telexistence, and spatial awareness are described showing the impact of such means to expand body's limited perception and actions.

2.3.1 The Essence of Presence

Presence is the state of awareness of body and its surroundings. Enhancing the state of presence or altering it is to change the physical constraints or to convince the mind of being at a different state than what the body physically is. Kim [46] highlights on that as "feeling of being a part of the phenomenal environment created by television and not being a part of the physical environment surrounding the viewer and the television set". Similarly, with the use of virtual reality tools, it's possible to lead the mind into a "suspension of disbelief that they are in a world other than where their real bodies are located" [96]. The state of presence is a flexible term that can stretch over time and space, as well as the form of body perceiving the various sensory modalities which develop the understanding of where and how the body is located at.

In physically teleoperated environments, several work and research discussed the requirements to provide a physical presence to operate and to be embodied (as being part of the remote location). Minsky [68] proposed telepresence systems that function to provide "feeling like you are actually 'there' at the remote site of operation". Minsky highlighted the role of teleoperated systems to empower the functions of human operation and labor to perform an effective operation for the tasks at hand. The essence of providing the sense of transferring body awareness to remote slave system was proposed by Tachi in his paper on Telexistence systems [101]. "a system that at the remote control-site a human operator can perform remote manipulation tasks dexterously with the feeling that s/he exists in the slave

36

anthropomorphic robot in the remote environment." which is referred to as Tele-existence system (or Telexistence) under the condition of "real-time sensation of remote presence". This definition provided a requirement of the subjective experience of body presence "...perform remote manipulation tasks dexterously with the feeling that s/he exists in the slave...". The requirement of mapping and matching the state of the body between the human operator and the slave system creates a strong sensation of agency and ownership. Rheingold [81] described Telexistence systems as a "form of out-of-the-body experience" once he looked at himself through the slave system. Telexistece systems can be viewed as a way of presence augmentation since the body's physical constraints are extended over distance, and the effect of manipulation and perceiving can be done remotely. The usage of mediated technologies can help to reproduce the sense of presence, and thus can be a mean to enhance and alter this sensation.

State of presence can also be stretched to non-physical environments. Virtual presence is usually regarded to the state of being in a non-tangible space which is being represented by digital means. Sheridan [94] created the distinction between physical and virtual, in which physical presence is characterized as "physically being there" while virtual presence as "feeling like you are present in the environment generated by the computer". The master/slave system of Telexistence also proposed to operate under virtual environments as if being in a physical one [99], extending the concept of presence into virtual, non-physical spaces.

2.3.2 Connected Awareness

The story of "Waldo"

"Waldo", a fictional novel by Robert A. Heinlein 1945 [35], tells a story of a mechanical genius who lived his entire life in a self-imposed exile: his own body. Struck with a neuromuscular disease (myasthenia gravis) in which it locked him inside the shell of his weak body, he was capable to develop dozen of mechanical hands (called as Waldo F. Jones Synchronous Reduplicating Pantograph) which are operated via sensory devices attached to his hands, and extend his body and enable him to amplify his weakened muscular strength. These mechanical hands were depicted as in various sizes which can be as small as micrometers in span allowing him to micro-manipulate, and all the way up to six meters from finger to thumb in size allowing him construct buildings and facilities while imitating his hands motion and actions. Despite of the fact of being crippled in his weakened body, Waldo considered himself as a superior to humanity since his augmented skills and his body extensions enabled and empowered him to do far more superior tasks than what a fully functional human can do.

The story is considered as one of the very first original novels which highlight the effect of technology and tele-body-representations to augment the motor skills of our bodies. Also, it highlights the effective adaptation of connected bodies and awareness to amplify our cognitive skills and enhance our physical capabilities.

Thanks to the never-ending expansion of interconnected networks and Internet, the boundaries of location and space has expanded from local to a global scale and created a true ubiquitous information accessibility. Mark

Wiser [107] introduction to Ubiquitous Computing helped to create a new paradigm of information accessibility and networked computational design. These technical innovations in network media and information accessibility opened a wide door for the emergence of digital-based Teleconferencing and Teleoperation tools at a wider scale compared to the older analog-based media transmission. These tools provided the capability of transporting human beings virtually from one place to another to get involved in a remote discussion, or to transport his skills for solving remote tasks.
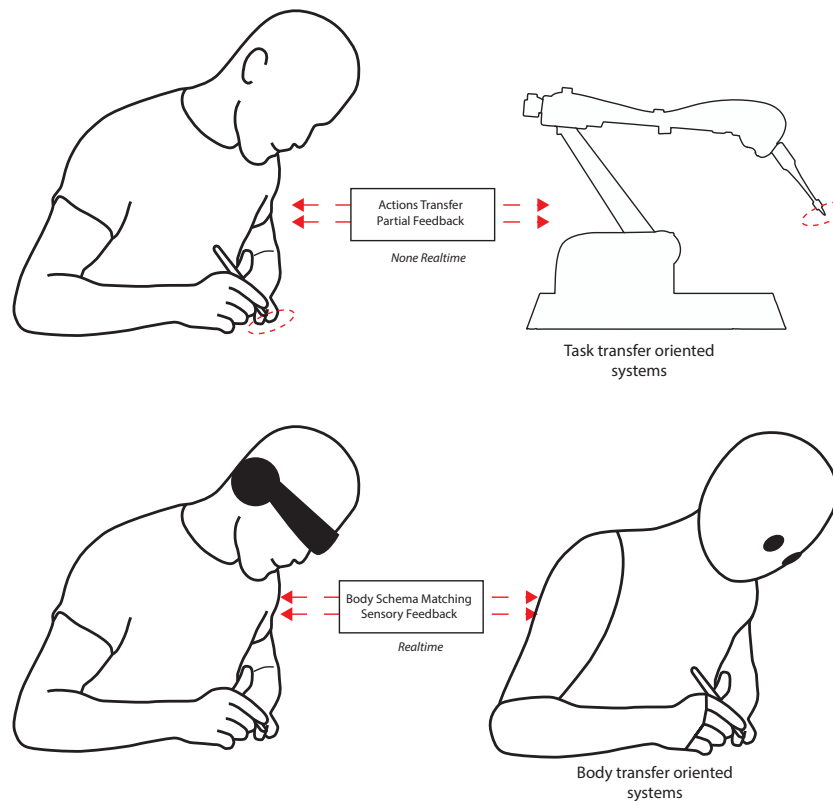
Figure 2.8: Task vs Body transfer oriented systems.

Basic teleconferencing tools, such as Polycom video conferencing [85], Skype, and Google Hangouts provide an easy to use means to connect remotely, however they lack to represent the physical attributes of the operator, and rather have limited number of modalities that are engaged (commonly video/audio feedback). The other paradigm aside from the basic

teleconferencing tools could be categorized to Telepresence and Telexistence technologies Figure 2.9 . Each of these addresses different type of cognitive transfer, the prior is task transfer and the latter is body transfer.

Telepresence systems emerged to the market and gave higher levels of flexibility for operation and representation, such as MeBot [2], Kubi [84], and Double Robotics [83]. These systems would allow motion and teleoperation and even can transfer non-verbal cues via robotics hands, or facial expressions [69]. Telepresence systems, in general, can be referred as task transfer-oriented systems, which focuses on the use cases of the system instead of the full embodiment of the operator.
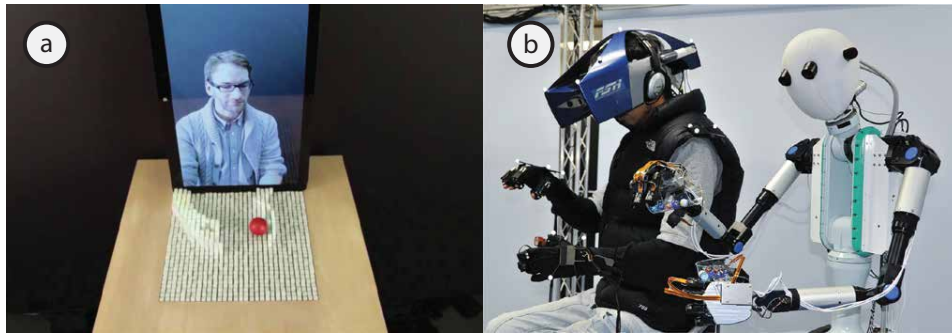
Figure 2.9: Embodying teleoperation technologies (A) inFORM as an example of Telepresence systems, and (B) Telesar 5 as an example of Telexistence systems.

Photo © (a) MIT Tangible Media Group, and (b) Tachilab.

In Telexistence related applications, an avatar that represents a substitute representation of human body using robotic systems and digital forms, the human operator can remotely access and have immediate control and real-time feedback sensation from the remote surroundings as if the operator is actually located there. Although current Telexistence systems [24, 34, 43, 100, 101, 105] provide a partial representation of the human body, however, they achieve a strong sense of self-consciousness and sense of presence in the remote place, which Telepresence systems lack.

In some sense, we can say that both Telepresence and Telexistence systems are a one-to-one mapping of the user to his avatar (best case when a full

body representation is achieved). Based on this definition, although avatar representation provides enhancements to the human operator and overcomes several difficulties that direct human operation cannot serve, these methods of direct one-to-one mapping has upper bounds of what an augmented human can do. The question is, what happens if a physically disabled person tried to use such systems? As its a direct one-to-one mapping, the disability will also arise in the representation too. In Waldo's case, his body was weak but not disabled, and his hands were still being represented as direct one-to-one mapping with his waldos. Without his basic motor skills, he would not be able to operate the body amplifiers he invented.

2.3.3 Presence Embodiment

Media interfaces are progressively growing to embody us within it, affecting our sensations of physical presence, social presence, and self-presence inside. This progression in telecommunication technologies creates a tighter coupling between the body and the interface, in which the body becomes a part of the physical space and the cyberspace. This progression creates an adaptation between the body and the interface [8]. Adding to that, research in Virtual Reality and cognition showed that sense of presence can be provided through technical means and psychological experiences [82]. Technological means are the physical representation of sensory organs, such as stereo vision (eyes), binaural audio array (ears), tactile system (touch), ..etc. The psychological experiences define the relationship between the technological means and the human body. Thus by altering both the means and the experience, sense of presence can be altered or enhanced, while maintaining the seamless experience of presence, or of being there.

2.3.4 Beyond Presence

The question of the possibility to leverage what our bodies can perform, or expand the perceptual capabilities of our visual, auditory, and tactile modalities has been a focus in human augmentation research area.

Recent systems were introduced to share the presence, and can be sort of extension of telepresence systems. Kasahara [42] proposed a JackIn system to allow vision sharing, thus enhancing participants bodily and visual presence into being remotely presented. An extension of this work [41] is to share vision into multiple locations, thus multi-casting the presence. Cohen [16,17] proposed to use an architecture of sources and sinks to direct the auditory feedback from multiple sources into a single observer. Although this architecture does not define a general way of mixing or layering information from these multiple sources, it proposes a framework for multi-sensory feedback systems and can be viewed as a way of presence alteration.

Recent research builds on these tools in order to enhance the presence and expand body's perception. Fan et. al [22] proposed a video-based image blending using two-way see-through HMD that provides the user an awareness of both locations behind and front of him by blending the visuals of both views. This type of augmentation can be viewed as remapping vision modality into multiple channels of input. Suzuki et. al [98] used pre-recorded 360 visuals and blend it using see-through HMD by phasing the visuals in and out through a specific scenario resulting the user experiencing past as being present, and thus altering the sense of presence for the user.

2.4 Meta-Modeling Design

In this thesis, a design approach is proposed in Chapter 3 for Embodied-Driven Design to define the body transfer and schema. This section provides a brief overview of the related work and research that uses meta level modeling approaches in various design fields, and gives the necessity of adapting relevant approaches for the proposed meta-modeling environment.

The idea of using tools such as diagrams for abstraction and modeling has been a commonly adopted approach to provide a middle ground between engineers and designers, and as tools for quick prototyping and re-usability of previous models. This can be attributed to the simplification of the target system or interaction, avoiding the designer to be involved in technical

details thus attracting a wider audience to adapt to the tools. The original idea of visual programming was proposed by [95] to address the previous concerns. Max/MSP [78] and Pure-data [79] uses visual programming approach for modeling, and they are widely used by designers who do not have programming skills, like audio or visual designers, accelerating their prototyping and testing process. These approaches have also been adapted for educational purposes due to the easiness, symbolic representation, and immediate results when performing actions within the model. Scratch [59], LittleBits [6], and MakerWear [45] used similar approaches for children oriented workshops using tangible or on-screen charts.

2.5 Impact of the Proposed Research

By the time of writing this thesis, there has been no known method or tool which provides the possibility to redefine the function of human body, or rewiring the body functions with external tools. The work established in the area of human augmentation has been mainly function driven, that is a bottom-up approach to the design of the systems.

Embodied-Driven Design builds on the research area described previously, and it provides a unified framework for defining and forming a systematic representation of the human body and its representation. Body representation and the way its mapped to operator's body plays an important role to maintain or enhance the presence experience and thus the operation quality. Embodied-Driven Design addresses this point by providing body schema transfer blocks into physical/virtual or hybrid representations. Physical representation blocks provide essential modalities for remote presence, such as visual/auditory feedback, physical interaction, and so on. Virtual representation blocks provide the capability to deform body visuals remotely to enable non-physical interactions, or to substitute available physical representations remotely, such as hands representation.

II

Embodied-Driven Design

"Our Bodies are the Medium of Interaction,
the Bridge of our Minds"

Chapter 3

Concept & Design Strategy

"one sees the environment not just with the eyes, but with the eyes in the head on the shoulders of a body that gets about. We look at details with the eyes, but we also look around with the mobile head, and we go-and-look with the mobile body"

JAMES J. GIBSON, THE ECOLOGICAL APPROACH TO VISUAL PERCEPTION 1979

This chapter will address the design aspects and process for body driven applications by first providing a systematic model of human body schema, and then deriving topological models for the postural and sensory body schema. The idea of decoupling and re-coupling is also discussed here as a part of LCM design pattern which allows the reconfiguration of body schema mapping by using a topology of the various body modalities. And finally, EDD framework is described along with the meta-modeling tools and modular blocks that reflects LCM design pattern. Three meta-modeling design categories are discussed that summarizes the major use-cases and design scenarios that EDD can serve. To facilitate the design process, an embodiment toolkit is proposed which can be used along with EDD framework to prototype and reflects the meta-models into a physical form.

3.1 Engineering Body Schema

In order to create a system which supports the ability to modulate our sense of embodiment into a new representation, the topic of body schema is important to fundamentally analyzed and to design a systematic model that reflects the properties of the body schema.

As described in Chapter 1, body schema is an internal/mental representation of the medium which embodies the sensory and decisions of one's self. The medium of representation holds both the kinesthetic or postural model that reflects the position of body's joints and structure [36], and the sensory or surface model that captures the information from the surrounding [54]. The body schema can be considered as a collection of processes which registers the postural model and validates through capturing information from the sensory model [31,61]. The body schema can be attributed as being plastic, that it can be altered and modified. As an example, the use of external tools effectively alter the mental structure of the postural model and thus reflecting a modified version of the body schema from the one originally maintained, incorporating the spatial and dynamic properties of the tools [7, 15, 60].

The body schema can be characterized by the following features [70]:

- Spatially Encoded: the body schema maintains a three-dimensional postural memory of representation's parts.

- Distributed and Modular: it uses a process of interaction between several modules (perceptual and kinesthetic parts).

- Inter-modality and Super-modality: the internal schema is based on the sensory integral of proprioceptive and exteroceptive information to maintain the mental representation of the body. This representation can be modulated by the sensory channels, such as vision, based on the task and the environment which the body is located at. This sensory integration has the responsibility to define a modal representation which maintaining the coherence of task operation [31].
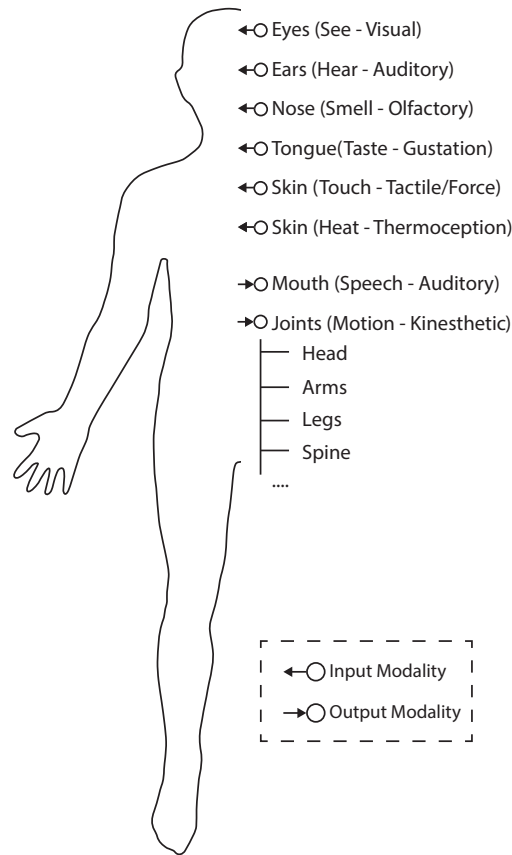
Figure 3.1: Human Body Schema Input/Output Modalities Map

- Plasticity: body schema maintains a short-term alteration, which is measured by the length of operation [60].

These characteristics help to derive a systematic model of human body schema, in which the structure consists of modularized subcomponents or, as defined herein this thesis, as **Modalities**. A modality is a single component of body schema that contains an information channel as input or output which interacts with the body representation and surrounding space. Figure 3.1 highlights the various modalities human body schema can be characterized with. Input modalities refer to the sensory model of human body, and output schema is related to postural model (with an exception to the speech modality).

In order to engineer body schema, and create a transfer model which alter the sense of body representation and how we interact with it, we can decouple the body schema representation into a set of modalities. In this thesis, the process of decoupling is called as "Loosely-Coupled Modalities".

3.2 Loosely-Coupled Modalities

As in our daily interaction, we rely on cross-modalities to have a clear comprehension of the tasks involved in. Human body physical structure imposes tight coupling of these modalities, and our mental representation is mapped to this tight coupling. Ability to isolate and separate these modalities can help to redesign new sort of perceptual augmentation interactions by modulating the information channels from the representation and feed them back to the human body through perceptual proxy devices (such as HMD, haptic interfaces, ... etc).

Modality	Data Representation	Capture Device	Representation
Vision	Imagery	Cameras	HMD
Hearing	Auditory	Michrophone	Speaker
Speech	Auditory	Michrophone	Speaker
Kinematic	Spatial	Motion Tracking	Servo Motors
Touch	Vibratory Signal	Force/Tactile Sensors	Haptic Display

Table 3.1: Examples of human body modalities & corresponding representations

The idea of Loosely-Coupled Modalities (LCM) resides on treating each Input/Output modality as a separate channel of interaction. LCM simplifies the body schema structure into primitive modalities, such as joints that can be represented by position and orientation, visual input for each eye, auditory input and output for the ears and the mouth, ...etc. The modalities have a predefined information type that only accepts, for example, the

joints only accept spatial information but not an audio signal. Table 3.1 shows various modalities with the corresponding data type which accepts. A channel accessor (symbolized by a circle and an arrow) represents the information flow of the modality. Two types of accessors are defined: input and output. An input accessor directs the information flow from the representation into the body through perceptual proxies, and the output directs the internal state of the body to the representation through tracking and sensing devices. The information flow is defined by a digital form, which is sampled through capture devices, and reproduced through perceptual proxies. Table 3.1 lists the modalities with the corresponding representations and information type that can be used to define them.

In this thesis, the body schema modalities are categorized into two types of modalities:

- Postural modalities: focuses on the spatial, three-dimensional attributes related to the body and its representation. Mainly consisting of the joints and shape of the body.

- Sensory modalities: contains the various sensory attributes of the human body.

Before describing each set of modalities categorizes, the concept of decoupling and re-coupling of modalities is explained. An abstraction model of the postural and sensory modalities are defined and referred to as Topology models of human body.

A *Topology* was originally derived in mathematics from the fields of geometry and set theory which defines spatial relationships that connects the matter being studied regardless of the dimensional differences (such as length, scale, ...etc) or descriptive properties (such as color, weight, ..etc). For human bodies, two bodies would have the same topology as long as they have the same number of limbs, eyes, ears, ...etc. If a body has experienced certain deformations (such as a lost limb), then its corresponding postural or sensory topology has also been changed.

3.2.1 Postural Modalities Topology

In order to create a model defining the body structure, a hierarchical model describing the distribution of the joints in the human body is defined. Figure 3.2 provides an approximation of the body joints topology. In this model, the following region categories are used:

- **Spine Region**: Contains the vertebral column joints which control the overall orientation of the body. The human spine structure consists of 33 vertebrae joints [20], the first 9 joints (from the bottom) are fused into two regions: Coccyx (4 vertebrae), and Sacrum (5 vertebrae). The rest of joints contributes mainly in the motion. Instead of modeling the 24 joints, they were categorized into three regions:

Body Joints Topology Approximation
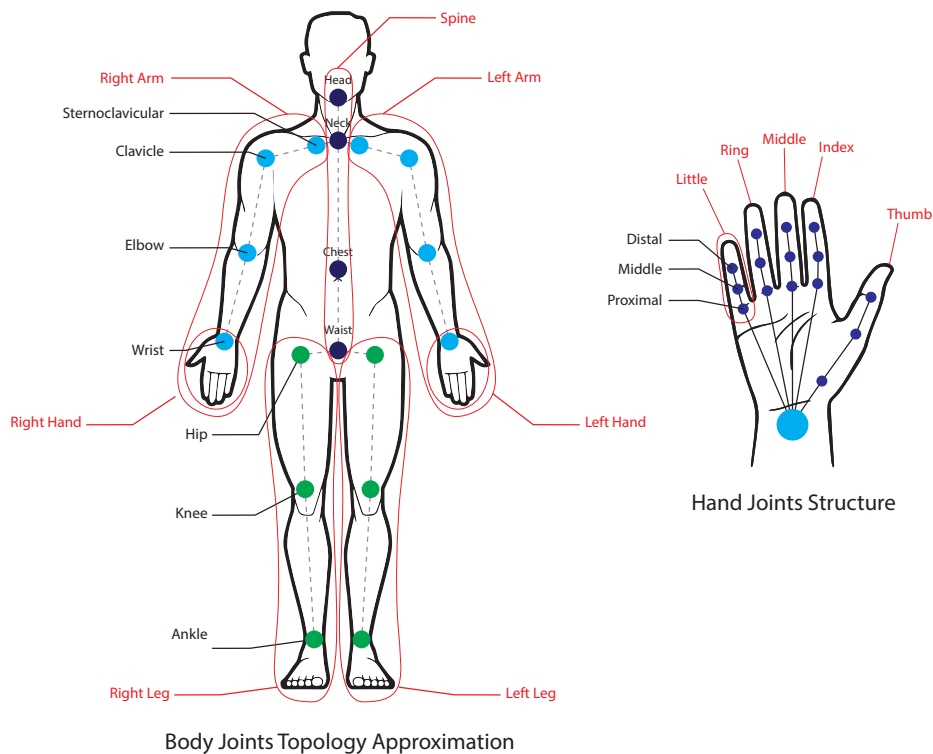
Figure 3.2: Body joints topology approximation diagram, and hand joints structure.

Lumbar spine (5 vertebrae) mapped into the Waist, Thoracic spine (12 vertebrae) mapped into the Chest, and Cervical spine (7 vertebrae) mapped into the Neck. Lastly, to provide an individual control to the head, an extra joint to the structure is added as Head.

- **Arm Region**: Has the structure of the human arm, and it consists of four joints in total to define the motion of the arm: Sternoclavicular joint, Clavicle joint, Elbow joint, and Wrist joint. The arm also contains a sub-region defining the hand.

- **Hand Region**: A sub-region which belongs to the Arm region. Contains the definition of hand joints structure. The hand region is defined by 5 fingers, each finger has the following structure:

    - Proximal Joint: the base joint of the finger.
    - Middle Joint: the second joint of the finger.
    - Distal Joint: tip joint of the finger.

- **Leg Region**: Defines a general structure of the human leg, and consists of three joints: Hip joint, Knee joint, and Ankle joint.

Using the previous categories, it's possible to define a postural body schema as: Spine, Left Arm, Right Arm, Left Leg, and Right Leg. In total, the number of joints defined by this approximation is:

$$
\begin{aligned}
Joints\ count &= Spine + 2 \times (Arm + Hand) + 2 \times (Leg) \\
&= 4 + 2 \times (4 + 15) + 2 \times 3 \\
&= 48\ joints
\end{aligned}
\tag{3.1}
$$

The joint, which is considered as the minimal, atomic representation for postural modality, is defined using two attributes:

- Position: 3 axis (X,Y,Z) tuple defining joint's spatial position.

- Orientation: a quaternion defining the direction of the joint.

3.2.2 Sensory Modalities Topology

For the perceptual schema of the body, major sensory modalities are characterized based on their function and data type. Figure 3.3 highlights four different types of perceptual modalities with their corresponding data they accept. The sensory modalities considered in this thesis are as follow:

- **Eyes**: provide the visual feedback to the user. Consists of two channels: left and right eye. Data type they accept is 2D image data.

- **Ears**: provide auditory feedback, and consists of two channels: left and right ear. Corresponding data type is audio signal (sampling rate 44100Hz or higher).

- **Mouth**: captures the speech audio of the user. A single data channel with data type of audio signal (sampling rate 44100Hz or higher).

- **Skin**: provides tactile and force feedback. Number of channels is defined based on the model (spatial resolution). Accepted data type is a vibrotactile signal.
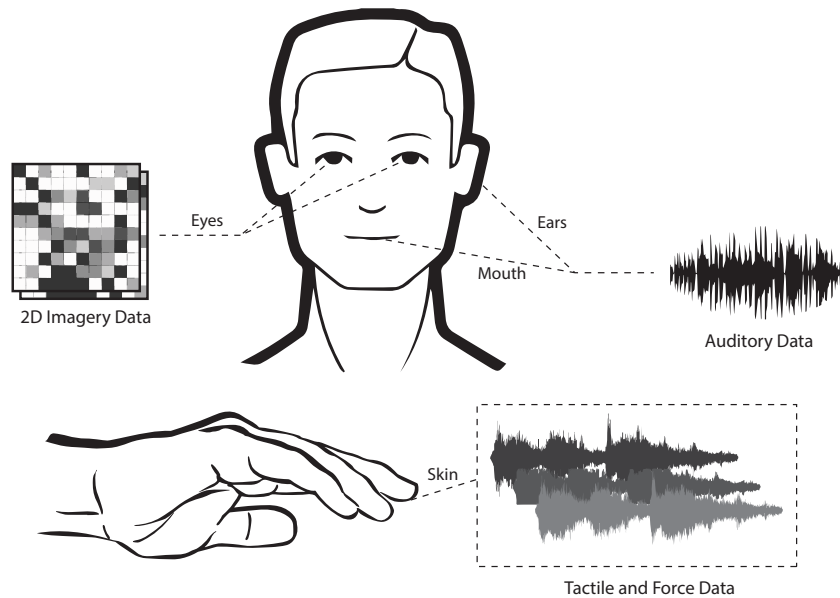
Figure 3.3: Sensory modalities categorization with their data type.

These modules provide the sensory mapping for the human body. The use of digital data as input/output helps to transmit and alter the sensory information using computer mediated communication between the human operator and his representation(s).

3.2.3 Decoupling and Re-coupling of Modalities

In order to simplify the process of modeling the interactions of body schema, the various modalities described previously should use a hierarchical structure to access a region of modalities or the individual atomic level of each sub-modality.
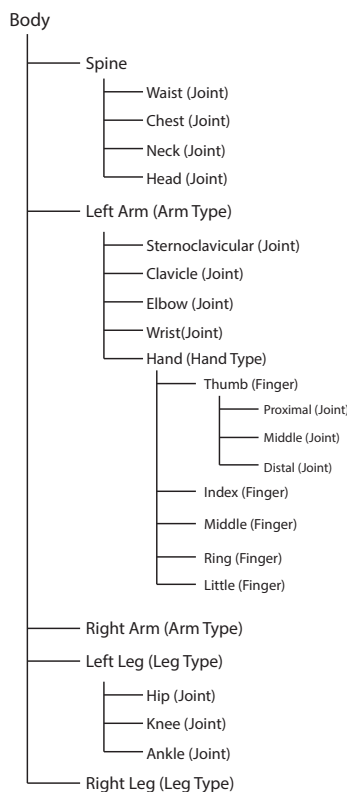
Figure 3.4: Postural schema hierarchy

The use of decoupling/re-coupling in postural schema modeling, it is possible to work at any level of the body structure to map or control the motion of the regions or individual joints Figure 3.4 . For example, its possible to map one arm to another directly, which results to mapping the individual joints of one arm to the corresponding joints of the other joint. Or it is possible to decouple the region into its atomic structure (joints), and reassign them to a different joint, for example controlling the thigh joint using wrist joint. Similarly, for the sensory schema modeling, the stereo eyes can be decoupled into individual left and right eye which are considered the atomic level of visual modalities, and mapped to different sources. The ears can be decoupled to left and right ear, and processed individually.

53

3.3 Body Schema Desgin

3.3.1 Model Oriented

Embodied Driven Design (EDD) aims to provide a systematic, designer-oriented, human driven body schema design and mapping. By using the idea of loosely-coupled modalities described in the previous section, body schema can be remodeled based on an alternative representation. EDD is based on two design models which defines the target application:

1. **Representation:** The representation of the modalities using physical and virtual means, such as robotic avatar, virtual body, non-humanoid representation, ...etc.

2. **Meta-Model:** An abstraction of the relationships between the modalities of human body and their corresponding representations. The meta-model defines a transfer function of the schema that can be direct channel mapping, or alternates the information flow of the modalities with their representations.

The overlapping of these two regions defines EDD framework Figure 3.5 . Using such tools of body and representation abstraction, its possible to
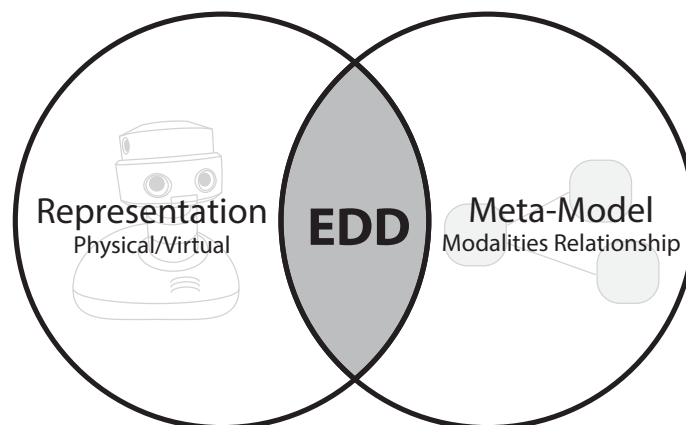
Figure 3.5: Embodied-Driven Design framework areas.
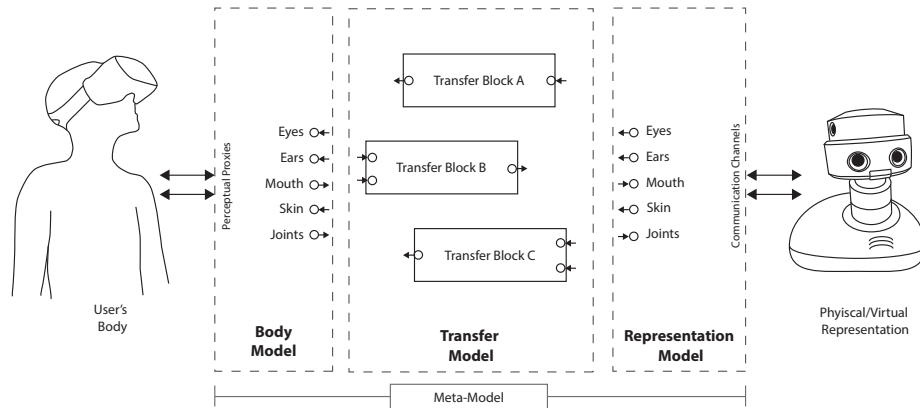
Figure 3.6: An overview of a general EDD meta-model.

design new body mapping. A meta-model defines the channel flow between the various modalities that represents the user body in relationship to a representation. As shown in Figure 3.6 , the meta-model consists of three sets of schematic models:

- **Perceptual Model**: provides an abstraction of user's body and reflects its state into the meta-model. Contains blocks that interacts with the user's body modalities via perceptual proxies devices.

- **Transfer Model**: an intermediate model residing between the perceptual and representation models. Contains various set of blocks that manipulates the information flow of the used modalities.

- **Representation Model**: an abstraction of the representation used (physical or virtual). The blocks contained in this model exchange the information flow from/to the perceptual and transfer models with the used representation. Set of communication channels (e.g. network) exchange the information flow between the model and the actual representation.

Using these three types of models, its possible to define the flow of modalities information of the system. If the transfer model was not used, then a direct schema mapping between user's body and an avatar representation can be defined. As an example of direct schema mapping, Telex-

55
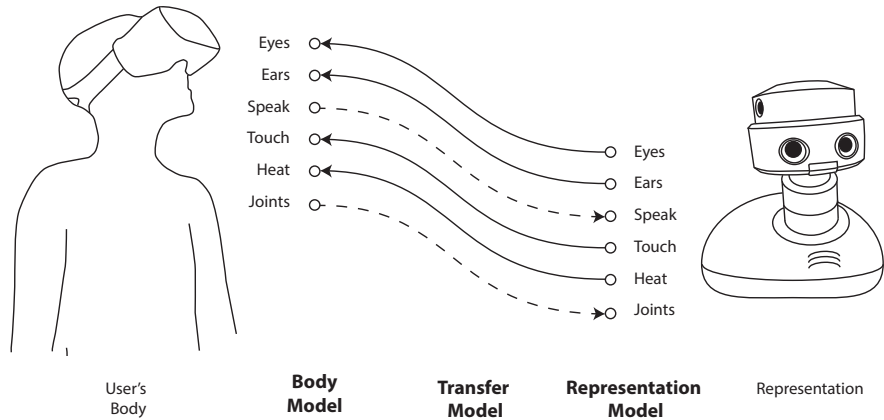
Figure 3.7: An example of EDD for direct one-to-one Telexistence mapping.

istence systems case shown in Figure 3.7 reflects one-to-one sensory and postural mapping configuration. Since the representation robot offers exact body schema modalities to the user, thus no intermediate transfer blocks are required and the flow between user/representation is direct.

3.3.2 Deriving LCM: Modular Blocks

In order to reflect LCM design pattern of body schema in EDD, each or the three models defines its own set of blocks that handles separate types of modalities and information flow. The following blocks categories define the overall formats that can be used within EDD modeling:

- **Perceptual blocks:** driven from the user postural and sensory modalities such as vision, auditory, tactile, motion, ...etc. Inputs/Outputs of these blocks are used with the perceptual proxy devices (HMD, motion trackers, ...etc).

- **Transfer blocks:** acts as a function for the modalities, in which they transform or alter the data flow between the perceptual and representational blocks. Using these blocks, its possible to alter the

limits of certain modalities, augment perceptual modalities, ...etc.

- **Representation blocks:** representation dependent blocks, defines the interface and data flow between the used representation (physical, virtual, or both) and the rest of the model.

3.3.3 EDD Knowledge Areas

In this proposal, four areas of knowledge (illustrated in Figure 3.8) collaboratively contribute to the design process of EDD meta-models, and described as follow:

- **Body Engineer:** a person who uses or develops perceptual proxies and focuses on the experience provided to the user (feedback, tracking method, ..etc). This role is responsible to define and create the perceptual blocks needed in the body model side.

- **Representation Engineer:** this role is responsible on creating the representation blocks used for the target body schema, such as a robotic or virtual representation.
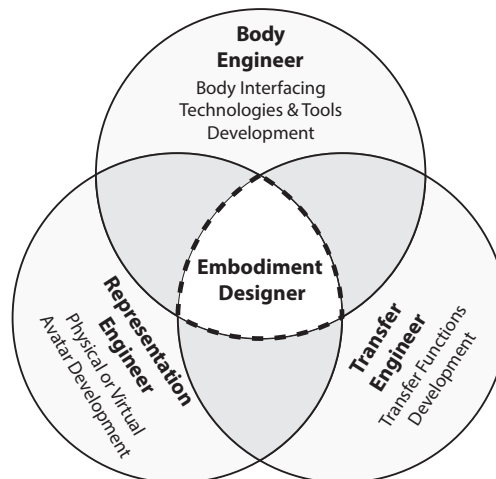
Figure 3.8: EDD Areas of Knowledge

- **Transfer Engineer:** a role to define and create transfer blocks used
  to alter the representation or the perceptual blocks. The transfer
  blocks defined should similar data representation for the other modal-
  ities used in the model.

- **Embodiment Designer:** this person creates the overall experience,
  that is, to use the three types of blocks to define a new body schema.
  The embodiment designer has an overall knowledge of the percep-
  tual,representation, and transfer blocks to be used. The result is a
  meta-model designed based on the specific application.

In respect of each of the three engineering roles listed above, the meta-
modeling process can be more dependent on one role than the other. In
the following section, Meta-Models Categories, further description of the
use-cases and scenarios in which EDD modeling process contributes into.

3.4   Meta-Models Categories

Designing a presence or embodied experience requires defining the scope
of the application needed, and the corresponding representation and percep-
tual proxies used. Representation design reflects the modalities which the
used representation model affords. The representation exposes these modal-
ities as a set of blocks that can be used in the model. The meta-model defines
the connection between these modalities, and user's perceptual modalities
through a mediated transfer model.

Four categories are described here that summarizes the general usabil-
ity of EDD meta-modeling for overcoming certain physical limitations, or
body limitation using a representation. These meta-models describe design
patterns for body mapping and configuration, and can be used individually
or combined within the same meta-model.

3.4.1 Direct Body Transfer

In this meta-modeling category, the model directly maps body modalities with the corresponding representation's modalities. Figure 3.9 shows an overview of a meta-model mapping. The modalities can be directly connected to the corresponding ones, or can a transfer block could be used to alter the data (for example visual enhancement blocks, audio filter blocks, motion gain blocks, ...etc). This meta-model is generally used for cases in which there is no need to change or alter the topology of the body, and the required application requires minimum adaptation time to the representation. As an example, Telexistence systems use direct body transfer so the user can immediately use the representation (or the avatar).

This meta-model is commonly used along the other meta-models to maintain certain modality mapping without alteration.
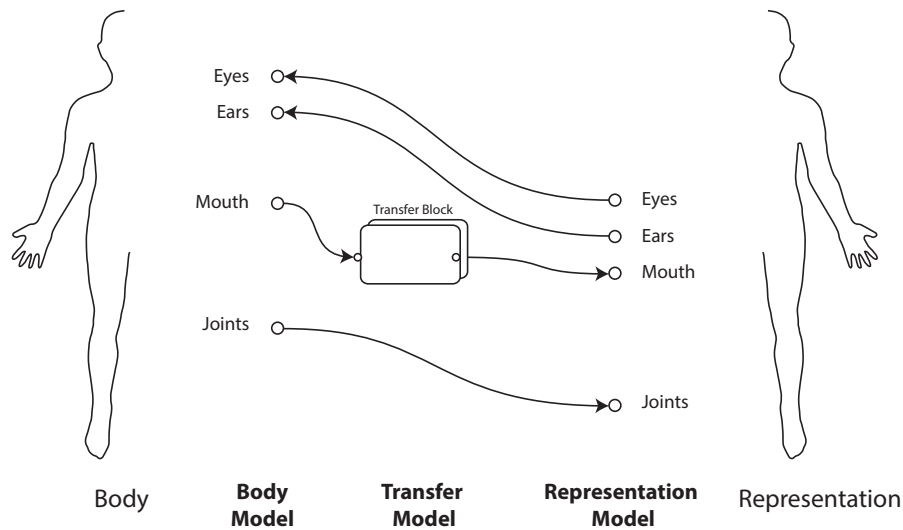
Figure 3.9: Direct body transfer meta-modeling.

3.4.2 Representation alteration

Our bodies are physically limited, and our reaching space is predetermined by the length of the arms for example. Same case is reflected when

operating a remote avatar with physical arms, the same physical limitations are presented. In contrast, virtual reality can easily overcome such limitations since objects in these environments are not represented using tangible materials, but instead, a digital body representation can be used.

Representation alteration design category focuses on redefining the physical representation which can have different properties than what our bodies can provide using a hybrid virtual/physical model of human body. In this design approach, only certain modalities are physically represented (such as vision and auditory modalities) depending on the application, and other modalities are virtually defined. For example, providing a different body visuals or size than the actual.
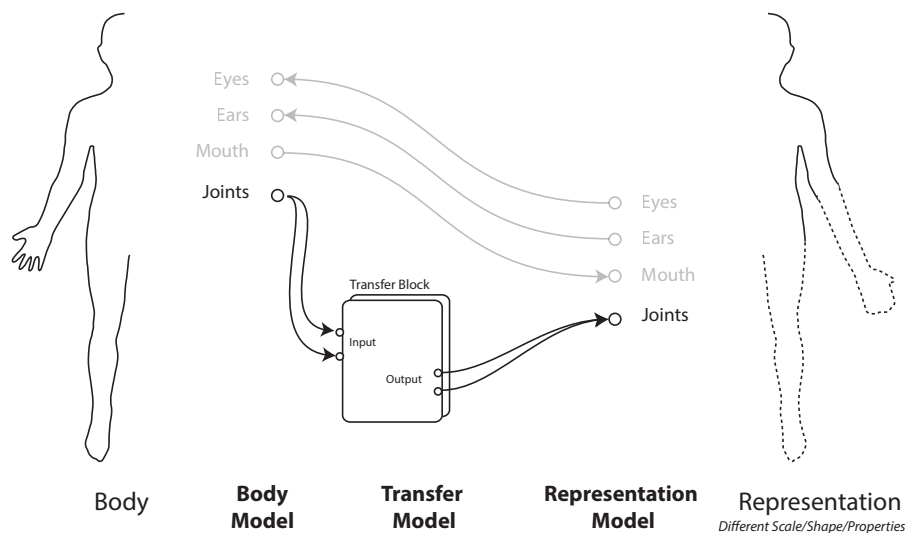
Figure 3.10: Representation alteration meta-modeling.

To visualize the representation alteration, Figure 3.10 shows an overview of this meta-model. The representation has different attributes than the human body has, and can be represented by a hybrid physical/virtual body. For example, the representation can have different arm length than the body has, or dynamically changing its length. The mapping between the body and the representation is altered using transfer blocks that, for example, calculates the required trajectory of the representation body, and maps into

the new limbs. Although the body shape changes, its topology is maintained without any alteration.

In summary of Representation alteration modeling category, the design process for this meta-model is mainly derived from the body limitations. By altering the representation, and reflecting this alteration into a set of meta-modeling blocks, it's possible to match them with the modalities provided by the perceptual model.

3.4.3 Topology Reconfiguration

When physical limitation prevents certain modalities to be used, or when a new bodily function that is not driven from the human body in the representation requires matching with operator body, then the design process is focused on the body topology side. Topology Reconfiguration design category focuses on altering the postural modality data flow from the user side and mapping it into the intended representation. This design category can be used to enable physically impaired users or people with lost limbs to alter their body schema and use different functional limb as a substitution
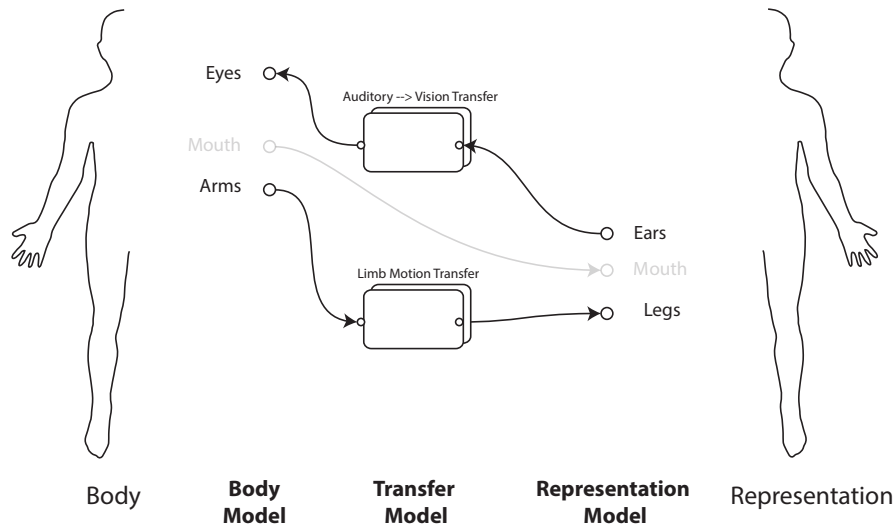
Figure 3.11: Topology reconfiguration meta-modeling.

for a modality. For example, a mobility impaired person (cannot walk), can use his head motion as an alternative method in order to navigate using the remote representation. This would provide direct and embodied operation, and would require a minimum amount of training to use.

To illustrate this meta-model design pattern, Figure 3.11 shows an example of reconfiguring body topology (postural and sensory topologies) by redirecting the mapping of the body's modalities into different modalities in the representation. To achieve such mapping, transfer blocks are required to translate the information flow between such modalities. For example, sensory substitution can be considered as topology reconfiguration, in which one sensory organ is mapped into another by converting the source information into the target modality.

There are considerations when topology reconfiguration is intended to be used. Certain modalities should not be arbitrarily mapped to a different one. For example, mapping wrist tilts rotational axis into panning axis of the head motion. This type of modulation would cause disembodiment although the motion is driven by the wrist. To avoid such disembodiment cases, its possible to define set of constraints on the joints channel mapping:

- **Rotational mapping**: Joint axes should be mapped as:

$$DestJ_{tilt} = F_A(SrcJ_{tilt})$$
$$DestJ_{pan} = F_B(SrcJ_{pan}) \qquad (3.2)$$
$$DestJ_{roll} = F_C(SrcJ_{roll})$$

  $SrcJ$: Source joint (driving joint from the user side).
  $DestJ$: Destination joint (target joint at the representation side).
  F_A, F_B, F_C: are functions that can be applied on each individual axis to alter the motion (e.g. linear mapping $F(X) = X$)

- **Spatial mapping**: Similarly to rotational mapping constraint, the spatial mapping constraint the mapping between axes (X,Y,Z) to the corresponding ones in the destination.

- **Directional mapping**: The spatial and rotational mapping should

follow the sign of the source motion (not to use inverse function on the joint motion).

In summary, this design pattern is intended to remap body modalities into different ones in order to overcome a physical disability imposed by a certain modality. This design pattern can be combined with other meta-modeling design patterns when designing a system.

3.4.4 Modality Expansion
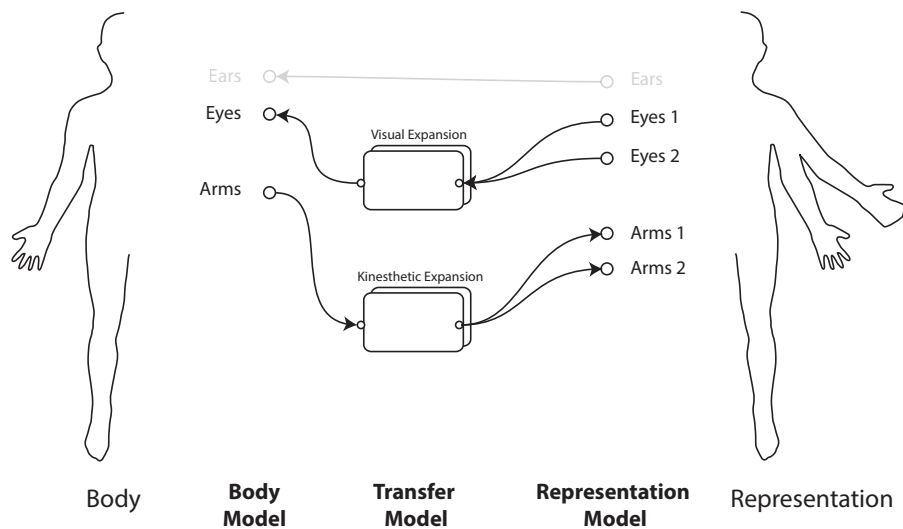
Figure 3.12: Modality expansion meta-modeling.

The previously proposed design patterns have mainly addressed direct one-to-one body schema mapping in which each body modality communicates with a single modality at a time. In that type of systems, the sensory and postural topologies are constrained to a single representation. *Modality Expansion* design pattern addresses systems with non-linear mapping, that is the operator can access multiple modalities and receives feedback from them, or controls them simultaneously. This control and feedback mechanisms in this meta-model should address the following considerations:

- Real-time simultaneous representation of body, visuals, and auditory feedback in multiple locations.

- Natural mechanism of presenting the visuals from the multiple sources to human user.

- Intuitive mechanism for switching the visual perception between the locations, while maintaining the awareness of the other locations.

To illustrate this design pattern, Figure 3.12   shows mapping body modalities into multiple modalities in the representation. The representation can be single with multiple modalities, or several representations that the user connects to and controls simultaneously.

To design a meta-model that can provide such non-linear mapping of body topology and representation topology, the concept of "Modality Layer" is used. The layer contains both the postural and perceptual modalities of the representation used, and thus when N modalities are required to interact with, we would have N layers $[L_1...L_N]$ that are derived from these modalities. These layers are aggregated together to produce the input to the perceptual or postural modalities of the user. These layers can be mixed (or blended) and delivered to the user and representation using corresponding transfer blocks.

This design pattern helps to expand body's limitations to perceive and control multiple representations that can be located locally, or connected remotely (ubiquitous).

3.5 Physical Embodiment Toolkit

In order to facilitate the process of prototyping meta-models and experiment the effect of altering the mapping of various modalities, an embodiment toolkit is proposed. The toolkit is a general purpose physical representation designed in a robotic form, which provides essential modalities to access through EDD framework. This toolkit can be looked as being a tool to reprogram our bodies and our presence. The design considerations for this proposed toolkit are described in this section.

For this thesis, this toolkit was used completely or partially in several projects (discussed in Chapter 5) and provided an agile design and testing process for these projects.

3.5.1 Human Driven Design

Human bodies provide high degree of complexity and constraints. The toolkit should provide a balance between the complexity of replicating the human body structure and the affordability to create a general purpose system. The following modalities are mainly considered to be available in the design of the hardware:

- **Vision feedback**: as in teleoperation and presence applications, the main sensory feedback we would rely on to gather information about the surrounding environment is the vision, or the eyes. The representation should provide means of capturing real-time visuals and provide them in bodily acceptable format, that is a stereo vision. As in Telexistence systems, stereo vision is an important design consideration that is required by the slave system to provide. For general use case purpose, the interpupillary distance (IPD) for the vision system is to be fixed on an average of 64mm [19]. Field of vision should be wide enough to cover at least 60 deg of user's field of view, so the user can obtain feedback from the peripheral vision. The visual feedback should provide real-time feedback with corrected distortion in order

to avoid "Motion Sickness" cases, that is, the disagreement between visually perceived optical flow and motion, and the vestibular system's sense of movement which caused the motion. In this thesis, the latency considerations should be less than 100ms for the visual system. The responsiveness (or framerate) of the vision system should be high enough to minimize the motion sickness as well as balanced to maintain low latency feedback (the higher the framerate, the more performance demand from the system is required, and thus higher latency will occur). Within these constraints, the spatial resolution of the images can be considered depending on the target application (for teleoperation and inspection purpose, low resolution can be sufficient to be used).

To use in EDD, the toolkit should expose a block that provides access to the eyes modality of the representation. In the representation model category, [Eyes] block is used to output stereo visual data channel from the toolkit into the meta-model.

- **Auditory feedback**: The auditory system of human body is an important cue to have spatial awareness of the events occurring around. The binaural feedback provides an understanding of the sound sources locations. For this modality, the system should provide a similar experience of human ears. To maintain high auditory spectrum feedback, the audio system should provide high-frequency sampling of the audio. Human ears can perceive a range of 20Hz to 20,000Hz of auditory spectrum. Thus the sampling rate should be at least twice the frequency (based on NyquistShannon sampling theorem), so 40,000Hz audio sampling is the minimum requirement for the audio system. The closest supported standard audio frequency by the hardware is 44,100Hz. The audio transmission should be synchronized with the visual feedback in order to avoid any disengagement of visual-auditory cross-modality during operation.

  The auditory feedback is integrated into EDD framework by providing

[Ears] block under representation model category to output binaural audio data channel from the toolkit into the meta-model.

- **Verbal communication**: Since the representation is generally used as a replication of human body, auditory communication system would be required to provide bi-directional audio from/to the toolkit along with the auditory feedback system. The design verbal communication system has fewer constraints than the previous modalities. Single audio channel output is sufficient to be included in the toolkit design. For frequency considerations, this channel mainly transmits the verbal audio from the human operator, thus the frequency sampling can be less than the auditory feedback channels. Old telecommunication systems (such as the telephone) used 8,000Hz sampling rate, although it was sufficient for human speech however it cannot support sibilance (ess/s/ and eff/f/ sounds). To overcome this, Wideband frequency 16,000Hz sampling rate is considered for the transmission and playback.

  In EDD framework, under representation model category, a [Mouth] block is provided by the toolkit to communicate the auditory data channel from the meta-model to the representation hardware.

- **Postural design**: the mechanical structure of the toolkit reflects the degrees of freedom (DOF) and limits of the operation and mapping with the user. The postural schema of the toolkit should provide visual-kinesthetic cross-modality, thus a structure that mimics the head motion is needed. A three DOF head structure would provide rotational motion along the tilt, pan, and roll axes are sufficient to produce the motion for the eyes. Speed and joints acceleration of the head design is required to match the speed and acceleration of human neck. If an inadequate speed and resolution were used, the system would result in mechanical latency and thus motion sickness would occur. To address these speed limitations, a user-study is required to calculate the maximum rotational speed of human neck.

For the parallax motion of the upper body, the addition of three extra DOF would provide the lower motion of the human torso [106]. However, the usage of these extra joints would increase the size and affordability of the toolkit. The addition of these DOFs is optional, and is application specific. Two versions of the toolkit can be designed: Head only (3DOF) and Head+Parallax (6DOF).

The postural modalities are exposed into EDD under representation model category as a set of blocks representing the body. For the neck modality, the [Neck Joint] block is used to communicate the target head orientation and position (if parallax motion was required).

3.5.2 Extensibility and Modularity

Although the toolkit provides a minimal amount of modalities as described in the previous point, however, it should also support the ability to modify and extend its functions by integrating new types of modalities into it. A modular design is considered for both the software and the hardware. The design of the software architecture on the toolkit should be capable to handle the addition of new types of functions, while maintaining the existing ones,e.g. adding an extra set of vision module into the toolkit. A common use case for such modularity is the haptic sensors. The number of haptic sensors can vary depending on the application, for example, its possible to integrate a tactile sensor on the bottom of the toolkit to capture the ambient vibrations of the location the toolkit is placed at (table top for example).

For each of the newly added modules, a separate information channel is generated to communicate with the EDD model. Also, a dedicated block is to be designed and integrated into the representation model (defined by *Representation engineer*).

3.5.3 Accessibility

In order to design remote presence alteration experiences using EDD meta-modeling, the toolkit should provide network accessibility. Thus, all the communications between the toolkit and the user side should be carried over dedicated network channels and in a digitized format. The network medium can be done through wired local area networks (Ethernet), or a wireless medium (e.g. Wi-Fi IEEE 802.11 standard) in cases such as mobility is required. This design consideration helps to allow flexibility in connecting to the toolkit, either as a single toolkit at a time, or multiple toolkits simultaneously.

3.6 Design Summary

This chapter has introduced the overall concept of Embodied-Driven Design (EDD) and meta-modeling design process. EDD derives the various modalities of the human body in a loosely, hierarchical structure in order to engineer a mapping between a human body and an alternative representation, such as telexistence system. This loosely representation helps to alter the input and output of the modalities, or remap them with non-corresponding modalities at the representation side. Meta-modeling is the process of defining body schema mapping by using various types of perceptual, transfer, and representation blocks. Four meta-modeling design categories were described: Direct body transfer, Representation alteration, Topology reconfiguration, and Modality expansion. These categories provide the design process for the systems that will be discussed in Chapter 5 . However, EDD is not limited to these meta-models design categories. To expand the usability and utility of EDD, a physical embodiment toolkit is proposed that it can be used along with modeling framework to achieve customizable body mapping and interaction alterations. This toolkit is also used in the projects discussed within this thesis.

Chapter 4

System Implementation

In the previous chapter, the overall design requirements for developing an embodiment and body schema transfer system were specified. This chapter will discuss the implementation details for a designer friendly embodiment modeling framework. The major components for the framework are discussed here: the Meta-modeling editor, modalities information channels and meta-blocks, the integration with an existing game engine (Unity 3D), performance optimizations, and physical embodiment toolkit.

4.1 Embodiment Framework Overview

The embodiment framework contains various components and blocks that defines the topology of the system and representation. An overview of the framework is shown in Figure 4.1 . The framework consists of: **User side** which communicates the state of the operating user with the model
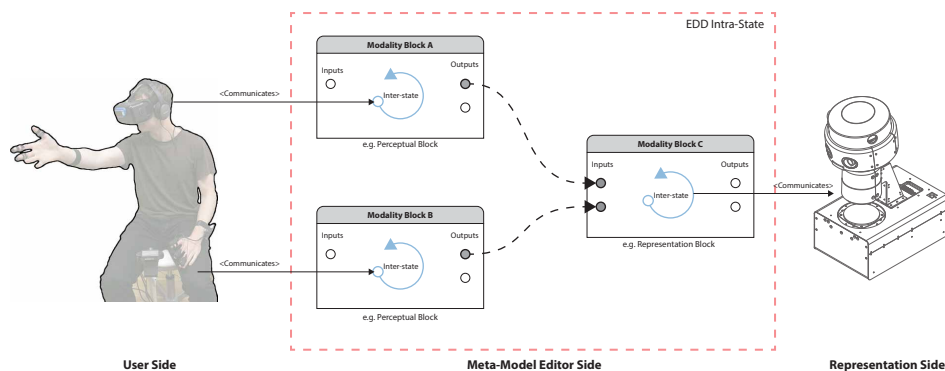
Figure 4.1: An overview for Embodiment Framework components.

using tracking and perceptual devices, **Meta-Model Editor side** which defines the interacting blocks and modules of the target schema of the system, and **Representation side** which communicates the state of the used representation (virtual or physical) with the model.

The user side contains modular tracking and perceptual hardware in which it can be customized depending on the intended application. For visuals presentation to user's eyes, a head mounted display is mainly for presence transfer systems (derived from telexistence systems). For body tracking, several tracking systems has been integrated into the framework in order to assist the degree of body tracking based on the number of joints intended to be tracked. For example, when head tracking is only needed, then HMD built-in tracking functionality can be used without the need for an external tracker, but when eyes, arms or legs tracking was required, then different tracking approaches would be used (individually or combined, e.g. optical tracking and data gloves systems). The perceptual and tracking modalities are exposed into the meta-modeling editor as set of blocks that transfer the state of the modalities from/to the model. For example, a [HMD] block would have an input representing the images to be displayed, and an output for its spatial position and rotation.

Representaiton side is dependant on the type of the body used in the representation. Two types of representations are addressed: Robotic representation, and virtual bodies. These representations can be used individually or in a combination, for example superimposing virtual arms on the visual feed from a remote robot representation. The representation blocks are responsible to communicate with the used body to transfer the modalities data channels from the meta-model and the representation. For postural modalities, joints data are used to control the position and orientation of the links used in the body (e.g. servo motor for robotic form). Similarly, sensory modalities uses transfer blocks to communicate the captured information flow from the representation to the meta-model. [Eyes] transfer block, for example, provides an image stream from the cameras used in the representation.

The meta-modeling editor handles the inter-states and intra-state of the modalities channels flow of the blocks used. The inter-state corresponds to the internal processes of the block throughout a procedure of capturing the inputs through the available slots, applying internal calculations or alteration to that input channels, and expose the results though output slots with a corresponding data type. The intra-state of the meta-model controls the flow of the channels, and constraints the data flow and types used. The intra-state is defined by the embodiment designer.

4.2 Meta-Modeling Editor

The meta-modeling editor provides a designer oriented approach to model the interactions between the user and the representation by using various types of meta-blocks. Figure 4.2 shows an overview of the meta-modeling editor developed in this thesis. This editor has been implemented under Unity3D. The reason of choosing this game engine in particular is that this engine has been widely accepted platform by both inexperienced and professional designers and programmers, also this engine supports cross-platform deployment which would increase the usability of the meta-modeler system. The design consideration for this editor is to focus on graphically oriented visual programming with minimum amount of writing code to cre-
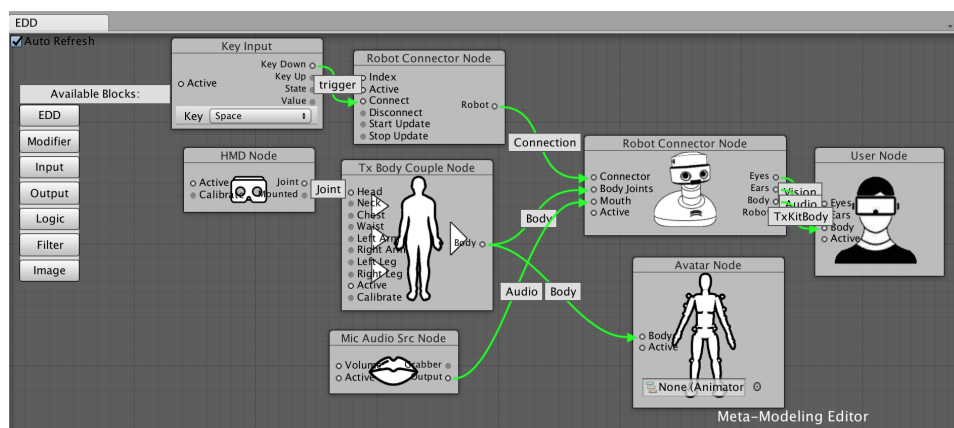
Figure 4.2: Meta-Modeling Editor Overview.

ate the model. The editor contains several categories of meta-blocks which can be basic arithmetic or geometric operation blocks, or higher level of interaction such as body tracking blocks. Figure 4.3 shows how the meta-blocks are organized in a tree structure based on the category and functionality. The designer can access the blocks directly from the side panel in a hierarchical structure, or directly search for the required block by using the search function (Ctrl+Mouse Click) and entering part of the block name in the search box. Using this mechanism, the designer would avoids the confusion of finding the required block.
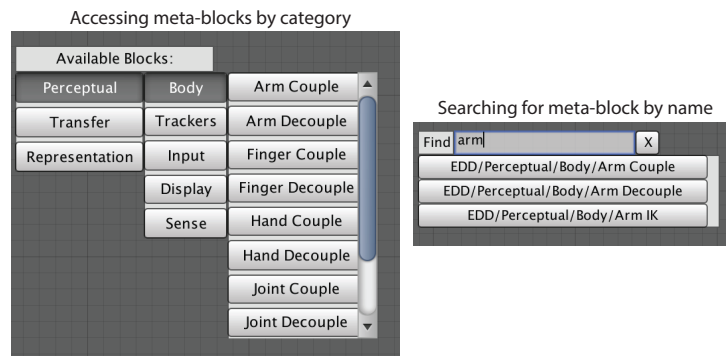
Figure 4.3: Accessing meta-blocks by Searching or Category.

In order to use the meta-modeling editor, the designer needs to create a new [EDD Model] object under Unity3D editor by choosing: *GameObject Menu* →*EDD Model*. Once an EDD Model is added, its possible to access the Meta-Modeling editor by selecting the model and clicking [Open Model] from the inspector. The Mea-Modeling editor supports to run multiple models on parallel, so its possible to design dedicate each model into different application or use case. For example one meta-model handles the auditory transmission and alteration, and another one for the body control and mapping. Thus the design process can be distributed.

When a meta-block is added to the model, the block exposes the available input and output channels (or slots) that provides the accessibility from/to its functionality. As a matter of making a convention, the data flow of the blocks always goes from left to right, so input slots are located
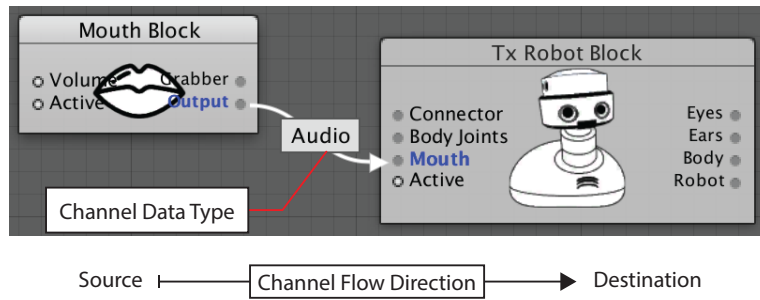
Figure 4.4: Channel data flow between the blocks.

on the left side, and output slots on the right. When more than one block is added to the model, its possible to connect their channels as long as the channel data matches. Although it constraints the modeling process, however it assists the designer to quickly understand which channels can be connected by highlighting the acceptable input channels in all the blocks added to the model. When channels are connected, a data flow is registered between them. Figure 4.4 shows an example of connecting [Mouth] block into a [Telexistence Robot] representation block. All blocks in the meta-editor contains *Active* state in the input which accepts boolean values, simply used to enable or disable block functionality.

4.3 Meta-Modeling: Blocks Implementation

As was described in the design chapter, EDD consists mainly of three set of blocks that can be used to define the body schema transfer model. The design considerations were reflected in the implementation following the same categories for creating meta-blocks. This section describes the details of each block category with the underlying details of the mechanism used in the inter-state of the blocks.

In order to organize the accessibility to the blocks, each block should define a path in form of "[Category Name]/[Sub-path]/[Block Name]". This path is set when the block is implemented in the script. Listing 4.1 provides

a sample template to declare new meta-blocks. The block class require an attribute of type [ModelBlock] in order to automatically register the block into the graph using the path specified in its argument. Two optional arguments can be added for the icon image and width of the block in pixels in the editor.

```
//mandatory attribute, declare this object as being a
    block in the editor using the specified path, and
    set an icon image with width of 100px
[ModelBlock("Transfer/Modifier/Block
    Sample","Icon",100)]
//the block should be derived from BlockBase class
public class BlockSample : BlockBase
{
    //internal state variables
    //.....
    //Declare an input channel with data type of
        XDataType
        [Inlet]
        public XDataType Input{
            set {
            //set internal variables
            }
        }
        //internal block data used for output channel
        YData _outputData;
        //declare an output channel.
        [SerializeField, Outlet]YEvent OutputY;
        void Start () {
            // Initialize block state
        }
        // UpdateState is periodically
        void UpdateState () {
            //process the internal state of the block
            //.....
            //signal outputs
```

```
            OutputY.Invoke(_outputData);
        }
}
```

Listing 4.1: a definition template class for meta-blocks objects

4.3.1 Perceptual Blocks Category-set

This category of blocks are mainly concerned with the user side. The blocks provides a representation of the operator body's state and posture by using **body state blocks**, which is captured using **body tracking blocks**, as well as providing to the user information and sensory feedback using **feedback blocks**.

Body State Blocks

Body postural representation is designed as a set of blocks that contains the individual joints state. The structure of the blocks follows the previously described Loosely Coupled Modalities (LCM) design pattern in which the structure can be decoupled into smaller regions or grouped back to reconstruct the original region. The minimal region that body structure can be decoupled into is a [Joint] block Figure 4.5 . The joint block contains the spatial attributes represented by a position and rotation channels, which can be captured using [Joint Decouple] block. The position channel accepts a data type of *Vector3* and the rotation accepts *Quaternion* data type. The coordinate system which the joints are represented with is Left-Handed system. Using a singular coordinate system among the spatial
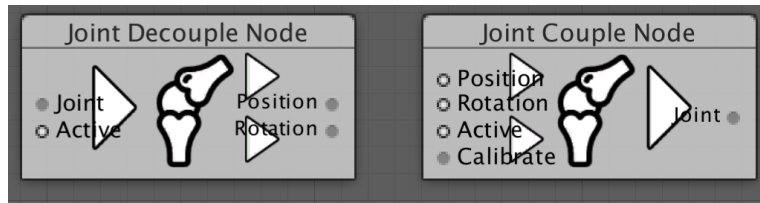
Figure 4.5: Joint block, the minimal level of postural schema.

blocks is important to decide, so the tracking and representation blocks can use the same system to provide or read the joints data from.

Each joint contains a *Calibrate* input which, when it is triggered, would set the current spatial state of the joint as its initial postural. That is:

$$Position_{Initial} = Position$$
$$Rotation_{Initial} = Rotation^{-1}$$

(4.1)

And for each new input value:

$$Position = -Position_{Initial} + Position_{Input}$$
$$Rotation = Rotation_{Initial}^{-1} \times Rotation_{Input}$$

(4.2)

The entire body structure is represented using Body block Figure 4.6 which contains a hierarchical structure of the main joints of the spine (waist, chest, neck, and head), two arms and two legs (Left and Right). The body block follows LCM design pattern in which it can supports the decoupling and re-coupling of its sub-structure. Arm block structure is shown in Figure 4.7 . The arm block can be decoupled into four joints (shoulder, clavicle, elbow, and wrist), and a hand. The hand block is structured using five fingers, and each finger has three joints. Finally the leg block is shown in Figure 4.8 and can be decoupled into three joints (hips, knee, and foot).

To clarify the benefits of using such highly hierarchical body structure, Figure 4.9 shows an example of how the body parts can be selectively
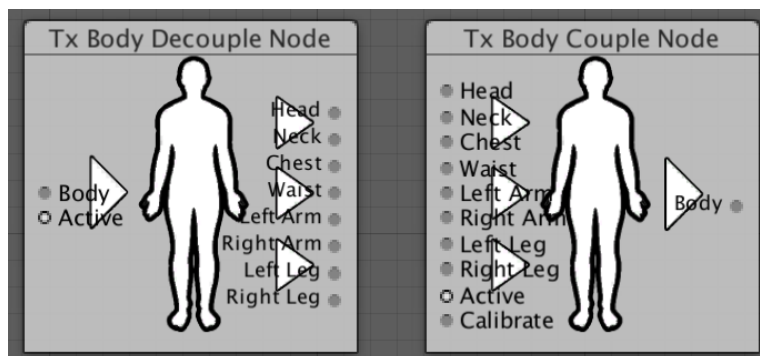
Figure 4.6: Body decoupling and coupling blocks.

77
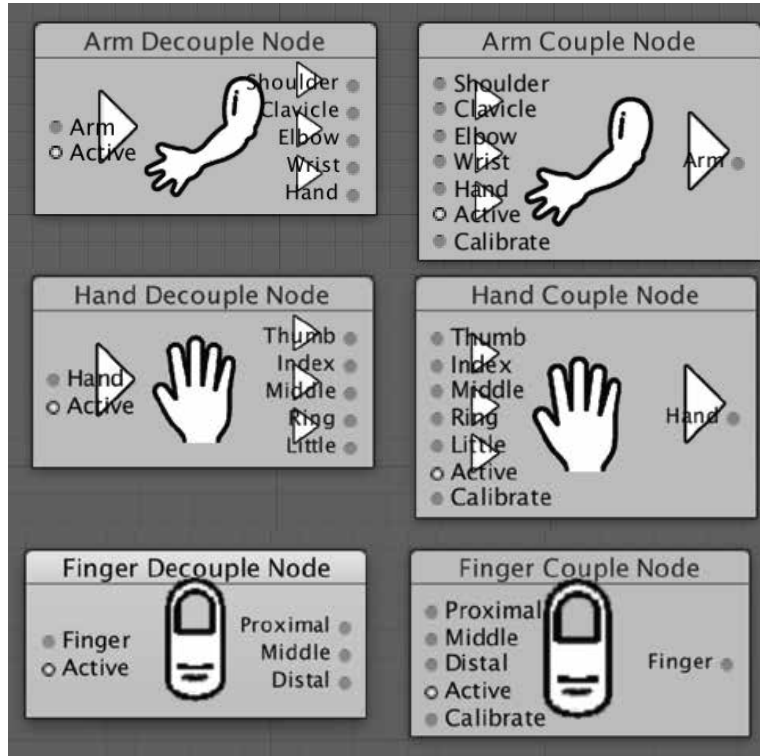
Figure 4.7: Arm, Hand, and Finger decoupling and coupling blocks.

Figure 4.8: Leg decoupling and coupling blocks.

controlled, mapped, and structured. This example shows a single joint controlling two different joints at different levels (head and left arm elbow), while the leg block is mapped into both left and right legs of the body block. Similarly, a single finger block is used to drive three different fingers of the left hand (thumb, index, and little fingers). The data flow type is highlighted on the channels outputs of the blocks.
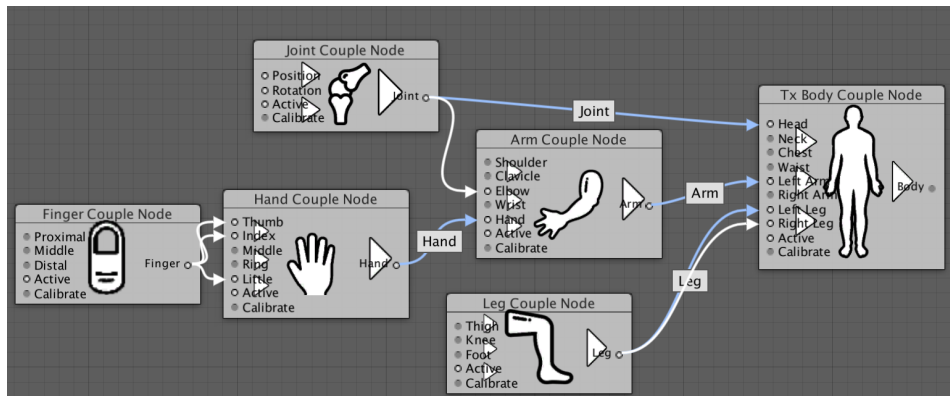
Figure 4.9: An example of Loosely-Coupled Modalities (LCM) to Restructure Postural Schema of the Body.

Body Tracking Blocks

For body spatial tracking, various technologies and systems has been tested and integrated into the meta-modeling editor. Each tracking system provides advantages and use cases which varies its functionality from the other systems. Table 4.1 summarizes a comparisons between the used systems in term of target body regions required to be tracked (usability), the degree of complexity the system would require to setup and use, and finally the accuracy and robustness of the tracking. Within the tested systems, a trade off between the three criteria was required when deciding the target system to be used in the model. Each of these systems uses different method and procedure for motion capture, and each has its advantages and disadvantages which would constraint the design of the system. In the listed systems, two distinct categories of tracking mechanism are used:

79

Tracking System	Usability					Complexity	Robustness
	Head	Hands	Fingers	Body	General		
Oculus HMD	√	×	×	×	×	Low	High
Oculus Touch	×	√	△	×	×	Medium[1]	High
Leapmotion	×	√	√	×	×	Low	Medium
5DT Data Glove	×	√	√	×	×	Low	High
Neuron Perception	√	√	△	√	×	Medium	Medium
OptiTrack	√	√	×	√	√	High	High

[1] Oculus Touch require the user to hold them while operation, which is not a seamless tracking method.

Table 4.1: A comparison between body tracking systems based on the usability, complexity, and robustness.

1. **Optical tracking**: In this category, the motion capture system uses a single or several cameras (or light sensors) to triangulate the 3D position of a set of points defining the tracked object. Based on this triangulation process, the points are matched with a predefined model representing the tracked object, and then the position and orientation of the object can be calculated. These systems require a calibration process for the cameras to extract the intrinsic parameters (related to the lens of the camera, sensor size, ...etc), and extrinsic parameters (related to the spatial distribution of the cameras when multiple were used, and the location and orientation of the reference ground plane). Since a continuous visibility of the points by the tracking cameras is required, these systems deliver poor tracking when the points were occluded.

   Two types of tracking systems uses optical tracking:

   - Active type: the tracked points of this system have an illuminating light source (usually within the infrared spectrum) that can be distinguishably tracked from other non relevant points in the environment. Both Oculus HMD and Oculus Touch uses this mechanism for tracking.

- Passive type: in this type of systems, the tracking cameras are equipped with a special light source (usually infrared light source) that illuminate the environment. The markers or features that needed to be tracked usually reflect the illuminated light back to the camera (usually using a retroreflective material), and thus can be extracted. In this type of systems, usually high noise occur and would require tuning of the environment and the camera parameters. The advantages however is low cost and the flexibility to define custom tracking props. Both Naturalpoint OptiTrack and Leapmotion systems uses passive tracking.

2. **Non-optical type**: In optical tracking category, the motion capture systems required a reference tracking camera and continuous visibility of the tracked features. The non-optical type however, does not require tracking source, thus can perform better in terms of occlusion and freedom in motion. The tracking mechanism however is different for each type of systems. For the used systems, two types of motion tracking is used:

   - Inertial type: systems of this category use intertial motion sensors, or most commonly, inertial measurement units (IMUs) that combines three different sensors: gyroscope, magnetometer, and accelerometer on a miniature device, and uses sensor fusion algorithms to measure the rotational rates of the unit. Since these units use relative motion and accumulate the results of the internal sensors to calculate the current orientation, a small error also accumulates resulting orientation draft. This error is due to the non-perfect accuracy of the sensors. The system would require calibration procedure (system dependent posture, usually T-Pose) whenever draft occurs. Neuron Perception system uses IMUs to measure the orientation of the body. This system can support various layouts for tracking depending on the number of

Figure 4.10: Full body tracking system (Neuron Perception).

IMUs used and their location on the suit shown in Figure 4.10 : 1) Full body tracking using 11,18,21, or 32 IMUs, 2) Single arm tracking using 3,6,9, or 11 IMUs, or 3) Upper body tracking using 6,11,21, or 25 IMUs. The calibration procedure for this system is shown in Figure 4.11 .

- Mechanical type: these systems uses directly mounted sensors that measure the mechanical motion of human joints. The sensors are located directly at the joint location to measure its bend. The problem with these systems is they are highly constrained due to the mechanical design, and require exact alignment with the joint. 5DT data glove system uses fiber bending sensors for tracking finger motion.

A combination of multiple systems is also possible in order to acquire the highest robustness with minimum complexity while maintaining the number of joints tracked. For example, to use Oculus HMD to track the head motion

(1) Check sensors connectivity   (2) Apply calibration postures   (3) Confirm tracking precision

Figure 4.11: Calibration process for Neuron Perception system.

while providing the visual and auditory feedback to the user combined with 5DT data glove to track fingers posture, and overall body tracking using Neuron Perception system.

As part of body state quantification modeling, eye gaze modality provides important information about the area of interest the user is looking at. This modality can be used as supportive input for multi-modal interaction as will be described later in operation and model alteration meta-models. An pupil tracking system was used by embedding dual cameras into the head mounted display. The system uses infrared cameras located behind the lenses of the HMD, and a half mirror placed at 45 deg. Five infrared LED emitters are embedded into the lens which illuminate the eye pupil. The camera captures the reflected light of the eye via the half mirror. An illustration of the system is shown in Figure 4.12 . A custom plugin for Unity3D was developed and integrated which communicates with the camera software. The plugin carries on the calibration procedure which is required per-user before start using the trackers. A 9-points eye fixation calibration is integrated, which maps the captured pupil locations with predefined points in 2D space. 120 samples per point (which is about 2 seconds per eye for 60 FPS tracking, in total 18 seconds for all points) provided good balance between calibration time and accuracy of tracking.

The last modality for capturing user's state is the speech modality. The

Figure 4.12: Eye tracking module used (PupilLabs + Oculus DK2).

audio from user side is captured from an attached audio input interface. Although in most cases only a single audio input interface is needed to capture the state of user's speech, however the model supports the use of multiple audio interfaces. For example, in multiple audio input case, its possible to triangulate the position of the audio source using the volume. These types of use cases are application dependant.

To use in the meta-modeling editor, the tracking systems were encapsulated and integrated as meta-blocks shown in Figure 4.13 :

- OptiTrack Block: Communicates with an optiTrack server that runs the tracking system. The node requires to specify the target server IP address. For each block, the name of the target tracked object should be specified in its block parameters. When the meta-model runs, the object data is output as [Joint] data channel.

- HMD Block: Handles the tracking of the connected HMD. The block provides two outputs: [Joint] for the spatial position and orientation of the HMD, and [Mounted] as a boolean channel indicating whether the HMD is being used or not (compatible with Oculus CV1/HTC Vive HMDs).

- Neuron Block: Similar to OptiTrack Block, it also requires to specify the server IP address running Neuron Perception service. This node requires to specify the ActorID when more than one tracking suit is being connected. The block outputs [Body] data structure containing

84
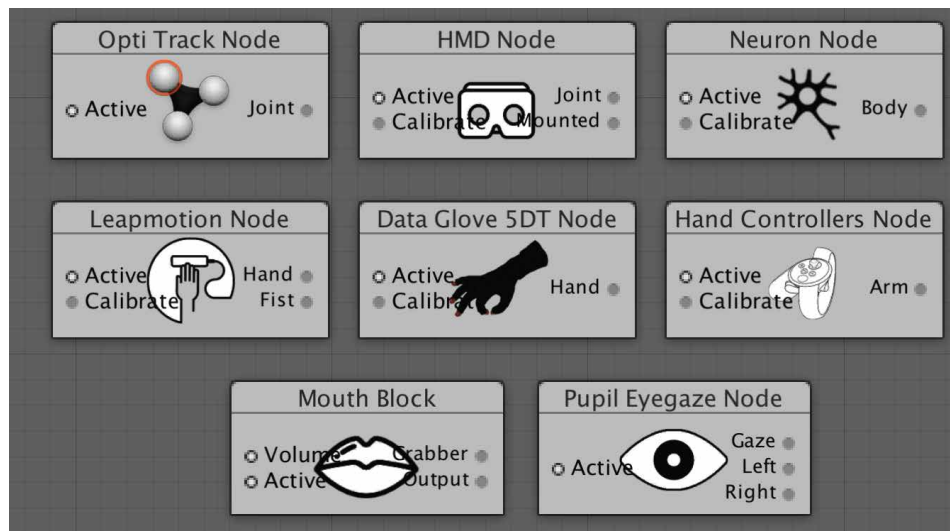
Figure 4.13: Body state tracking blocks.

the entire tracked body joints data.

- Leapmotion Block: Provides accessibility functions to leapmotion tracking. Requires to specify the target hand to track (Left or Right), and it outputs [Hand] data. [Fist] provides whether the hand is closed or not as floating point value.

- 5DT Data Glove Block: Similar to Leapmotion, requires to specify the target hand to track, and it outputs fingers information on the [Hand] output channel.

- Hand Controllers Block: Uses Oculus Touch to track hands position and fingers bending. Outputs [Arm] structure containing the estimated position of the hand relatively to the HMD position.

- Mouth Block: Connects to an audio input interface, and provides the sampled audio on [Output] channel. It is possible to indicate the sampling rate in its settings, default set to 16,000Hz.

- Pupil Eyegaze Block: Uses the functionality of Pupillabs eye tracking hardware to measure the pupil gaze position. Outputs both [Left]

and [Right] values for both eyes, and the estimated looking position on [Gaze] channel. Data used is a 2D value vector [Vector2].

Feedback Blocks

Feedback blocks mainly communicates the meta-model perceptual states back into the user. Within the current scope of this thesis, three different types of feedback modalities are explored and integrated into the EDD framework and meta-modeling editor:

• Visual Feedback: transforms stereo images input into a displayable format to be presented into the user's vision. The visuals are directed into [Vision] data type which contains dual textures of type Texture2D. These images are either captured from representation blocks, or are generated from transfer blocks. Depending on the display device used (HMD, Monitor, ...etc), parameters related to the field of view and aspect ratio of the device should be specified. In the meta-blocks, these functionalists are encapsulated and are automatically calculated, however it is still possible for the designer to alter these settings.

• Auditory Feedback: plays back audio samples stream into an output audio interface at the user's side, and represented as [Ears] module. Samples are encapsulated into an [Audio] data type that can handle multiple audio channels, and different sampling rates. For audio playback, the sampling rate by default is set to 48,000Hz, but can be changed by the designer.

• Haptic Feedback: communicates with a haptic interface connected to

Figure 4.14: Fingers force feedback display.

the user's side, and can provide both force and tactile feedback information into the device. Both data are represented using [Haptic] data structure. Force feedback is a three-directional vector3 which provides the 2D (X, Z) tangential force along the surface of the display, and the perpendicular (Y) pressure force on the display surface. Currently, Gravity Grabber [67] type of displays are supported for force feedback. A custom display has been developed which can provide horizontal (X) shearing forces, and the perpendicular (Y) pressure forces for three fingers of each hand Figure 4.14 . For tactile feedback, the data is stored as an array of floating points within the range of [-1,1] representing the signal captured by a vibrotactile sensor, or generated in software using a transfer block. The sampling rate for this signal is less than 2000Hz samples (representing a maximum frequency of 1000Hz which is the limit of Pacinian corpuscle receptors in the skin). To display the tactile data, a vibrotactile actuator can be used by directly connecting it to an output audio interface and specify the interface ID in the settings of the haptic block.
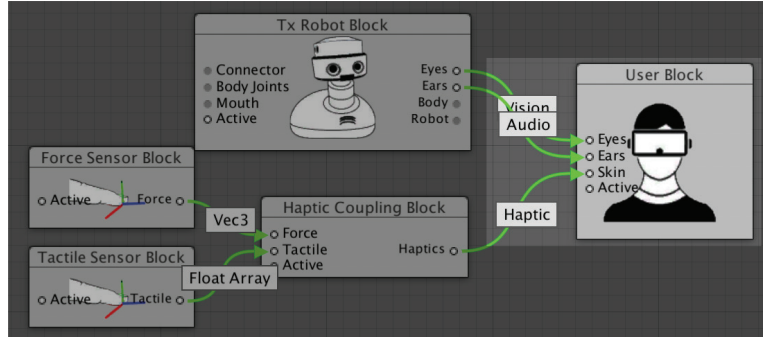
Figure 4.15: User feedback block meta-model example.

Since all these modalities are related to the user feedback, a single [User] block Figure 4.15   that provides the functionality to receive the various perceptual input has been designed and integrated into the meta-modeling editor. The block receives three inputs representing the Eyes, Ears, and Skin modalities that corresponds to the visual, auditory, and haptic feedback respectively.

4.3.2 Representation Blocks

This set of meta-blocks communicates with the used body representation. It facilitates the process of transferring the information flow between the meta-model and the representation type used. Two main types of representations are used in this thesis: physical and virtual type. The physical type communicates with a robotic form located at the user side as a peripheral device, or remotely connected over the network, and controls its physical state such as servo link driving or captures sensory information from it, such as tactile sensors. And the virtual type mainly communicates with Unity3D editor into a virtual character (humanoid or non-humanoid) to control its body postural. Generally speaking, representation model blocks are considered as mirroring the functionality of perceptual model, that is the for each input of representation blocks a corresponding output from perceptual blocks and vice versa.

The first block type for the body model is the postural blocks. The control mechanism for the postural of the representation is driven by the [Body] data type that defines the new body schema from the user. Different approaches for control is dependant on the *Representation Engineer* choice. For example, the driving can be done using Forward Kinematics (FK) model (direct drive) in which the joints angle is directly applied from the input body joints data. For example, 3 DoF telexistence robot uses forward kinematics to drive the head motion of the robot from the head joint of the body. A 6 DoF robot (with parallax motion type) would use a different approach to calculate robot's joints angle from head position, using Inverse Kinematics (IK) model to drive the parallax joints from head position and rotation. The postural blocks should provide compatible data channel to the representation used, for example a network-based communication is used, a combination of UDP and TCP protocols can be used to bridge the data flow from Unity3D to the robot side, and for peripherals connected directly to the user side pc, a serial communication protocol can be used to deliver the control signals to the used peripheral. The EDD framework provides an

abstraction to the communication protocol used, so the same block can be reused for different representations. The *Representation Engineer* can either develop new protocols to communicate with the hardware used, or to build on to of the exciting protocols used in the framework. The physical toolkit in Section 4.4 provides building blocks for developing the representation hardware while maintaining the support to EDD framework representation blocks. The toolkit discussed also includes important details about media streaming optimization for various types of sensory modalities, mainly for the real-time visual feedback.
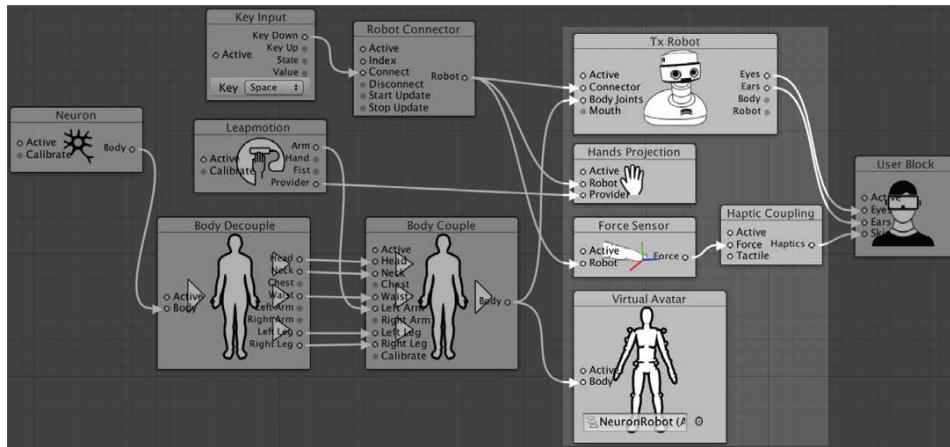
Figure 4.16: Using multiple blocks types for representation.

The other block type for the representation blocks is the sensory blocks. These blocks are responsible to deliver various modalities information back to the meta-modeling editor. Figure 4.16 shows an example of using both postural and sensory representation blocks used in parallel for two different types of bodies. [TxRobot] block encapsulates the postural modality function, and provides the sensory feedback on its outputs. The [Force Sensor] is used as a separate module since the number of sensors can vary based on the application. [Hands Projection] block provide the functionality for representation alteration meta-modeling in which it streams the virtual representation of the hands into the robot side. Finally, [Virtual Avatar] block communicates with Unity3D character model to control its postural state based on the shared [Body] input. In this example, both representations

are driven by the same body input.

4.3.3 Transfer Blocks

The final category of meta-blocks is the Transfer type. This category contains a wide variety of functions and modifiers which can be general purpose type (e.g. mathematics, image processing, body mapping alteration, ...etc) to a very application specific blocks (e.g. layered perception, motion transformation, ...etc).
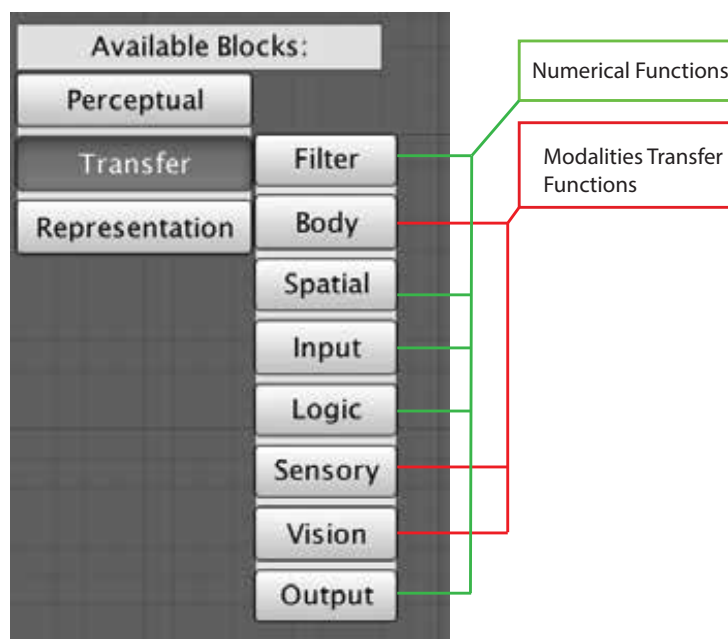
Figure 4.17: List of transfer blocks sub categories.

The transfer blocks are used as a middle function to modulate the data flow between the representation and user blocks. Also it can be used to apply logic operations on perception, for example activate or disable certain modalities channels flow based on behaviours or body state. A list of transfer blocks categories is shown in Figure 4.17 .

Examples of using transfer blocks to modulate body mapping and perception feedback are shown in Figure 4.18 . The first example highlights the use of spatial alteration filter block [Rotation Filter] to modulate head joint

Motion transfer blocks example. The rotation of the head motion is modulated before mapping it to the body module.

Vision alteration example. The visual feedback perceived by the user is provided is modulated using several transfer blocks.
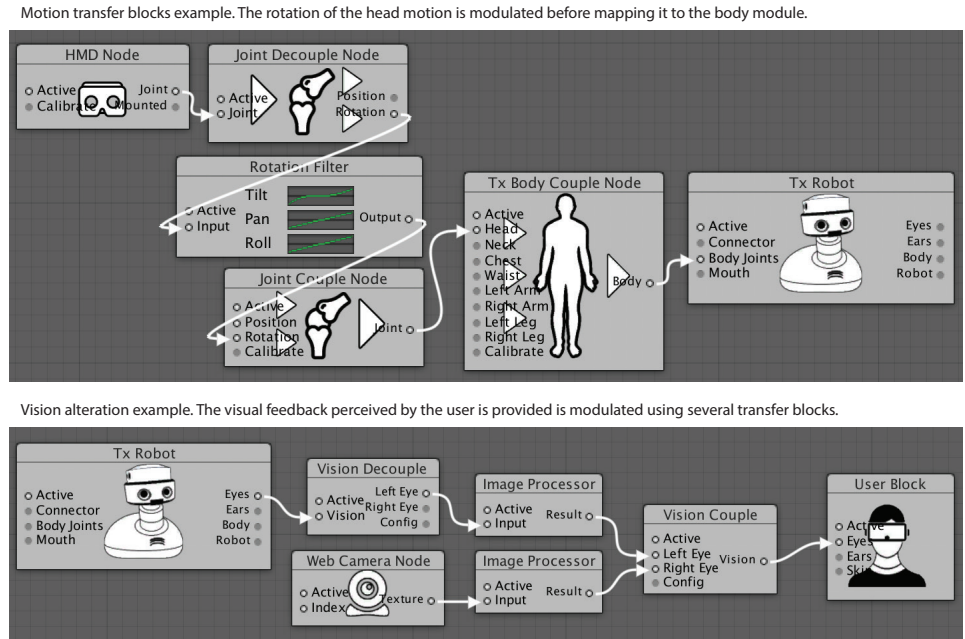
Figure 4.18: Examples of using transfer blocks to modulate body and visual modalities.

rotation into a nonlinear mapping. Three different mapping curves can be defined for each rotational axis (Tilt, Pan, Roll), which acts as a function for each angle. The second example shows a higher level of perception editing. In this example, the vision input modality for the user is altered using two different sources. The robot block's vision is decoupled into separate eyes using [Vision Decouple] transfer block, and then is processed using [Image Processor] transfer block. The image processor block uses a program written in shader language (HLSL) which runs inside the block process to produce a new image on its Result output channel. Robot's vision is mapped into the left eye of the user, while the right eye is mapped to a different source using [Web Camera] representation block.

4.4 Physical Toolkit Development

As a part of realizing the idea of Embodied-Driven Design on a larger scale, a general purpose physical toolkit was developed through several iterations and prototypes along the development of embodiment framework and the meta-modeling editor. The toolkit design considerations were discussed previously in the design chapter in Section 3.5 . The implementation process and details of this toolkit is discussed here.

An overview of the different prototypes that has been developed is illustrated in Figure 4.19 , and the description of each iteration considerations and requirement is shown in Figure 4.19 . Iterations A-B were highly application specific designs and mainly used for representation alteration purpose. Iteration C was a base step for developing higher performance toolkit using better optics and vision system than the previous generations, and has been used as a part of model alteration meta-modeling. And iteration D is the current design which provides a balance between performance and usability.

Along the hardware development of the robots, a corresponding software architecture was developed to facilitate the usability of the toolkit and to make it compatible with the Embodiment framework. Both the hardware and the software are highly modular and reusable.
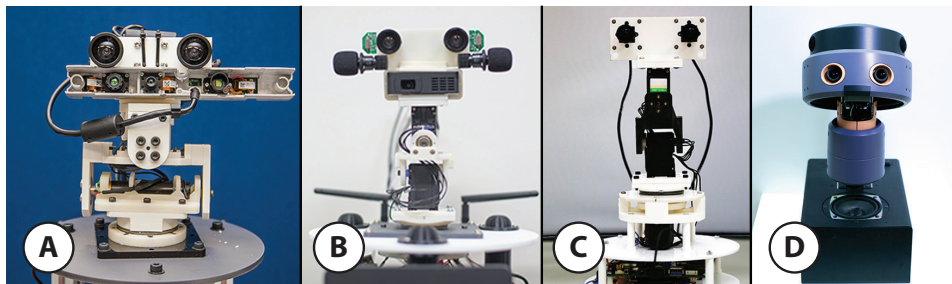
Figure 4.19: Physical toolkit design iterations.

Iteration	A	B	C	D*
Development Year	2014	2015	2016	2017
Application	Enforced & Mutual Telexistence		Layered Presence	General Purpose
Requirement	Depth Sensor	Projection Module	High performance	Usability
Meta-Model Category	Representation alteration		Model alteration	-

*Developed in a collaboration with Karakuri Products, Inc.

Table 4.2: Toolkit design iterations and considerations.

4.4.1 Hardware Design

The schematic design of the latest toolkit is shown in Figure 4.20 . This design provides the essential modalities for creating general purpose presence experience:

- **Vision:** Stereo camera with fixed interpupillary distance of 65mm (using camera model Ovrvision Pro, or See3CAM CU130 by e-consystems). Each camera has different specifications, but both are suitable for human oriented vision. Table 4.3 provides a list of the important details regarding the selection of the cameras.

- **Auditory:** Binaural microphone set (Audio Technica AT9911).

- **Communication:** Audio output speaker with an omni-directional design. A custom audio amplifier was added for controlling audio levels.

- **Body structure:** Motion control using 3 Axis head (Pitch, Yaw, Roll). Servo type motors were used to control rotation position of the robot head motion (Kondo Robots model KRS-4031HV ICS/KRS-4032HV ICS). The mechanical limits for each of the three axes is shown in Table 4.4 .
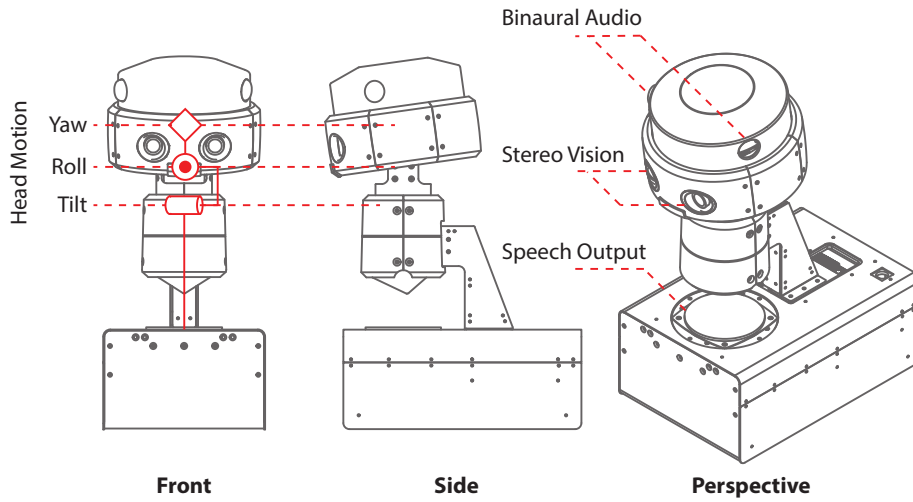
Figure 4.20: Physical toolkit hardware schematic design.

	See3Cam_CU13 +H0320KP (S Mount Lens)	Ovrvision Pro
Count	2	1 (stereo camera on board)
Image Stitching	Software stitch (unsynced)	Hardware stitch (synced)
Supported Formats	640x480@60 1280x720@60 1920x1080@30	640x480@90 960x950@60 1280x800@60 1920x1080@30
USB Connections	2 x USB3.0	1 x USB 3.0
Field of View	89x67	115x105
Quality	High quality (Raw image from sensors)	Low~Medium Quality (compression and optics)

Table 4.3: A comparison between the used cameras in the toolkit.

Axis	Min[-180,180]	Max[-180,180]
Tilt	$-20\,\mathrm{deg}$	$20\,\mathrm{deg}$
Yaw	$-90\,\mathrm{deg}$	$90\,\mathrm{deg}$
Roll	$-20\,\mathrm{deg}$	$20\,\mathrm{deg}$

Table 4.4: Head rotational mechanical limits.

4.4.2 Software Architecture

The LCM design pattern is reflected in the design architecture of the software for the framework. Figure 4.21 shows the major components used in the framework design and the implementation of the representation blocks. The software is distributed along two sides: robot side, and user side. In the current implementation, the robot side provides the following five services:

- **TxEyes:** responsible for camera access and streaming, handles image capture, encoding, and sending to the user side. Supports the following camera types:

  - Directshow UVC camera drivers, such as webcams.
  - Ovrvision Pro camera.
  - Ricoh Theta omni directional camera.

  Video latency in this system is measured to 180 ± 20ms when streaming dual images of size 640x480 at 60 frames per second on a LAN.

- **TxBody:** communicates with the hardware and servo motors. The following body architectures are supported:

  - Three axis head control (Pitch, Yaw, Roll).
  - Six axis body control, such as Torso type systems, controlled by head position & rotation six values tuple (X,Y,Z,Pitch, Yaw, Roll).
  - Drone type systems, driven by speed vector for position & rotation (Speed X,Y,Z/Rotational speed Theta, Gamma, Alpha).

- **TxEars:** provides access to audio hardware, capture, encode and streaming audio data over network to user side. Number of audio channels can be set, by default 2 audio channels (Left, Right) are configured.

- **TxMouth:** handles audio stream input from user side, and responsible for audio playback on the robot side.
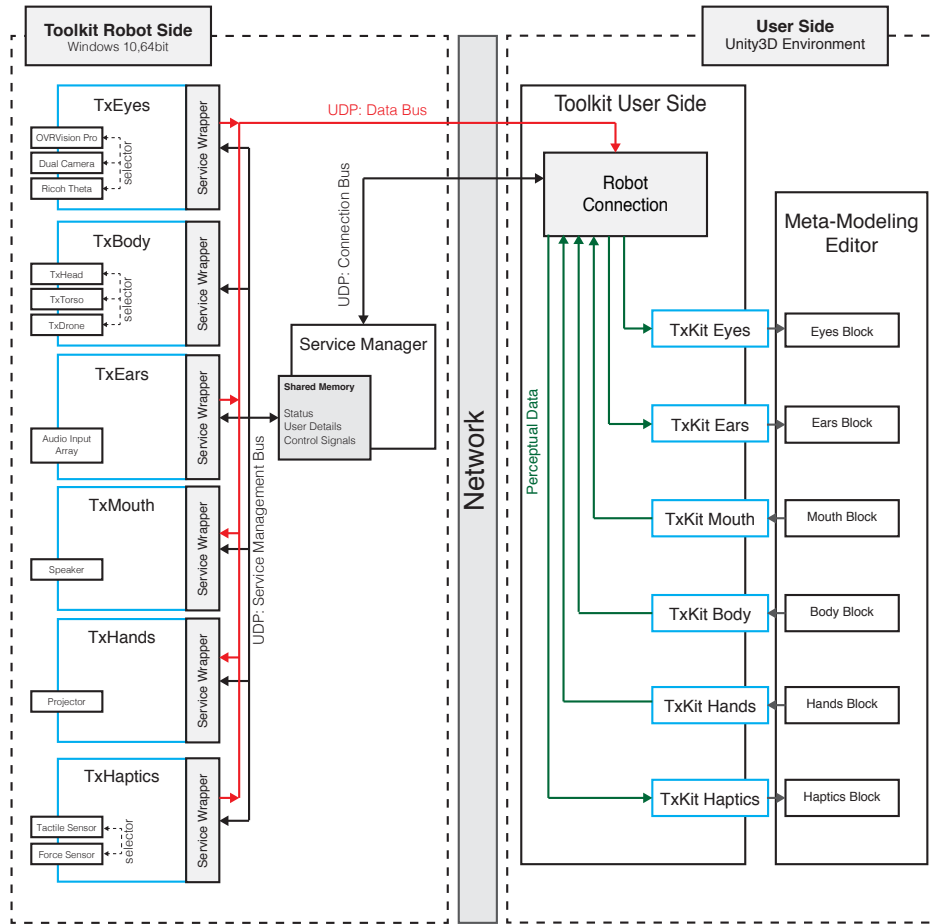
Figure 4.21: Toolkit software architecture.

- **TxHands:** responsible of displaying operator's body images from an egocentric point of view. Compatible with projector modules.

Each of these modules is encapsulated as a service, and runs on a separate process in the system, and communicates with a service manager that hosts them over UDP protocol. Figure 4.22 shows four different types of services running on the robot side. Each service provides information of their state, as well as the overall resources being consumed by the service (CPU, virtual & physical memory consumption). A shared memory is used between the processes to provide direct and quick access to the status of the system, user's address, and other control messages. This design provides

maximum stability over the system runtime, and reduces the chances of process crashes. Also, as mentioned before, these modules (or services) can run individually depending on the design requirements. For media streaming, encoding, and decoding, GStreamer library was used [1]. GStreamer is an open source multimedia library which encapsulates the core functions for general video and audio encoders/decoders. In the robot side, H.264 codec was used for video encoding, and Opus codec for audio encoding. Both encoders proved to be the most efficient encoders in term of low latency performance.

Figure 4.22: Service modules operating on toolkit side.

On the user side, modules are defined to handle the actual communication with the representation. Each of these modules corresponds to a representational block (Eyes, Ears, Mouth, ...etc). The software on the user side is implemented for Unity3D Editor integration under Windows operating system, as this game engine is considered as one of the most user friendly environments for virtual reality related application design.

---

[1]https://gstreamer.freedesktop.org/

4.4.3 Video Streaming Evaluation

For achieving close real-time perception streaming over the network, and to minimize the overall latency from the robot side to the user side, an optimized media tool-set were developed and used in toolkit. In the conducted tests, visual latency and image quality were the major considerations in the system design, which they affected the overall experience in robot operation.

In order to evaluate the streaming performance, a custom measurement software has been developed which calculates the capture-to-display latency of the cameras. The latency measurement system is shown in Figure 4.23 . To use this system, both the robot side containing the camera to be tested, and the user side should be located physically together. The system uses a display located at the user side that outputs a series of black/white frames which the camera of the streaming side captures. When a frame is displayed, the system records the current time stamp in milliseconds and waits before generating the next frame. The camera side captures, encodes the images into a compressed format such as H264 encoding, and streams the encoded images as a RTP data into the user side over the network medium using UDP protocol. The measurment side receives the RTP stream, decode it and compare the decoded image color value with the currently displayed frame. The system measures the cycle time (or latency time) by calculating the difference between the current timestamp and the displayed frame timestamp.

The camera module (Model no. See3CAM CU130 by e-consystems) used in the design of the third iteration has been evaluated for streaming performance over the network using the three different quality presets. Table 4.5 shows the measured latency for each mode.

4.5 Vision Bandwidth Optimization

In the previous section, the evaluation of the vision system showed an acceptable latency measurements (that is less than 100ms), however the ef-
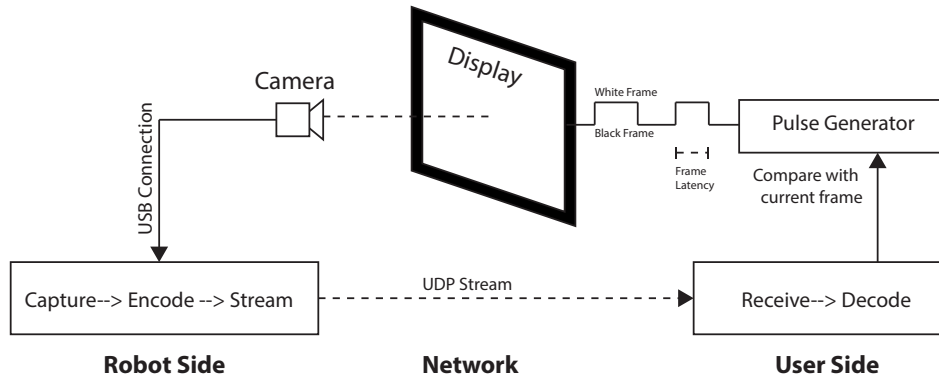
Figure 4.23: Camera capture-to-display latency measurement system.

Quality	Resolution	Framerate	Latency (ms)
Low Quality	640 x 480	60	64 ± 8
Medium Quality	1280 x 720	60	71 ± 8
High Quality	1920 x 1080	30	95 ± 16

Table 4.5: Camera streaming quality modes and corresponding latency.

fective spatial resolution of the encoded images is relatively low when high frame rate and low bandwidth were required. In this section, an optimization approach is proposed based on the human vision cues which would reduce the bandwidth and performance requirements significantly.

4.5.1 Problem with current streaming systems

Recent advances in virtual reality display systems began to offer end users immersive and high fidelity experiences in a rather accelerating manner. Current Head Mounted Displays (HMD) such as Oculus, Vive, and Samsung GearVR offer wide field of view (FoV) displays ($100 \deg \sim 120 \deg$) with a pixel density of $100 \sim 175$ per degree at refresh rate ranging from $60 \sim 90$ frame per second (FPS). In future, expected HMDs are to offer much higher resolution and refresh rate than now in order to enhance the visual experience and immersion towards human eye. However, such visual

experiences would require high processing bandwidth to deliver such information stream to the HMD, and performance would decrease exponentially based on the target pixel density and linearly based on the target frame rate.

On one hand, virtual reality applications that run locally would only require a system capable to operate at such high bandwidth, which is mostly addressed by hardware manufacturers. On the other, Telexistence and Telepresence systems would suffer more at addressing this problem due to network bandwidth limitations. Even with the current standard compression codecs such as MJPEG, H264/MPEG-4, and VP8 will have limited bitrate to accommodate for network streaming. Also, in this type of systems, encoding/decoding are required before streaming and after receiving the image stream, which performance is also dependent on the pixel resolution of the images. Extending the problem for omnidirectional type cameras which sends the entire 360 deg view, the resultant pixel resolution of the equirectangular image (the stitched omnidirectional 2D image) would increase 5 to 8 times more than the limited field of view pinhole cameras image type (at $90 - 110$ deg). As a consequence, the performance of encoding and streaming would increase the latency and decrease the overall experience for Teleoperation.

4.5.2 Foveated Streaming

To address the previous issues related to Telexistence and Telepresence systems, a perceptual driven approach was used. Human eyes have the highest visual acuity at the fovea in the retina, which is only about 2 deg of the visual field, and this visual acuity decreases the further it is from the center of the fovea towards the peripheral vision. By taking the advantages of this property, a multi-resolution image stream is generated from the original high-resolution images of the Teleoperated system. This multi-resolution stream maintains the high visual quality at the center of the eye gaze, while lower spatial resolution images are used for the peripheral area as shown in
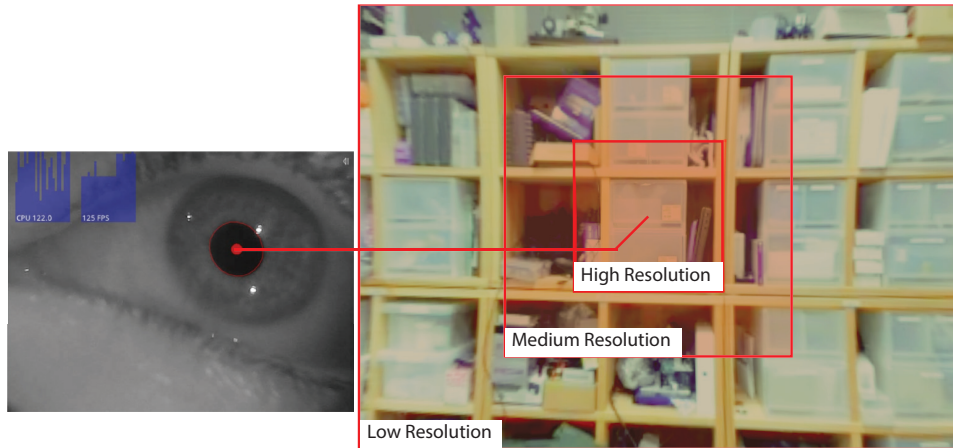
Figure 4.24: Foveated Streaming overview: multi-resolution regions based on eye gaze position.

Figure 4.24. Using this method, the image stream would result in a higher compression ratio based on the selected fovea size, and the encoding/decoding performance would increase according to that too. This paper is based on the previous research area of foveated rendering, and it contributes as follows:

- Providing real-time network streaming synchronization for eye-gaze using RTP stream.

- Quantitative and qualitative studies showing the effect of foveated streaming for performance and human perception.

The idea of using eye fovea as a driver to optimize the image spatial resolution has developed an important body of research in the area of multimedia and computer graphics optimization. In image compression, Wang et al. [103] proposed image coding system by taking into consideration the nonlinear decrease of spatial resolution in the human eye, and removing the high-frequency information from the peripheral area, thus improving the overall compression performance. Itti [39] used a different approach to address image compression by using saliency information rather than just the eye gaze position. And Ryoo et al. [88] proposed a video streaming service

based on eye gaze foveation to enhance network bandwidth. In computer graphics, Guenter et al. [29] used foveated rendering in order to enhance rendering performance by 5-6 times than direct rendering. Patney et al. [75] used foveated rendering in virtual reality applications to reduce the rendering cost, the work showed that foveated rendering did not introduce a significant difference in visual quality compared to direct rendering.

4.5.3   Encoding & Streaming

Foveated streaming pipeline is shown in Figure 4.25. This pipeline runs on the robot side, and the eye-gaze data is received from the user side. The pipeline generates a number of regions (N) which can be set in using system parameters. Region size in degrees is also determined by the Fovea Size parameter, in the conducted experiments, a value of 15 deg produced a good balance between streaming compression and perceptual awareness. The generated regions are combined into a single image frame as in Figure 4.26 (Streamed Regions). Then the images are compressed using the H264 encoder and converted to RTP stream to be sent to the user side. One important thing is to synchronize the eye gaze information in order to reproduce the same arrangements when combining the images on the user side. To do that, eye gaze information used for foveation is embedded into the first RTP packet of the image frame before being sent to the user. The main advantage of embedding this information into the RTP stream is to ensure exact synchronization of the data and the image stream when rebuilding the original image at the user side.

4.5.4   Unpacking & Presentation

User side receives the packed regions as an image stream, each frame contains foveation levels as was described in the previous section. When decoding the stream, first the eye gaze position is extracted from the first RTP packet of the packed frame and is used to arrange the regions and positioning them exactly as they were sent. After unpacking the images,
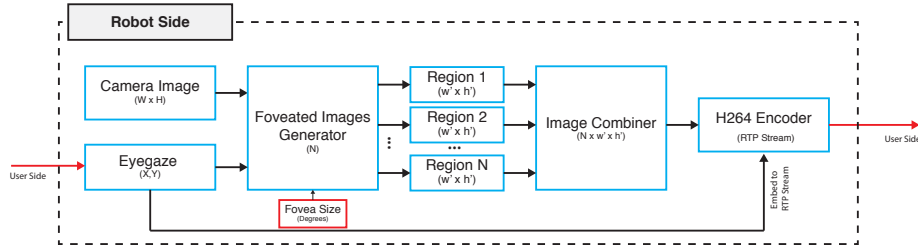
Figure 4.25: Foveated Streaming pipeline.

the regions are presented starting from the last region. Region number N is rendered into an offscreen image covering the entire field of view, next Region number N-1 is rendered at the original scale, and so on until Region 1 is rendered on top of them as shown in Figure 4.26 (Foveated Image). To mitigate the effect of change in resolution between two consecutive regions, a mask is applied to each region which helps to fade the edges and to reproduce similar effect of eye retina decrease of visual acuity, while maintaining high resolution at the center of the region as shown in Figure 4.26 (Region Masking).

4.5.5   Evaluation

To evaluate foveated streaming method, two types of evaluation were conducted:

- Performance Evaluation: to evaluate the overall performance of the system compared with full resolution image streaming, the bandwidth usage, and CPU consumption are tested.

- Qualitative Evaluation: a user study confirming the effectiveness of foveated streaming in maintaining the perceptual consistency of visual information.

Same system and hardware setup were used in both evaluations. USB3.0 camera module (Model no. See3CAM CU130 by e-consystems) were used to capture image stream, and connected to Intel NUC for video encoding and
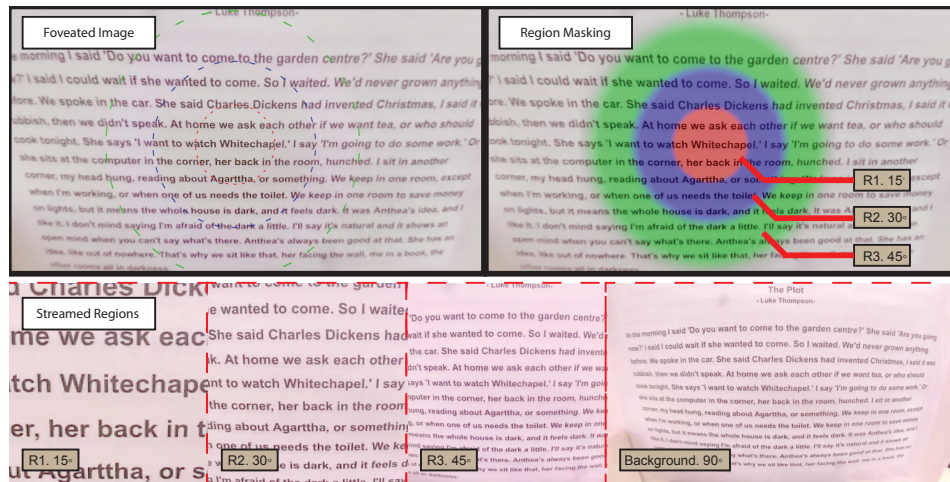
Figure 4.26: Foveated streaming decomposition (1) Foveated Image: final reproduced image at the user side, (2) Region Masking: masks used for combining the regions, and (3) Streamed Regions: streamed frame containing regions of foveation.

streaming over an Ethernet LAN to the user side. For media encoding and streaming, GStreamer library was used. The video stream is compressed using H.264 video codec provided by the library.

Performance Evaluation

The goal of performance evaluation is to measure the effectiveness of using foveated streaming network bandwidth, and CPU performance at the encoder side compared with full resolution streaming. In this evaluation, three different streaming modes were used: 640x480 pixels at 60 FPS, 1280x720 pixels at 60 FPS, and 1920x1080 pixels at 30 FPS. To maintain image quality among the images in both cases, adaptive bitrate quantizer set at 20% was used for the video codec. Foveated streaming parameters used were: 15 deg Fovea Size, and 3 levels of foveation. Figure 4.27 shows the quantitative results the proposed system. The results show the performance effectiveness of foveated streaming compared to full resolution streaming for both the bandwidth (compression rate varies from 5 to 8 depending on the
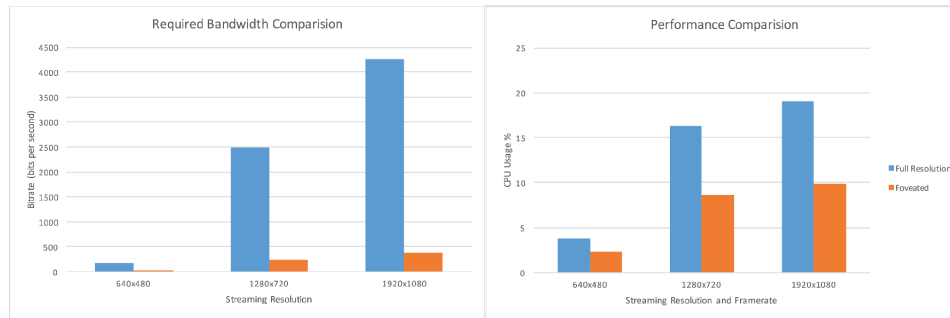
Figure 4.27: Performance evaluation of foveated streaming compared with full resolution streaming.

streaming resolution) and CPU load (about half CPU usage for foveated streaming).

Qualitative Evaluation

A user study was conducted to evaluate the visual perceptual effectiveness when using foveated streaming compared to full resolution streaming. The hypothesis is when using foveated streaming, there is no significant difference in reading speed compared when using full resolution streaming.

The study setup consisted of a three-axis telexistence robot head (providing pitch, yaw, roll motion) equipped with 90 deg cameras for video streaming. The cameras were set to capture 1920x1080 pixels at 30 FPS when streaming. For foveated streaming settings, fovea size used was 15 deg with 3 levels of foveation resulting a stream of 1528x320 pixels per image. An A3 paper containing one page of text (font used bold Arial at a font size of 24points, and 1.5 line spacing) placed in front of the robot at a distance of 30cm. Page was fixed that the first line of text is placed at the eye level of the robot. The total number of pages is 5, with an average of 225 words per page. The text used was a short English essay (The Plot by Luke Thompson). The user is connected to the robot over the network and uses an HMD to control the head motion of the robot and to perceive the visual information from it. The HMD is equipped with an eye gaze tracker (Pupil-lab for Oculus DK2) to measure eye movement when the foveated

streaming mode is enabled. In this setup, the user will be required to use both his head motion and eye motion to read the text from the beginning to the end of the page along yaw and pitch axes.
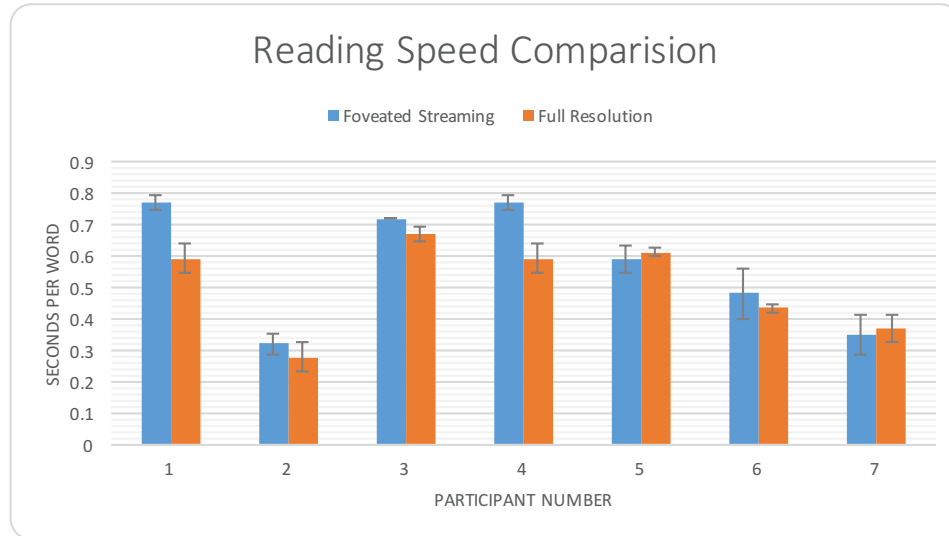
Figure 4.28: Subjective evaluation of reading speed for foveated streaming and full resolution streaming.

Prior to the study, participants are asked to fill a questionnaire containing basic information such as the age, eyes visual acuity for the left and right eye, and frequency of using HMDs in general. The participants are only informed of the task of reading an essay at one page a time and were asked to inform the experimenter once they finished reading the active page. After wearing the HMD, the participants are asked to calibrate their eye gaze using 9-points eye gaze calibration procedure, then a waiting gray screen is shown before starting reading. Once they inform their readiness to start reading, the user gets connected to the robot, and they start reading the text from top to bottom, left to right. After reading the page is finished, the waiting screen is shown again and the procedure continues to the next page until the 5 pages are over. The streaming mode for the pages is switched between full streaming/foveated streaming when changing the pages, and the first page's mode is randomly selected between both modes. Reading time is measured per page and stored along with the streaming mode for

performance evaluation. After the study is over, qualitative questions are asked about the reading experience and whether anything has been noticed while reading and switching the pages.

In this evaluation, 7 participants joined the study (6 males and 1 female with an average age of 24 ± 3) from different ethnicities. Overall feedback did not include any significant difference in the reading experience of the pages. Two participants reported seeing blurry words around the edges of gaze point (participant 1 and 4), this is due to the eye gaze tracker which could have been moved during the experiments (HMD was moved from the calibration position), the performance results show the impact of the calibration mismatch to their reading speed in foveated streaming. For performance results of the study, Figure 4.28 shows the average time per word for both full resolution and foveated streaming per participant. Overall reading speed can of both modes shows no significant difference between both modes. Reading speed varies per page, and overall the speed slightly increases when proceeding in the study, this can be due to the adaptation time of reading over a Telexistence robot. From these results, it can be shown that foveated streaming provided similar results to the full streaming under the condition of consistent eye gaze tracking.

Chapter 5

Meta-Modeling Evaluation

This chapter discusses the use of Embodied-Driven Design (EDD) meta-models to address body limitations in five different projects. Each of these projects highlights a specific scenario which EDD meta-modeling resolves. The four meta-modeling approaches described here are:

- Direct Body Transfer: Modalities are directly mapped from the body to the representation, used for functional body modalities (no alteration is required).

- Representation alteration: Different type of representation is used to overcome a physical limitation of the body.

- Topology Reconfiguration: Change the association and mapping of the modalities between the body and representation to overcome a physical disability, or to alter modality's function.

- Modality Expansion: Augment a sensory modality by using different sources of feedback from multiple representations, or expand modality's operation to multiple modalities simultaneously.

These four approaches are illustrated in Figure 5.1 along the association with the five projects discussed here. Each of these projects uses one or more meta-modeling approach to resolve body related criteria. Design considerations and implementation of these projects using EDD, along the conducted observations and results are detailed here.
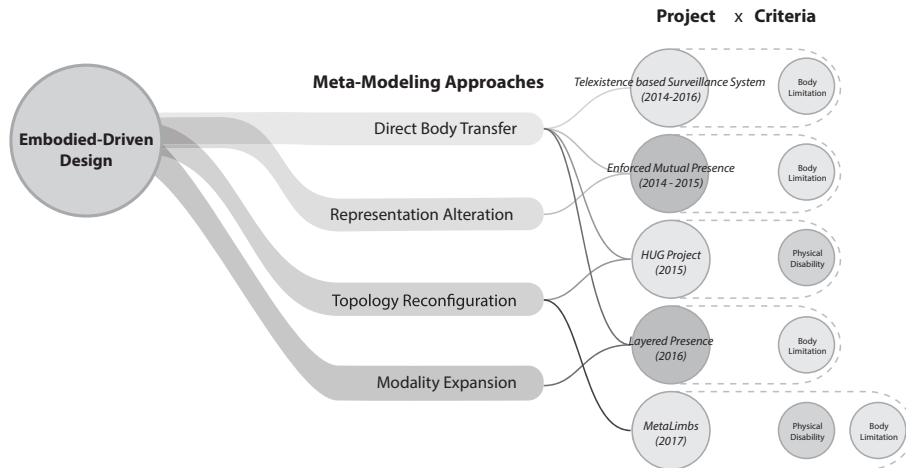
Figure 5.1: An overview of the four meta-models along the five projects discussed in this chapter.

5.1 Telexistence based Surveillance System

Teleoperation, in general, has begun as a mean enabling human to operate a sub-system from a remote and a safe place. The disasters that occurred during the course of the past 50 years (mainly nuclear and hazardous disasters) has sparked the need to build safe systems which experts can use and access the hazardous regions for maintenance or collect data and samples regarding the reason of the fault. Most recently, Fukushima Daiichi nuclear disaster in 2011 has resulted to launch several government-funded research programs to avoid similar cases in the future, and to build systems which the current teleoperated and AI-based systems failed to provide adequate access to the reactor. One of the most important requirements for such systems was the quick deployment to the scene to collect early information, and to allow expert people not used for teleoperation to quickly manage these systems with a minimum amount of training.

This project was established to address the previous requirements and
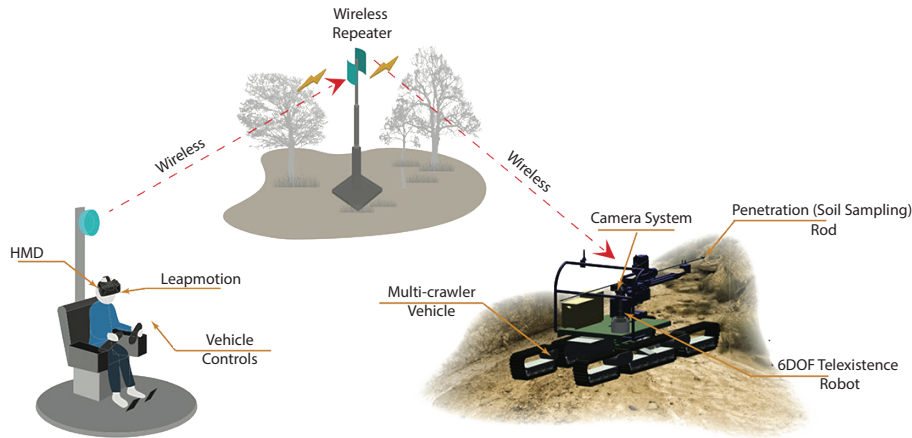
Figure 5.2: An overview of Telexistence based surveillance system.

need through a remotely operated navigation systems that can be deployed in the damaged location which human can not access. A collaboration between Keio University and Obayashi corporation started in September 2014 and continued until the end of 2016, funded by New Energy and Industrial Technology Development Organization (NEDO). During this collaboration, we developed an unmanned survey robot that enables remote investigation of the collapsed site. An overview of the developed system is shown in Figure 5.2 . The system mainly consists of a remotely operated vehicle through a telexistence based system located in the driver seat of the crawler, a penetration rod used to collect samples from the ground, a stable and long range wireless network access points and repeaters, and an operation cockpit which contains the tracking and control tools of the vehicle and robot, and the perceptual devices to deliver feedback from the representation. Each of these sub systems was handled separately by different corporations and research facilities. An overview of the general specifications for this system is listed in Table 5.1 . Our main contribution in this project is the development and deployment the operation cockpit as well as the telexistence system used in the crawler.

Specification	Details
Vehicle Dimensions	4340mm x 1740mm x 1650mm
Vehicle Net Weight	1900kg
Image Acquisition	Stereo camera system + rear omni directional camera
Vehicle Driving Mechanism	Operator hand held controller
Visual Presentation	Head Mounted Display + screens for external monitoring
Stereo Visuals Resolution	480p@60FPS - 720p@60FPS
Operation Mechanism	Head motion (position + rotation)
Visual/Operation Latency	Less than 150ms
Wireless Communication	2.4 GHz wireless LAN
Communication Distance	Up to 2 km (with a wireless repeater)

Table 5.1: General specifications for the surveillance system.

5.1.1 Design of Remotely Operated Representation

For the development of the telexistence robot, the challenge was to maintain high degree of operation of the robot in terms of both performance time and quality of operation in the remote location. Vehicle driving was a part of the tasks to be operated, thus 3DOF telexistence robot is not sufficient since parallax motion is required when driving. To achieve such requirement, a custom made 6DOF telexistence system was developed (mechanical design by KAWABUCHI Mechanical Engineering Laboratory, Inc.) as shown in Figure 5.3 . The robot operation system follows the same protocol for communication and control as the one designed for the toolkit (described in Chapter 4), thus the visual, auditory, and body control modalities are compatible with EDD framework requirements.

The system development and testing took several iterations and run-
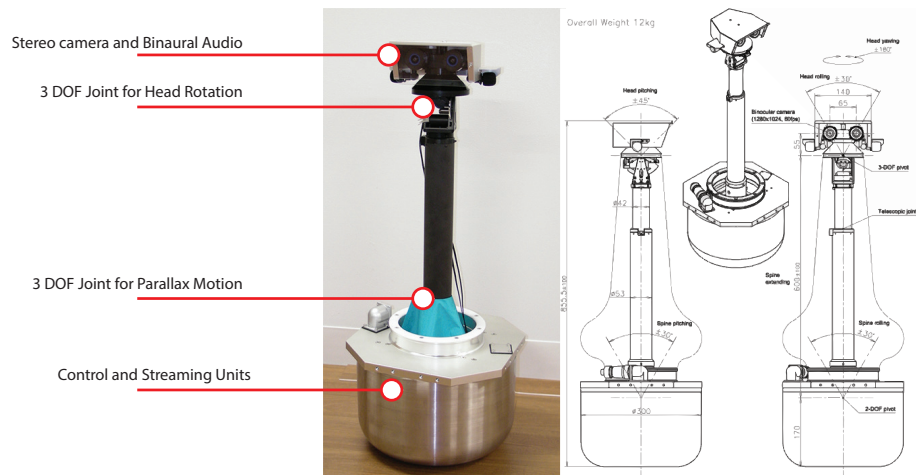
Figure 5.3: 6DOF Telexistence system (TORSO).

tests to evaluate the degree of control of the robot. Initial experiments to
operate the vehicle is shown in Figure 5.4 . The operator vision is mapped
to the robot visual feedback, and his head motion controls the representa-
tion movement. The modality data flow is managed by EDD framework.
The evaluation layout for the test field of the initial experiments is shown
in Figure 5.5 , and located at Obayashi corporation machinery factory in
Kawagoe City. The operator had to navigate through set of land marks
(cones) without touching them while driving. The vehicle used was a back-
hoe that has a dedicated remote control panel.

Figure 5.4: An initial set of experiments of using remote operation.

Total of 40 users operated the system, 35 out of them had no previous

experience at driving the backhoe. Before operation, the users would be taught how navigate the vehicle using the remote controller. Afterwards, the pilots would access the remote robot via HMD and drive through the field. As a result of the tests, we found that the pilots were capable to maneuvering precisely through the landmarks since they had a correspondence mapping between the visual modality and the robot stereo cameras, thus the depth perception was correctly matched. Also, the low latency transmission played an important role at reducing the visual sickness of operation, and at precise decision making when making turns. Pilots even used to see through the mirrors located inside the backhoe to have wider awareness of the location, cones, and walls nearby, just as if they were in the driving seat.
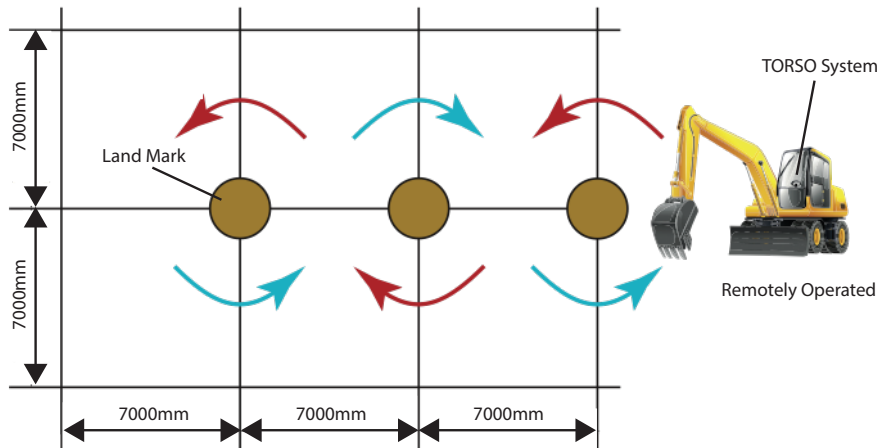
Figure 5.5: Remote navigation using TORSO system.

The next stage of the project was carried to deploy the robot into the crawler vehicle. The vehicle operation differs from backhoe operation in terms of scale and operations that should be conducted using it. The vehicle was designed to operate in terrain environment, and under various weather conditions (rainy, windy, ... etc). The finalized system after being integrated with the vehicle is shown in Figure 5.6 . The operator is provided with two types of operation: using the HMD for driving and inspection, and external screen connected to an omni-directional camera mounted on the top of the vehicle to gain a global view of the vehicle and the inspection equipment.
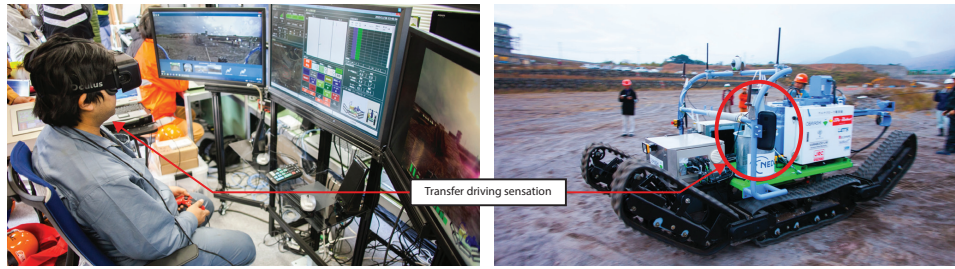
Figure 5.6: Second phase of remotely operated driving system.

To support the operation in large-scale environments, a virtual navigation system was developed Figure 5.7 that uses the concepts of virtual presence to replicate the same experience of driving the actual vehicle, and which communicate in real-time with the physical system. Various point of views can be used in the virtual environment which would provide a higher degree of flexibility in navigation. A photogrammetry based model of the actual location is generated offline (developed by a third-party) and rendered in the virtual space. The vehicle coordinates are captured via a GNSS module (u-blox Neo 7) embedded into the vehicle and mapped into the coordinates of the virtual environment. This transitional environment between the first point of view (robot side), the third person of view (simulator side), and bird point of view (simulator side) would expand the spatial awareness of the operator beyond what traditional teleoperated systems can provide.

Figure 5.7: Virtual navigation and operation platform.

5.1.2 Telexistence System Meta-Modeling

In this project, direct body transfer meta-modeling approach were used to map operator's body with the used representation. Figure 5.8 shows the target meta-model and body mapping between the user and the robot. Since the system is intended to be operated same as being physically in the vehicle, and by operators without physical disabilities, the modality mapping is direct. The head modality contains spatial information of user's head (position and rotation), which is mapped to TORSO robot. The robot maintains same coordinates with the user, and in return provides the visual (stereo images) and auditory (bin-aural) information. Such mapping directly matches the operator with the robot.
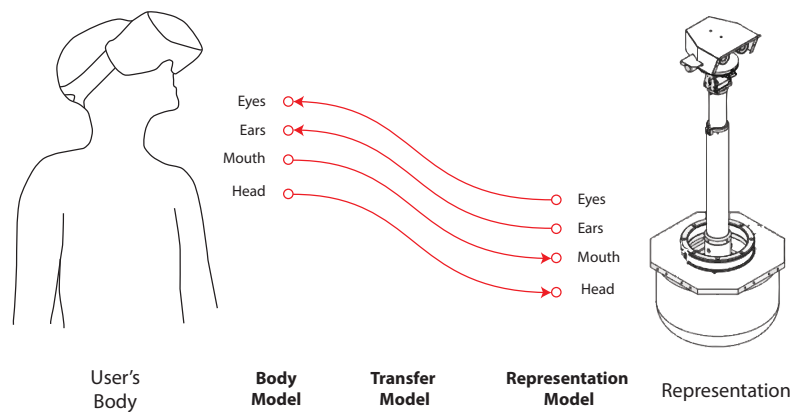
Figure 5.8: An overview of the target meta-model for Telexistence type systems.

The actual meta-model developed is shown in Figure 5.9 . The meta-model contains two types of blocks: Body blocks (HMD Node, Body, Microphone, and User Node), and representation blocks (Robot and Simulated). For this particular design, no transfer blocks were used.
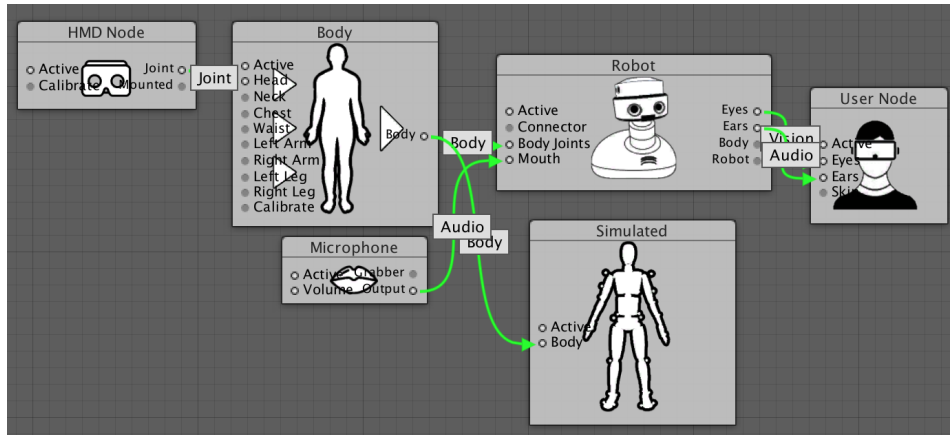
Figure 5.9: A screen-shot of EDD Meta-modeling editor describing the association user's between body and two representations: physical and simulated.

5.1.3 Discussion

This project was an evidence that the presence of certain modalities (postural, vision) in the representation is sufficient to mediate body state in a bi-directional manner. In scenarios which the operation of the system is required to be immediate without any body reconfiguration or change in operation, direct body transfer can be used. This maintains the quality of embodied use of the representation. Compared with telepresence type of systems, the subjective evaluation it showed that the use of such body mapping with a representation has enhanced the quality of navigation and spatial understanding at the remote sites.

The digitization of body and operation and the use of mediated representation can also leverage the possibilities of what a body can perform, can sense, and can operate. For example, when the simulator environment was deployed, sensory data from drones were possible to integrate within the same proxy the operator used in his operation of the robot. Initial tests were also conducted to integrate Enforced Presence into the operation so the operator could see his hands and controller. The operators did report they had better awareness and less motion and visual sickness, however in

later stages of the project, virtual hands were removed due to some design constraints of the cockpit.

5.2 Enforced Mutual Presence

Enforced Mutual Presence uses representation alteration blocks to compensate non available physical representation using a virtual one. This use-case leverage the current telepresence systems which lacks the physical attributes of presenting human body, in which the user fails to observe his body being immersed in the teleoperated robot side. Also, as an extension of this work, a mutual body presentation is used to compensate the remote body representation, so the remote participants would be aware of the state of user's body and his actions. Enforced Mutual Presence addresses the following points:

- Users body representation in a physical environment.

- Preserving body ownership during teleoperation.

- Presenting users body visuals to remote observers.

Three design aspects [8, 13] contributes in achieving mutual presence and the awareness of being presented in a specific environment: spatial presence, self presence, and social presence. Spatial presence is to be capable to interact with the environment. Self presence can be described as to be able to observe our bodies being presented in the environment, and aware of its posture at any given moment. The social presence is how we observe social interaction with other people within one environment.

To address the first key point "spatial presence" the user should have direct access to the environment as if he is located there, with the freedom to move and navigate via an alternative representation of his body. The user should be able to move freely and independently his head and body, and according to that, the slave robot should follow and update users visuals of the remote place. We avoid using any physical or tangible controller (such

117

as a joystick or keyboard) to control motion and rotation speed of the slave robot. This is important because if the user is aware of the presence of a physical controller, then the coherence between the local and remote place will break. So an intuitive and natural interface is required to maintain spatial coherence. Body as a joystick concept was used to drive the robot spatial motion in the remote place. The concept uses body motion (leaning forward,backward, and sideways) to control the motion vector of the robot. This motion correlation helps to maintain the same motion/optical flow between the body and the visuals observed remotely.

The second point "self presence" is the fact the user should have physical awareness of his bodys presence. The user validates his existence in a specific place by observing his bodys visuals as he expects, maintaining the ownership relation with his body. In this work, we found that observing body visuals is an effective factor to maintain the seamless sense of presence for the user, so image-based method is developed which captures egocentric images of users body visuals, and superimpose it into the remote place.

The final point addressed is "social presence". In order for the user to communicate effectively with other people in a different location, mutual communication between both sides should be maintained. As in "spatial presence" the user is aware of the surroundings and people around, those people should be capable to understand what the user wants in return. It is commonly to user an LCD panel only to show users body, however this method is not capable to provide spatial interaction in the 3D space. As an alternative, we propose to project users body visuals in robot side, so the user can visually affect in the remote place, allowing remote observers to visualize his body.

Figure 5.10 shows a physical telexistence system combined with projected body visuals of the operating user. Figure 5.10 (A) shows what the user sees in the remote place while using his hands. The body visuals are preserved and projected back into his field of view in a natural and correspondent manner to his actual body. While Figure 5.10 (B) & (C) shows what the remote participants observe: an extension of operator's body vi-

118
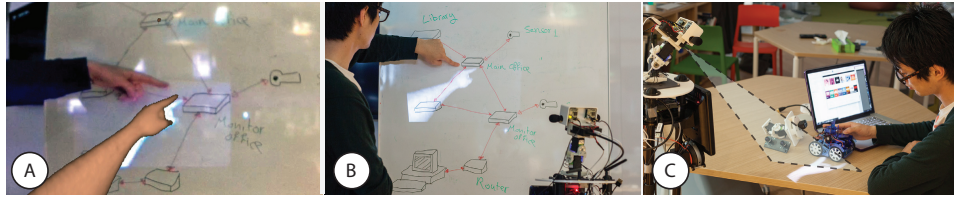
Figure 5.10: Enforced Mutual Presence overview (A) the first person view for the user showing superimposed virtual hands, (B) user's hands are projected externally so others can mutually see, and (C) virtual hands are projected on an arbitrary surface.

sual projected into the remote surfaces making it easy to understand what the user is pointing at. This design helps to overcome body physical limitations of the representation as well as increase sense of bodily presence.
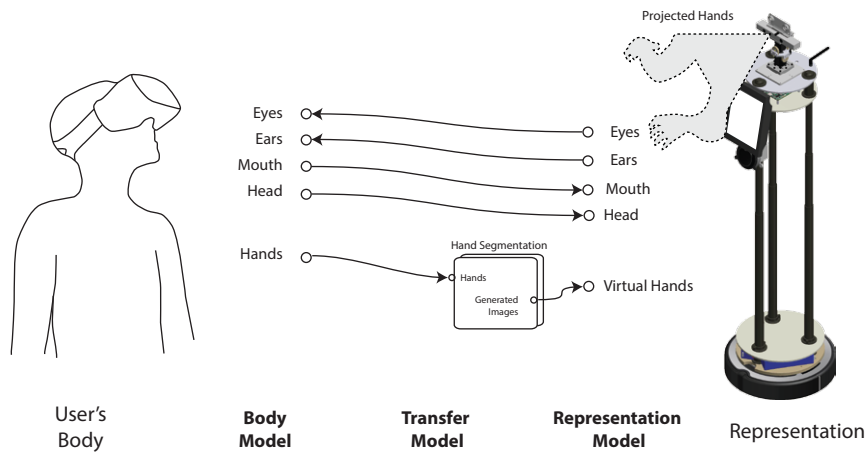
Figure 5.11: An overview of the target meta-model for Enforced Mutual Presence.

To reflect such considerations using EDD design, two design approaches were used to address the intended meta-model:

- Direct body transfer: used to map the available modalities in the representation (Eyes, Ears, Mouth, and Head).

- Representation alteration: used to compensate the missing arms of the representation by changing the mapping of user's body model

119

(Hands) into an alternative hands representation (virtual hands) using
transfer model block (Hand Segmentation).

An overview of the meta-model that combines both approaches is shown
in Figure 5.11 highlighting the main modality flow between the body and
representation. The process for developing the components of this model is
discussed in the following sections.

5.2.1 Body Representation

In this scenario of body representation alteration, a hybrid representa-
tion was used that composes a physical form providing the necessary el-
ements of spatial presence: visual/auditory feedback and navigation, and
virtual form that reflects user's body into the physical representation that
lacks the arms. Although when operating the system with single physical
representation (without the virtual arms) the user gains sense of presence
awareness via the visual motion perception provided by the robot head used
(that is sense of agency), however the lack of body visuals presentation de-
creases the subjective sense of ownership of the operated system.

In the user side, the hands are captured using an IR camera mounted on
the front of the HMD. The camera provides 110 deg field of view which covers
HMD FoV, and thus it is possible to capture user hands with no cropped
areas. Though the resolution of the cameras are relatively low (640x240),
up sampling step is necessary to smooth out the edges. The advantages of
using IR camera compared with RGB camera is the possibility to capture
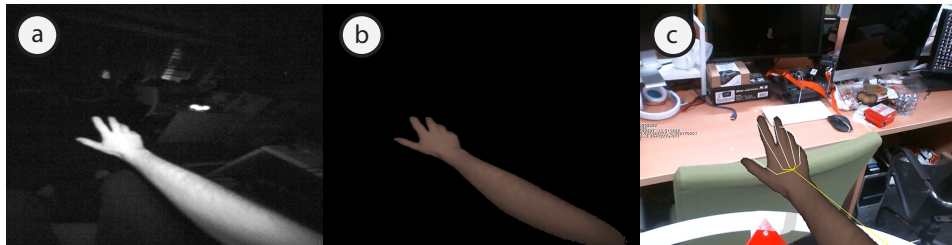
Figure 5.12: User's virtual body image segmentation and composition
into the physical representation.

objects close to the camera using the returned intensity, in this case, hands and body usually maintains close distance to the HMD, and thus their visuals are captured effectively. However there is a resulting noise from the background. A nonlinear filtering function is applied on the captured images, this function removes the pixels which color intensity are below a certain threshold:

$$Filter(P) = \begin{cases} P^{\frac{1}{Gamma}}, & \text{if } P^{\frac{1}{Gamma}} \geq threshold \\ 0, & \text{otherwise} \end{cases} \quad (5.1)$$

$$P \in [0, 1]$$

Figure 5.12 shows the procedure to extract and segment arms visuals from the background. (a) is the raw data captured by the IR camera, a noticeable amount of background noise is presented in the captured images. (b) shows the process of filtering the noise from the background. A color correction is also used to tint and match the visual appearance of the hands with the human skin color. And finally (c) is the super imposed hands image into the remote visuals after applying perspective correction to match the capture FoV from the cameras with FoV of the HMD. The user would observe the hands being placed correctly within his field of vision corresponding to the actual placement of his body.

5.2.2 System Diagram

The overall data flow and main components of the system are shown in Figure 5.13. The robot side follows the design aspects of the toolkit described in Chapter 4 . An additional component was added to this system at the representation side, that is a pico-projector mounted on the head of the robot which is used for hands displaying and projection. User hand images are streamed remotely to robot side, and are projected from robot's point of view using a pico projector mounted on its head. Those hands are aligned with user hands position and motion, and allows the remote participants to see the gesture of his hands. Figure 5.10 (B) & (C) shows
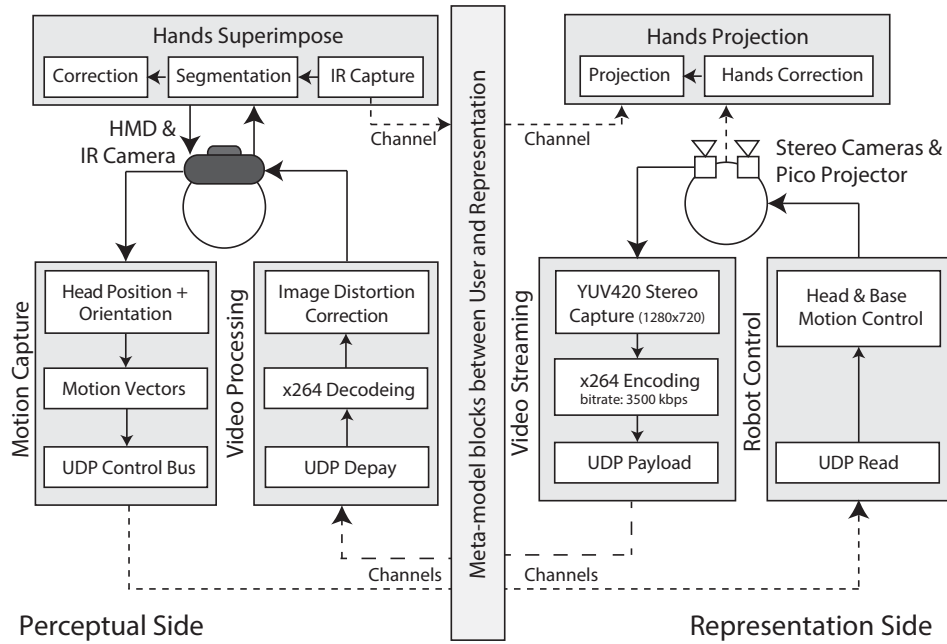
Figure 5.13: System data flow between user, representation, and the used meta-model.

the hands being projected on a trivial surface, in which user hand gesture can be seen remotely. Depending on projector's lumens, the hands might be difficult to see in a well lit room. In current implementation, a 65 lumen projector to render the hands was sufficient in dim lighting condition.

5.2.3 EDD Meta-model

The meta-model for this type of representation alteration system is shown in Figure 5.14 , which reflects the mediating layer in the previously shown diagram Figure 5.13 . This model uses two different representation meta-blocks for the virtual body representation: **Hands Display** block to provide visual feedback of user's hand locally, and **Hands Projection** block to project the body visuals remotely on the robot side. **Hands Capture** block handles the visual captures, and produces details about the Hands (hands images provider) that can be used to segment and extract hands information locally and remotely. These information are provided to the

robot side for projection, and to the user to be super-imposed locally. The flow of the information is highlighted by the green arrows. Using this meta-model, it is possible to alter the hands display or projection using transfer blocks if needed (changing hands model used, hands attributes such as color and length, and even the number of fingers).
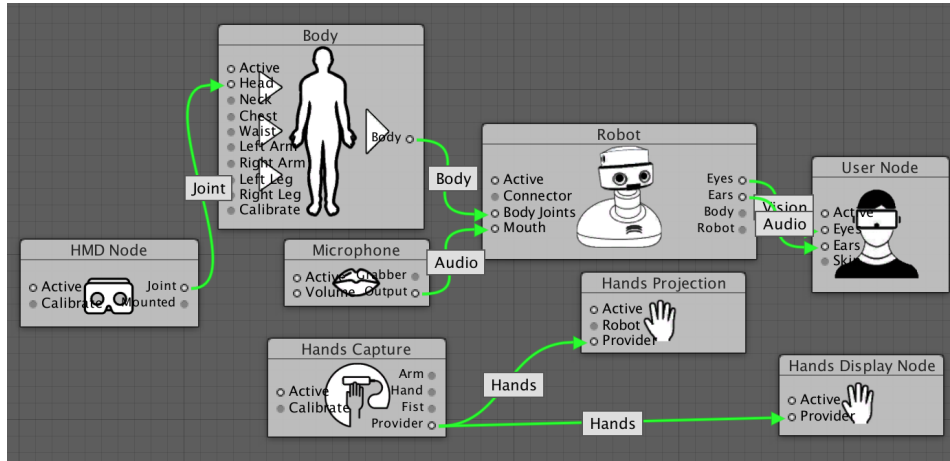
Figure 5.14: Screen-shot of Enforced Mutual Presence meta-model using EDD editor.

5.2.4 Evaluation of Presence

To evaluate the effectiveness of representation alteration, Enforced Mutual Presence meta-model was used to add virtual hands to the representation, and to provide the user visual feedback of his body visuals being rendered remotely.

Hypothesis: The addition of the virtual hands increases the sense of presence for the user, and enhances the engagement of operator in the remote place during the operation.

Experiment Setup: The experiment was conducted in a living room setup. The room was divided into tow parts. Avatar navigation room was well lit, and contained several objects, tables and chairs as can bee seen in Figure 5.15 (a). This setup reflects one of the intended applications of this system, the use of such platforms into our daily life environments. Several

participants did not have a previous knowledge about the layout of the room. The experiments were conducted individually for each of the participants. The participants were situated in an isolated small room that contained the tools for controlling the robot Figure 5.15 (b). Tracking devices for the head and hands are provided in the HMD. The avatar system is connected over a wireless network with the user side.

Participants: For this study, 10 subjects joined the experiment (8 Males and 2 Females). The age range for the participants is 22 to 32 years (Mean: 26 - SD: 3.6). Basic information were collected from each participant at the beginning of the experiment, 3 out of 10 reported they are unable to see stereoscopic contents using virtual reality devices due to their eyes condition.

Procedure: The experiment was divided into the following cases:

- Case 1: Operate the robot without hands presentation (enforced presence meta-model not used).

- Case 2: Operate the robot with hands feedback (enforced presence meta-model used).

To avoid evaluation bias, the participants were divided into two groups, each group experienced the system in a different case order from the other. Group A experience case 1 followed by case 2, while Group B experience case 2 followed by case 1. During the experiment, we observed user's behavior

Figure 5.15: Representation alteration evaluation room setup: (a) Avatar navigation room, (b) user operation room.
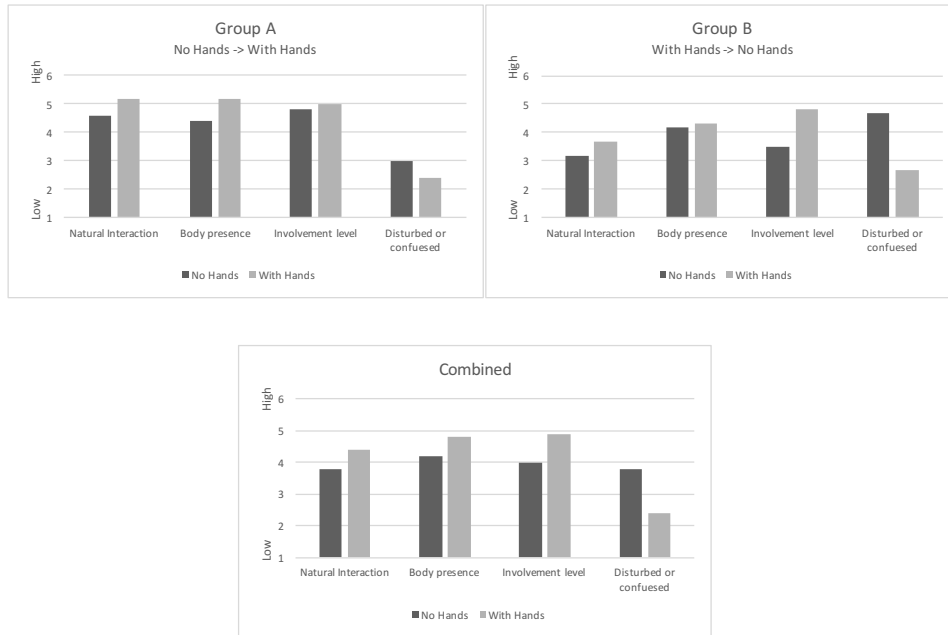
Figure 5.16: Representation alteration user study for Group A and B, as well as the combined results for both groups.

when the hands were enabled and disabled. A questionnaire was asked after each case to report their subjective experience of the system: interactivity, body presence, involvement level, as well as disturbance or confusion while using it.

Results: Figure 5.16 shows the results after evaluating Group A and Group B, and details of groups answers are listed in Table 5.2 . It can be noticed that the presentation of virtual hands helps to enhance the state of operation, and reduces the confusion about body presence. The data shows that the lack of hands in both conditions resulted higher deviation in the measured data mainly for level of disturbance and confusion. Group A which experienced the system without hands at the beginning showed a noticeable increment in body presence and natural interaction, and reported their excitement when they saw their hands being "available". When the hands was presented, participants were using their hands more actively for pointing, gesturing and communication. For Group B they started using the system with their hands being presented, several participants found it rather

125

	Group A			
	No Hands		With Hands	
Question	**Mean**	**SD**	**Mean**	**SD**
Natural Interaction	4.6	1.52	5.2	0.84
Body presence	4.4	1.14	5.2	0.83
Involvement level	4.8	0.84	5	0.71
Disturbed or confused	3	2	2.4	1.67

	Group B			
	No Hands		With Hands	
Question	**Mean**	**SD**	**Mean**	**SD**
Natural Interaction	3.17	1.33	3.67	0.52
Body presence	4.17	1.33	4.4	0.52
Involvement level	3.5	1.38	4.8	0.75
Disturbed or confused	4.67	1.03	2.4	1.52

Answer Range: 1: Low - 6: High.

Table 5.2: Questionnaire results for Group A and B with and without hands representation.

trivial for their hands to be there where they expect them to be. In contrast with Group A, when the hands were removed in Group B, they reported less involvement and more confusion about the operation (with hands 2.4 ± 1.52 compared to no hands 4.67 ± 1.03). Some participants expressed their feeling as "taking away part of their body" during the experiment.

From the previous experiment, it can be shown that the body presentation (despite being virtually substituted) helps to increase the overall sense of presence and minimizes the disembodied feeling of lacking some parts of the body (level of confusion). Using representation alteration meta-modeling design, it is possible to overcome such limitations of body presentation.

5.2.5 Public Demonstration & Discussion

The system has been demonstrated publicly in three major international conferences (SIGGRAPH Asia 2014 Emerging Technologies [91], Augmented Human 2015 [90], and ICAT 2015 [89]) in which attendees experienced several variations of the system (addition of haptic feedback) as shown in Figure 5.34 . The core concept was maintained in which the virtual representation of the body was added to augment the perception of presence, and to maintain the sense of body ownership. The benefits of this system were not just the enhancement of bodily presence, but also extended into social and usability situations.

Although the projected hands did not serve to provide any actual physical actions, such as moving, grasping, or manipulating other physical objects, however, they acted as an important social cue. The body visuals transfer has played a role of being a medium to communicate and transfer the fact of a human being in operation, not just a robotic/physical medium that used to capture and transfer information. The ability to use non-verbal communication to infer the need or the state of the user in operation to the remote participants.

The other advantage of using such hybrid system in which a non-physical component incorporated into was to overcome the physical constraints. The users found it much easier to express what they needed by just pointing their hands into the objects even in far locations. The projected hands served as a medium that expanded their reachability beyond the physical bounds of their real arms. This shows the advantages of using a mixture of virtual/physical representations to overcome certain constraints and bounds inherited within the human body.

127

Figure 5.17: Demonstration of Mutual Enforced Presence system: users interacting through virtual hands with others despite the lack of physical representation of the hands.

5.3 HUG Project

Overcoming the physical limitations of human body was one of the goals of human augmentation research. Physical limitations does not have to be just a form of a lost modality, but any physical constraint could be considered as a limitation of the body. *HUG Project* demonstrates a scenario

in which EDD meta-modeling was used to resolve a physical disability of a 90-year-old lady. The project has started as a collaboration with Ducklings Japan [21] in late 2015 to design a system which can provide the accessibility for the lady to participate in a remote wedding ceremony of her grandson (the CEO of Ducklings inc.). The project lasted for almost three months (August - October 2015) in which the remote presence system had to be tested and used at the wedding ceremony. The lady was hospitalized with a physical disability of her body that restricted her neck motion significantly.
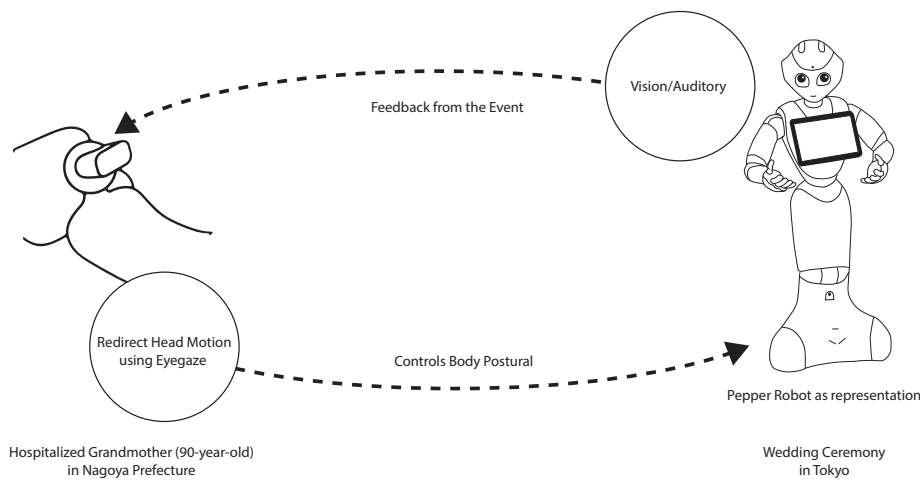
Figure 5.18: HUG Project design requirements overview.

An overview of the discussed scenario is shown in Figure 5.18 , in which the disabled lady uses a remote representation to access the wedding ceremony. In such situation which a physical disability is presented to the user, if a *Direct body transfer* meta-model was used then the same modality disability will be reflected in the representation. To address this, *Topology reconfiguration* meta-model is used to change the postural topology of the body schema. For this particular scenario, Figure 5.19 shows the target meta-model of this system in which the eye gaze is used to substitute head rotation (disabled modality) of the representation. Other modalities can also be used, however, the eye gaze was a suitable input for the head rotation to minimize the number of tracking devices attached to the lady.

129
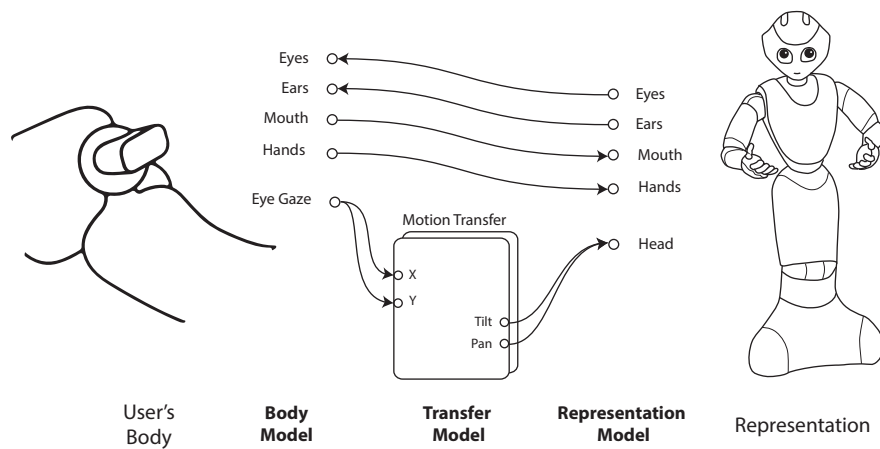
Figure 5.19: An overview of the target meta-model for modality substitution in HUG Project.

5.3.1 Meta-model Design

Initial tests were conducted using the toolkit discussed previously in Chapter 4 , however, it was not sufficient to be used in a social event such as the wedding ceremony. Thus an alternative avatar system was required. To resolve this, Softbank Pepper robot was used as "body host" which just mediates the postural schema of the grandmother (hands, heads motion), and a custom stereoscopic camera module was developed and embedded into the robot to provide the visual sensory mapping. Both postural and sensory modalities were integrated into the EDD framework to be used for interaction modeling.

Since the lady was not capable to use her head motion to operate the robot, an alternative operation modality was required to use instead of her physical head motion, or else the representation will also inherit the same physical limitation of the user. To resolve this, the eye gaze is used as an input modality to the system and was mapped into the head rotation of the representation. Thus the lady had full capability to rotate the head of the robot using two axes:

- Tilt axis: mapped with the vertical motion of eye pupil.

Eyegaze to Angle Mapping function
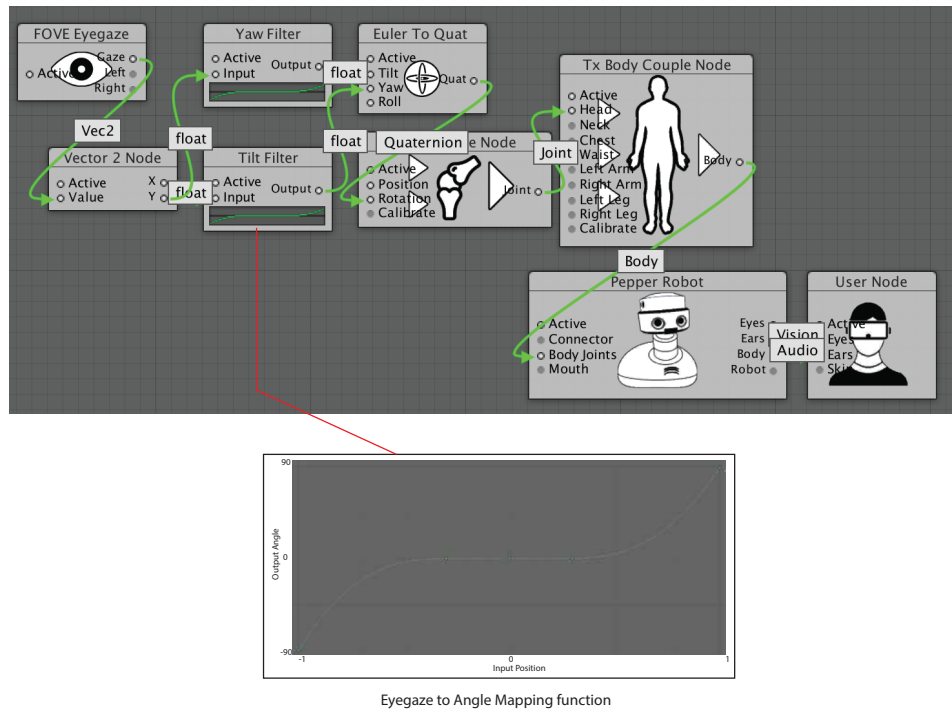
Figure 5.20: A screen-shot of HUG Projection meta-model using topology reconfiguration.

- Yaw axis: mapped with the horizontal motion of eye pupil.

FOVE HMD was used to capture the eye gaze movement converted into normalized 2D coordinates, and to provide her the visual feedback from the wedding. The eye gaze modality is decoupled into horizontal and vertical motion, which are mapped into rotation angles using Tilt and Yaw filter blocks. These filters acts as non-linear functions which are defined for
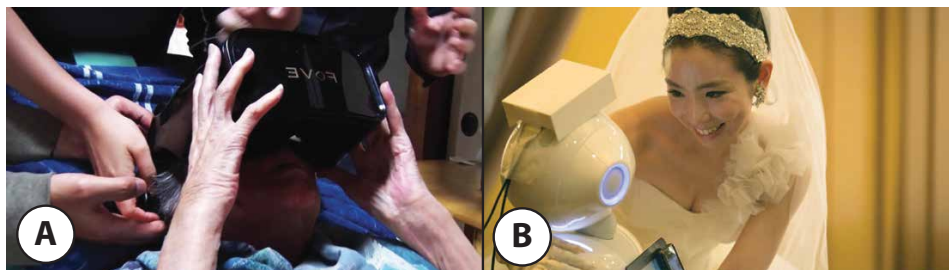
Figure 5.21: HUG Project: eyegaze substituting head motion.

131

each axis separately, and used to remap the motion of the eyes [-1,+1] into rotational values in degrees [-90,+90]. The chosen function (designed as set of key points) sustains the rotation within limited range of eye gaze motion, and provides non-linearity mapping along the edge to start rotating the head. The results of the angles conversion function are coupled to a rotation block (as Euler angles) and mapped to a joint which would correspond to the head joint of the robot. The overall EDD meta-model used with the transfer blocks to achieve topology reconfiguration is shown in Figure 5.20 . This EDD model acts as a high level design tool that can be easily manipulated and alternated to remodel the control mechanism and body structure.

The final results of this model were sufficient to produce sense of agency toward the robot although the control mechanism does not match how our body moves. This result was concluded based on the feedback received by the lady in which she expressed her feeling of being physically in the ceremony and her active involvement when interacting with the remote attendees Figure 5.21 .

5.3.2 Discussion

This scenario had showcases that the concept of operation alteration could benefit wide spectrum of physically impaired people to gain back the sense of the lost modality by re-configuring their body schema.

Several reasons could have contributed in her immerse with the system and biased her experience to be as "believable" sense of presence despite being aware that she was using mediated tools and she was not physically there in her own body. Psychological factors such as the emotional attachment toward the groom played an important factor to build up the social context of the experience. The second factor is the use of an immersive display (HMD) for the first time affected her sensory and perceptual experience of involvement. The sensory matching of her visual feedback (stereo vision camera module), and the correlated control of the robot head via through her eye motion created a closed loop system in which she can drive,

and gain feedback. This loop created a strong sense of agency toward the artificial body. Lastly, the social context in which the remote participants at the wedding ceremony expressed affection toward her avatar system. The use of human factors in the designing the experience creates mutual social involvement in which the observers can understand that the person within the shell of the robot is actually a human, which is expressed by body shape and canny motion.

Topology reconfiguration design approach would result losing the function of the source modality which is used as for the substitution.Although this argument is sound, it is not true all the time. Since in most of the time we do not use our entire body muscles for operation. For example, when seated, our legs usually at rest all the time. The decision of the body configuration is totally driven by the embodied designer which can be the user himself, thus the user can reconfigure his body schema depending on the scenario and state of his body (like body reprogramming/rewiring). This approach is effective in a sense that the users can adapt to the new representation by deploying an existing modality they are accustomed to. The plasticity of our body helps in temporarily reconfiguring our postural mapping and control of the new schema.

5.4   Layered Presence

Our bodies are inherently limited to operate or perceive to/from a single point of interaction. For example we see from a single location which our body is presented at, similarly our actions follows the same physical constraints. The use of Telepresence and Telexistence robots for example expands our spatial awareness, however the same limitations are also presented in such type of systems. To challenge such body limitation, Layered Presence provides an approach to expand our modalities function and from one-to-one mapping to one-to-many mapping.

The proposed system addresses the previous mapping point by using multiple Telexistence robots (based on the embodiment toolkit) that are

Figure 5.22: Layered Presence: combining visual feedback from two different locations.

synchronized with users motion, and provides real-time visual and auditory feedback to the user. Robots are represented as layers of perception (originally discussed in Section 3.4.4), in which the information represented by each layer can be visual, auditory, or haptic feedback. These layers are blended based on the saliency found in them, and presented to the users feedback displays. Users eye gaze is used as the main interaction modality for layers information presentation. Eye gaze is tracked and used to identify the target layer to be highlighted among the other layers based on the saliency information dominance. The layer user is looking at becomes focused while the other layers are defocused using an artificial depth of field effect, that maintains the optical flow in the peripheral vision of the user. Figure 5.22 shows the proposed system in action in which the user perceives two different locations simultaneously. This system is addressed using EDD meta-modeling as illustrated in Figure 5.23 . The meta-model reflects the idea of modality expansion by mapping body's modalities into multiple representations. Layered Perception transfer block is responsible to combine/expand the modalities inputs/outputs between the user and rep-

resentations. Details about the design and implementation of this transfer block are described in the following sections.
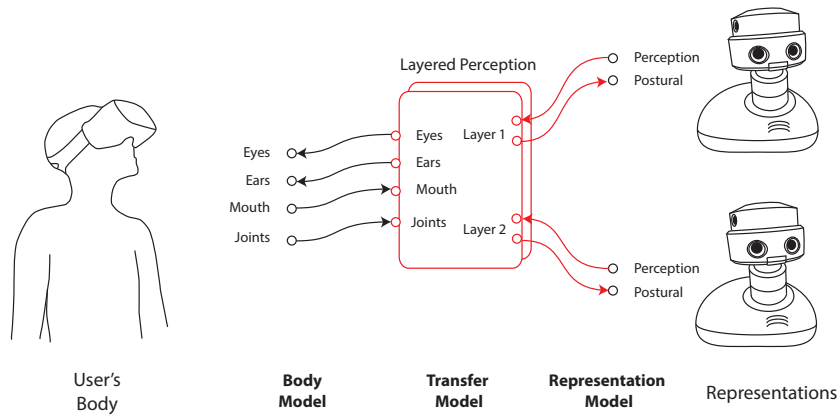
Figure 5.23: An overview of the target meta-model for addressing one-to-many body mapping situations.

5.4.1 Design overview

To realize the idea of overcoming the limited body perceptual awareness, a Layered Presence system has been developed using EDD framework. The system uses Telexistence robots (based on the embodiment toolkit) that are synchronized with users motion, and provides real-time visual and auditory feedback to the user. Figure 5.24 shows an overview of Layered Presence data flow from multiple representations. This interaction uses the eye gaze as an input to the system to determine which spot the user is looking at, and based on this input, the system controls the focus of the layers to the corresponding layer user is looking at. To determine the candidate layer that should be in focus, visual saliency maps are generated for each layer that contains the "interesting" information from the corresponding representations. Such blending mechanism provides seamless visual feedback of the multiple layers, and maintains the awareness of multiple locations.
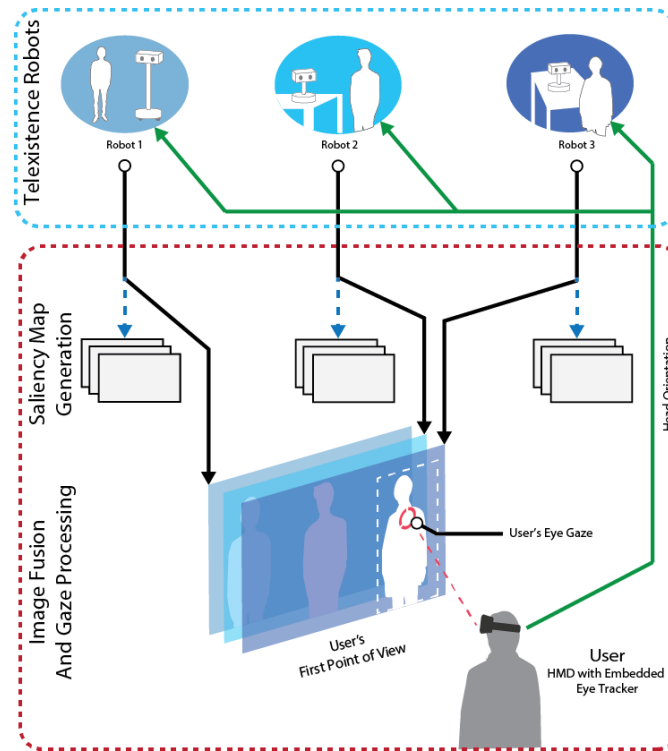
Figure 5.24: Layered Presence: perception flow from multiple avatars.

5.4.2   Saliency Map Generation

Saliency maps in this method are responsible to represent the presence of remote participants as a weight map generated for each captured frame while taking to consideration the temporal factor of the frames. The process of generating the saliency maps is done by a combination of two image analysis and features extraction methods. First the layers are processed for human presence, the procedure is done by applying Haar cascades classifier on each captured frame and the results are rectangles set representing the detected faces regions. These rectangles are expanded proportionally to their size to cover user body size (manually tuned, Width factor: 200%, Height factor: 400%). In practice however, using facial detection only fails to provide continuous tracking of presented people for several reasons such as partial occlusion of the face, lighting conditions, and resolution of the captured images. Also relying on facial detection only limits the visual
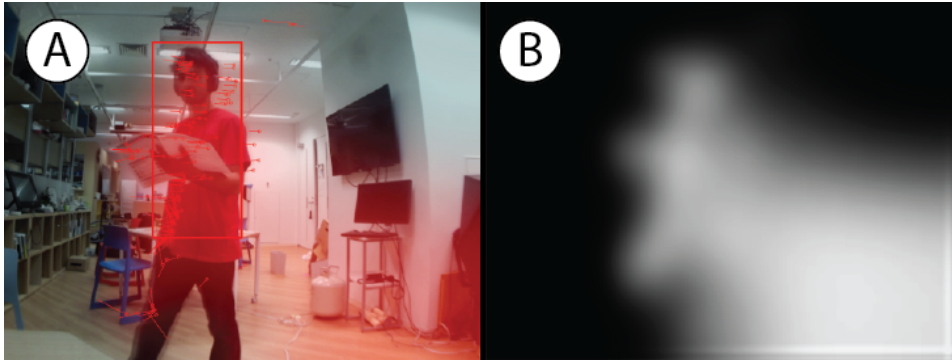
Figure 5.25: Saliency maps generation for the layers.

saliency maps to capture the information of moving objects in scene. We addressed this limitation by adding a second layer of tracking using optical flow detection in the layers. Lucas-Kanade method was used to track scene features points for changing. When local changes occur in the layers, motion vectors are recorded for later registration in the corresponding saliency map of the layer. Figure 5.25 (A) shows the detected motion vectors.

Next, the saliency maps are filled with weight of 1 for the pixels corresponding to the registered feature points and facial regions for each layer. To avoid the presence of hard edges along the detected regions, a Gaussian blur filter is applied to the saliency maps. To assist feature tracking consistency over time, a temporal process is applied to the calculated maps, a window of 500ms of previously calculated frames is used to calculate the weighted sum of the final saliency map. Using this procedure, the saliency maps maintained higher consistency in both tracking and representation of participants area in video frames. Figure 5.25 (B) shows the final saliency map for a single layer.

5.4.3 Layer Fusing

Fusing the layers, or mixing them, is considered the final step of this method to deliver the layers to the user view area. One of the main considerations in LP is the user should maintain clear visuals and auditory feedback from the location he is engaged at, while being aware of the other

locations simultaneously. Firstly, the layers are fused together based on the weight of each layer which is assigned based on user's eye gaze. Each layer's weight is calculated by sampling saliency map corresponding to each layer using eye gaze coordinates.
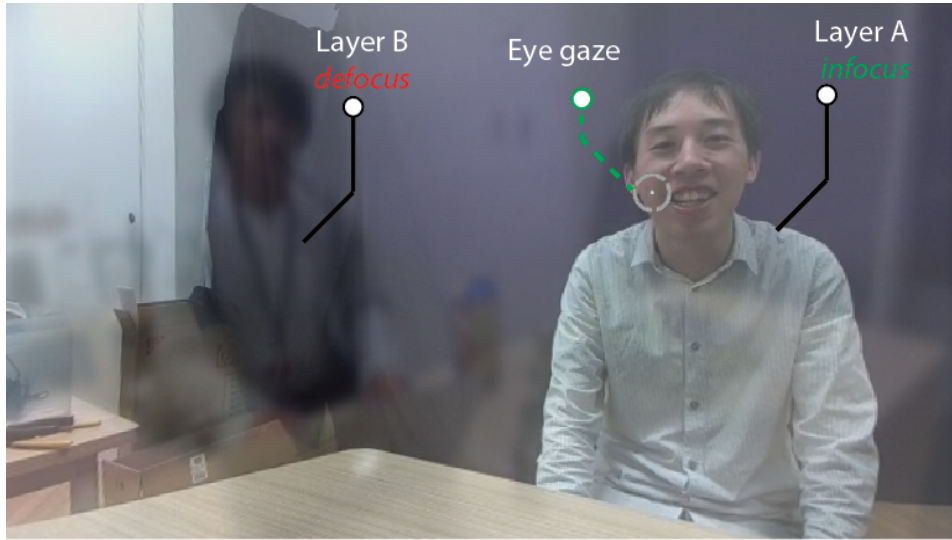
Figure 5.26: Two different locations fused together with an artificial depth of field.

In preliminary experiments of layer fusing, we used a basic alpha blending to all layers based on the calculated weight of each layer, however we found that its difficult for the users to clearly distinguish the visuals of the layers due to visual overlapping between all locations. To address this issue, we considered using a similar phenomenon seen in image reflection over window glass. Basically this phenomenon of reflection and transparency of window glass allows to see two different locations simultaneously as well as the ability to focus at different depths that would result blurriness of objects in the background of both locations (depth of field). Using this phenomenon, a pseudo model was defined that defines a focal value for each layer, that is basically driven from the calculated layer weight, is used to control the amount of shallowness of layers out of focus. Figure 5.26 shows the final results of fusing two layers using the proposed method.
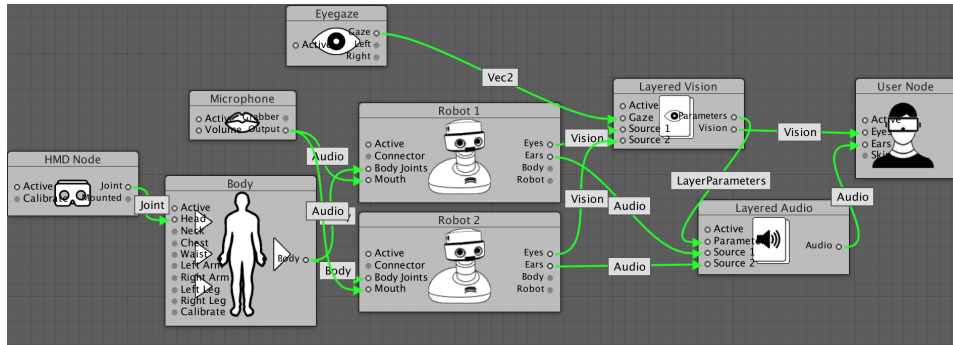
Figure 5.27: A screen-shot of layered presence meta-model using EDD editor.

5.4.4 Layered Presence Meta-modeling

The meta-model for layered presence is shown in Figure 5.27 . The model reflects the main components interacting with the user, the robot, and the intermediate transfer blocks. Two new transfer blocks encapsulates the functionality layering the visuals and auditory feedback:

- **Layered vision block**: receives several visual sources (can be determined by the designer) from the representations side. Calculates the saliency information based on the eye gaze of the user. The eye gaze modality block can be changed to a different modality type as long as it can provide a compatible output (Vec2). This block calculates the fused layers, and stores the results as vision data type, which can be linked to the user block.

- **Layered Audio block**: provides the functionality to layer the audio feedback from multiple sources based on the calculated layers weights from the [Layered Vision] block. The output of this block is audio type which can be directed to the user block.

5.4.5 Applications

This method can be expanded to more than just Telexistence related applications. By using the concept of layered perception, its possible to

139

generalize the layers to be media or even interactive applications, and apply the same procedure in combining them into a single space. Also, we define two modes of viewpoint (Egocentric & Exocentric), in which are based on user's point of view and how the user perceives the remote environment or the layers.

Egocentric Applications: The user in this type of application is immersed with the task he is doing, for example when the user is being engaged in a virtual reality game while wearing HMD. By using the concept of LP, the game it self is also considered as a layer that can be fused with other layers with eye gaze enabled. By using this method, the user is capable to be simultaneously engaged with the activity he is doing as well as with remote discussion. Figure 5.28 (A) shows an example of LP with virtual game. Also by using a video see-through HMD, its possible to consider the local location as a layer, and thus it can be seamlessly fused with a remote location as shown in Figure 5.28 (B).

Exocentric Applications: In this category of applications, the user perceives the layers from an external display with an external eye tracker mounted into the display. The contents of the display (application, visual media, etc.) can thus be considered as a layer of presence, while remote robots are blended with contents of the display. Figure 5.28 (C) shows an example of interacting with a software while being engaged in a simultaneous meeting in two different locations.

Figure 5.28: Layered presence applications. (A) interacting with remote location while being in VR environment, (B) use of video see-through while teleconferencing, and (C) layering media contents.

5.4.6 Public Demonstration & Discussion

After demonstrating Layered Presence system in SIGGRAPH 2016 [92], several observations and highlights were collected out of the attendees who experienced it. The demonstration consisted of three layers:

- Local layer A: provided a video see-through into the local space which the user is located at. The user can observe his body through it and his local surroundings.

- Remote layer B: located at a spatially different location from where the user is located at, in the demonstration it was referred to be in Japan (by setting a background image of Japanese site).

- Remote layer C: similarly to layer B, this layer was referred to be in Paris.

At each layer, a person was located which enabled direct communication with the user of the system. Figure 5.29 shows an overall layout of the experience setup used in which the participants experienced simultaneous multiple location perceptual presence. The low latency, consistent motion

Figure 5.29: Demonstration of Layered Presence at SIGGRAPH 2016 Anahiem. User connected to three different locations in parallel.

and optical flow that was maintained across the three layers provided seamless experience of spatial consistency. The corresponding representations (Layer B,C) were following the exact motion of the user head, which enabled mutual communication with remote participants. The usage of eye-gaze as an input to layer selection helped as an intuitive trigger to the region of interest. Most of the participants reported as being "co-located" with three persons, where in fact they were only facing one person. They even tried to reach by their hand to the remote participants during conversation to confirm whether they were actually located beside them or not. The fact of using such validation method shows the effectiveness of Layered Presence to convince the visual/kinesthetic modalities that the three layers coexist spatially together. When they failed to "touch" the remote participants, they reported as their body being a "phantom", not physically existed any more. This transitional state between physical and phantom body highlights the role of the body to validate the spatial localization, and can be viewed as an indicator to the effectiveness of perception layering. That is the user is aware his local space as well as the remote locations, and trying to interact in between them.

5.5 MetaLimbs

Figure 5.30: MetaLimbs: use of body topology reconfiguration to alter limbs mapping.

The last use-case discussed in this chapter is MetaLimbs Figure 5.30 . MetaLimbs addresses the ability to change body schema and postural topology in a way we can alter our limbs functions and model. We developed a proof of concept system that expands the number of limbs to four independently operated arms. Previous work discussed the topic of using an artificial prosthesis to augment or support body locally [74, 97] required a long period of time to adapt to the newly added arms in order to perform tasks, or was partially operated using an autonomous system which thus disembodies the newly added arms. In this work, we discuss the possibility of remapping our body to achieve immediate adaptation and operation for the new limbs, and to maintain the sense of embodiment towards these additional limbs.

As discussed in Section 3.4.3 , Topology reconfiguration meta-model is used to change the body schema topology by reassigning existing modalities to new ones in the representation. The goal is to maintain the sense of

143
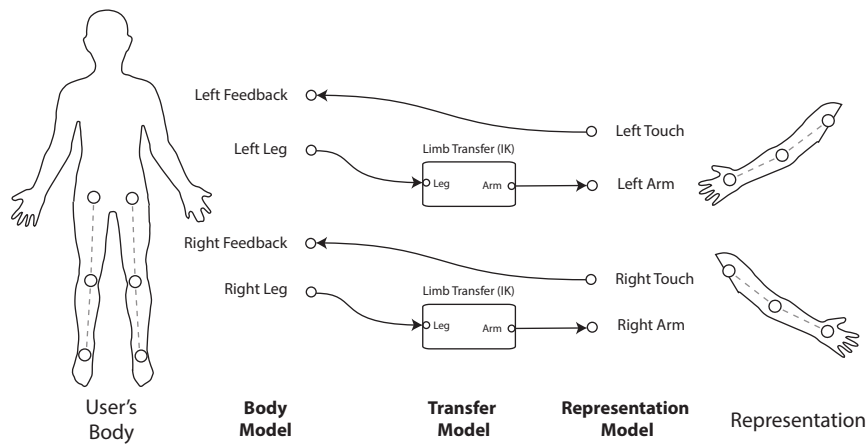
Figure 5.31: An overview of the target meta-model for remapping body limbs.

embodiment and ownership towards the additional limbs by mapping its functionality to existing joints of the body, thus when the joint moves, the artificial limb moves in accordance to its motion. For MetaLimbs, the topology reconfiguration meta-model maps legs motion to the newly added arms as viewed in Figure 5.31 . The arms also provide force feedback to the legs so the user can be aware of objects touching the hands (creating closed-loop feedback system). Further details about this system are described in the following sections.

5.5.1 Design Overview

MetaLimbs uses two 7 degrees of freedom anthropomorphic robotic arms that can be mounted on the user back. These arms can function as an extension of the human body, and function as extra limbs, thus expanding the number of arms into four arms. The main consideration in designing such a system is control mechanism of the new limbs in which it can provide independent motion of the four arms, and driven by the user to maintain the sense of ownership toward the new limbs. Existing work in Supernumerary Robotic Limbs (SRL) [73] uses a supportive approach for the additional

144
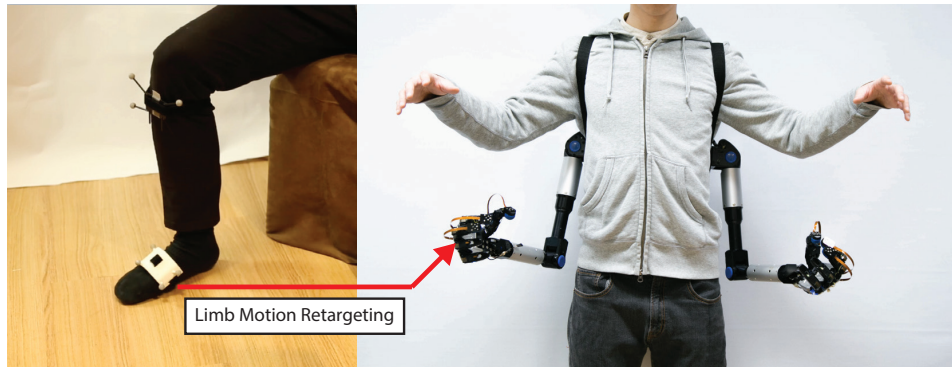
Figure 5.32: Re-configuring body schema mapping using feet motion.

limbs, which are mainly used as assistant limbs instead of being embodied with user's actions. For MetaLimbs, the approach used here is user-driven, that is to use existing body modalities to alter the operation of the new limbs. The approach used here is to use leg/feet motion to drive the arms, thus limb mapping model is used. Figure 5.32 shows an overview of body configuration in which feet motion are retargeted into the hand's motion.

5.5.2 Meta-model Design

In EDD modeling categories, Topology Reconfiguration meta-model addresses this type of interactions by redirecting an existing body modality output to a different part. For MetaLimbs, this can be achieved by redirecting legs motion into the artificial arms motion.

To realize this model, Figure 5.33 shows the meta-model for MetaLimbs system. Tracking markers were used to capture the spatial position and orientation of user's legs. One marker per feet, and single marker as a reference for the waist. **Arm IK** transfer block was used to reconstruct the arm posture based on the feet posture by calculating the inverse kinematics model of the arms. The resulting posture is provided to a body block which is directed into the robotic representation of the arms. The representation block handles the joints values (vector of angular data for each link) to drive the robotic arms. All calculations are done in real-time to match the intended actions of the user.
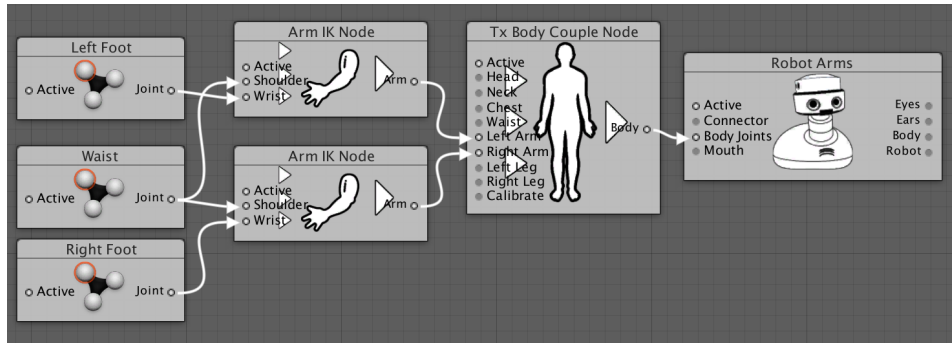
Figure 5.33: A screen-shot of MetaLimbs meta-model using EDD editor.

One of the advantages of using meta-modeling is the ability to apply non-linear mapping to the motion. For example, for high-precision control of the arms, the spatial resolution of the control can be scaled down so large motions are mapped to smaller motion trajectories. These types of control derivations can be easily done on a meta-level design.

5.5.3 Public Demonstration & Discussion

MetaLimbs system has been showcased in SIGGRAPH 2017 Emerging Technologies, in which more than 200 attendees experience the system for the first time. Generally speaking, the users were capable to adapt to the new body structure within the first 3 minutes of operation, in which they were able to hand shake, hold objects, and interact with their own arms Figure 5.34 . The main drawback in the operation was the mechanical latency of driving the arms. The participants spent the initial time to understand the speed limits and the spatial reachability of the arms (about 2 mins on average). Some participants were capable to develop fast adaptation to the new body schema and to interact using their own physical arms and the new limbs by doing coordinating actions. Usually, this type of motor-skill learning requires several trial and error until getting sufficient operation of the limbs.

This mechanism of body restructuring and schema adaptation would
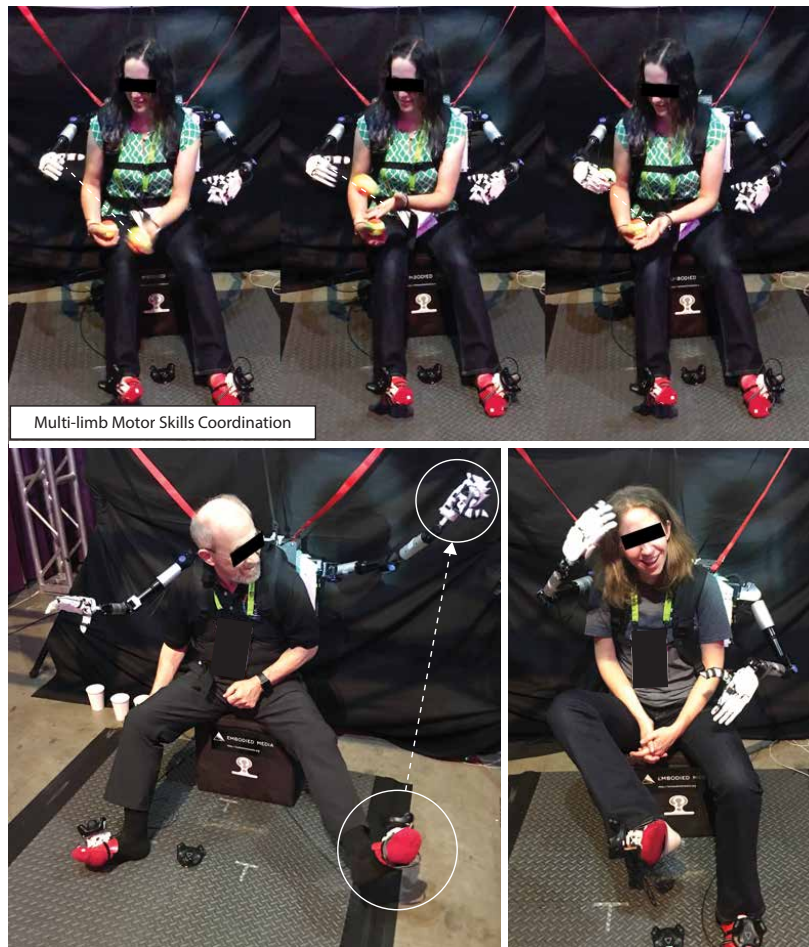
Figure 5.34: MetaLimbs at SIGGRAPH 2017 Los Angeles. Users adapted to the artificial arms are able to interact using them as being part of their body.

open the possibilities for new types of human augmentation functions to use on-body robotics extensions to facilitate a variety of tasks. Also, it can be used as an approach for prosthetic limbs control which can provide an immediate alteration and substitution of the missing limb.

One of the limitations this system lacks is the proprioceptive feedback regarding the posture of the arms. Currently, the arms are being driven one way by the legs move, and the user can understand their spatial position based on the legs posture (open loop system). However, when the arms are

adjusted by an external factor (like handshaking, or blocked by an object), the user will need to have visual feedback to maintain an understanding of the posture of the arms. Our biological arms, on the other hand, we are aware of their posture at any moment thanks to proprioceptive feedback, the inner structure and spatial position of each joint. To address this issue, we are considering to investigate in using EMS[1] feedback to create a closed-loop system, thus arms' posture would affect legs' posture.

5.6   Summary & Reflections

This chapter has demonstrated five use-cases and scenarios in which EDD meta-modeling approach helped addressing body limitations and physical disabilities and provided the tools to modulate the body schema and representation mapping. Also, it described how the use of such meta-level modeling approaches helps to alter the perception of space and physical limits while maintaining the embodied sense of ownership and agency during the operation. Concepts related to expanding the body physical limits and constraints using hybrid virtual and physical forms, exploring the topic of ubiquitous presence and expanded spatial perception, and reforming the topology of the human body structure by adding extra limbs and mapping them to existing modalities were discussed.

General observations of users who experienced the previously discussed projects have shared in common the active involvement through the newly provided embodied abilities and functions. In some projects, this effect was more dominant than others (Enforced Mutual Presence, Layered Presence, and MetaLimbs) in which users experienced new body function that was not presented before into their disposal (extended arms, expanded spatial awareness, and multiple arms respectively). In the scenarios which the body schema or representation has been altered from the original body form and interaction, users spent varied amount of time to learn and adapt to the

---

[1]Electrical muscle stimulation, the use of electrical impulses to cause muscle contraction.

new representation. The adaptation to the new body schema reflected the level of plasticity our bodies and cognitive abilities can achieve. When changing the body mapping, our understanding of the new body would also change as well as the way we embody our actions through it. For example, Enforced Mutual Presence showed that adding hands representation would significantly change our body awareness and level of presence. Similarly for MetaLimbs, when our body mapping was altered that the legs functions as arms, our cognitive understanding of body topology is changed accordingly. This alteration assisted to create new image of our new limbs affordabilities and what it can do, thus the users of the system quickly learned to use and maneuver through the new body schema.

One of the important means that EDD provides when addressing a body augmentation design problem is the flexibility to change the design and mapping of the body transfer model. Thus, the same system can be used by different users using different interactions and body mapping while, depending on the design, the embodied interaction is maintained. For example, for HUG Project scenario, it is possible to reuse the same system for a user with different physical impairment by changing body mapping and substituting the disability. Such flexibility is inherited from the meta-modeling approach used in the design of the system.

III

Conclusion & Future Direction

Chapter 6

Conclusion &
Future Direction

In this thesis, I proposed the idea of changing body schema mapping embody new functions through a mediated representations. This thesis outlined the concept of Embodied-Driven Design to create a range of embodied oriented applications through a series of use-cases and meta-models. In this chapter, a summary recapping the used methodologies of this research, outcome contributions, and current limitations. Finally, this thesis is concluded with an outlook to the future direction.

6.1   Summary

The topic of body schema modification and transfer have been studied before in philosophy to understanding human cognition of his/her body. Various cognitive driven theories have been proposed in this regard, and few of them has been put into evaluation and testing. Most of these studies were focused on testing specific phenomena of body perception. This thesis carries these theoretical studies and places them into a practical model that is derived from human body modalities. This work argues that human sense of his body schema can be altered and restructured by first overcoming the rigid coupling of his various modalities, and then use a meta-modeling approach to define the new model of the body schema. This approach aims to bridge between human body and digital representations to create transparent interaction through them.

This work approaches body schema altering by creating a topology of human body modalities, that act as input and output channels of the body. These modalities are represented as model blocks that offer channels of information to communicate with the human body and can be altered using a meta-model. EDD uses this meta-model to define the relationships between the various modalities used by the user or the target representation (physical or virtual) to define mapping and transfer functions between these blocks. EDD uses three types of modalities blocks to model the body schema: Perceptual, Transfer, and Representation blocks. The meta-model interfaces with both the user via various types of tracking and feedback blocks, and with the used representation using communication blocks. Three types of designing categories summarize the general use cases of the proposed meta-modeling: Representation alteration, Operation modulation, and Perception model expansion. Each category is concerned with a different type of body schema and perception mapping. Detailed design considerations of EDD has been thoroughly discussed in Chapter 3 .

Chapter 4 provided the implementation details for EDD framework. A meta-modeling editing environment has been developed and integrated into a commonly used tool, Unity3D, to facilitate the process of designing body schema models. The meta-modeling editor provides three blocks categories that can be directly used in the meta-model, or new meta-blocks can be implemented depending on the design requirements of the target application. These blocks provided in the editor have a wide support for various types of tracking and feedback technologies for the user side. To facilitate the design and prototyping process, a physical embodiment toolkit was developed. This toolkit provides the essential input/output modalities to create remote presence experiences and can be used directly within the meta-modeling environment. An optimization method for visual feedback "Foveated Streaming" integrated into the toolkit, which significantly enhances the overall performance, especially when multiple representations with visual feedback were required (such as Layered Presence case).

The proposed framework, meta-modeling approach, and toolkit were

152

used to resolve five different use-cases addressing body limitations and resolving physical disability. Chapter 5 showed the possibility to remodel body schema in a manner that allowed the users (depending on the use-case) to adapt to new body mapping that: extends their spatial interaction, alters their body representation, expands their perceptual feedback, or reconfigure their body postural topology. The first use-case in Section 5.1 addressed is a remotely operated surveillance vehicle which is targeted to be used in hazardous situations human cannot access. EDD provided a schematic matching between the operating user and the vehicle system, and assisted to extend spatial connectivity to the representation used. A direct body transfer meta-model approach was used to map operator's body with the corresponding representation, facilitating the embodied control of the robot without requiring adaptation time. The second use-case in Section 5.2 explored the possibility to use different mean to represent the body to overcome physical constraints and arm reach. Virtual arms were used to substitute physical representation, and to expand the spatial interactions of the arms by projecting them on the representation side. Representation alteration meta-modeling approach was used to facilitate this design. Third use-case discussed in Section 5.3 showed the effectiveness of using EDD to enable a physically impaired person to use the disabled function of her body. For this project, a topology reconfiguration meta-model was used to redirect a functioning modality (eye-gaze) to the disabled modality (neck), allowing the lady to use her eyes to control the head motion of the representation. Due to the correlated of eye and head, the lady was capable to adapt to the new mapping, and actively participate at a remote location. Fourth use-case discussed in Section 5.4 discussed the possibility to expand visual awareness from one location to many, leveraging sense of presence to ubiquitous level. Body schema and sensory topology are mapped to several representations simultaneously, which the corresponding feedback modalities are layered and provided to the user achieving parallel awareness of multiple locations. Modality expansion meta-model outlined the relationship between body and representations used. The last use-case explored

the possibility to reconfigure our body postural schema, and discussed in Section 5.5 . In this use-case, we demonstrated the plasticity human body has to adapt to new limb mapping, and to use it as an extension of the body instead of being a tool. MetaLimbs system was developed and used to expand the number of the arms. The meta-model redirects the feet motion to the arms trajectory, thus the user can maintain a sense of embodiment and agency towards the newly added arms.

Through these use-cases, we demonstrated the effectiveness of Embodied-Driven Design approach to alter the association between human and digital representations (or tools) from a meta-level perspective, to instead of using the tools, rather embodying their actions and feedback for effective and minimal adaptation.

6.2 Contribution

The design of body extension tools and technologies are necessary to be transparent and seamless with our actions in order to overcome a limitation or augment body functions. This thesis tackles this topic by providing an embodied oriented design approach with a modeling environment to allow body schema meta-modeling. Here I revise the main contributions and outcomes which this thesis provided:

- This thesis explored the theories of body and embodiment and investigated in the topic of the role of technology on our body and perception. A systematic model for body schema was established that translates the theories of body plasticity into practical tools. Embodied-Driven Design presented an abstraction medium for body modalities rewiring and design.

- Meta-modeling framework was designed and developed that realizes the concept of Embodied-Driven Design and has been integrated into a commonly used platform to enable a wide range of embodied oriented applications. This meta-modeling framework shifts the design

process of the conventional body augmentation systems to become a top level approach, that the design of the interaction is inherited from body capabilities and functions, and is possible to alter depending on the subjective use. Also, this framework can be effectively used prior the actual design of the system, so the designer could prototype and simulate the required mapping using virtual representation before deploying/designing a physical one.

- Four categories for body schema meta-modeling were proposed to facilitate meta-modeling design process: direct body transfer, representation alteration, topology reconfiguration, and modality expansion. These meta-models summarize the general scenarios which the EDD can modulate body schema to overcome a limitation or disability. These categories were evaluated using five scenarios.

- A physical embodiment toolkit was designed and developed that provides essential functions for remote presence applications. Toolkit functions were integrated into the EDD framework which can be accessed and mapped with user's body. This toolkit allows rapid prototyping for the embodiment designers to use and test of their ideas.

6.3 Limitations

Currently, the developed framework does not provide any constraints on the design process and the flow of information between the used modalities within the meta-model. This unconstrained design process is highly flexible during the modeling process, however, depending on the design used, the result interaction could generate unnatural and inconsistent feedback. For example, in HUG Project, when switching the eye gaze mapping axis in a way horizontal motion is mapped to robot's head tilt, and vertical motion to the pan axis, it would result in a highly unnatural feedback to the visual-motor system regardless of the correlated mapping. Although such design-error scenarios can be validated by experiencing the model and it-

erate until the inconsistency is resolved. Another approach is to deploy an embodiment grammar into the design process which can validate whether the model would result in such unnatural interaction. Thus reducing the space of possible meta-models down to a valid set of mapping operations that maintains the embodied interaction of the meta-model.

For the body operation modulation use-cases, such as discussed in MetaLimbs 5.5, the system lacks proprioceptive feedback regarding the posture of the artificial arms. Since the system as was shown uses an open-loop structure with no muscle feedback, the user relies on visual modality to understand the displaced posture of the arms when they are forced to a different posture externally (such as handshaking). We will further investigate using EMS feedback [52] in such operation modulation scenarios to produce proprioceptive cues to the body, creating a closed-loop system.

6.4 Outlook: Towards Radical Bodies

When writing this thesis, I established the topic of Embodied-Driven Design as a step of defining a systematic approach to leverage the limitations of human body through mediated tools based on the philosophy of human body and psychological studies conducted before in this area. Radical Bodies, the introduction topic of this thesis has presented the effect of the exponential progression of technology in reshaping our cognitive and physical abilities in tasks we were limited at. Some thinkers believes that this reliance on technology will not only be as a tool, but rather we would become the tools and lose our essence:

> "Every day millions of people decide to grant their smartphone a bit more control over their lives or try a new and more effective antidepressant drug. In pursuit of health, happiness and power, humans will gradually change first one of their features and then another, and another, until they will no longer be human."
> Yuval Harari, Homo Deus: A Brief History of Tomorrow p. 56

Despite the what can be seen as a pessimistic prediction of the progress of human augmentation, there are essential boundaries that can be used as a framework to shape this progress, such as using human-driven approaches to maintain the ownership and agency. Embodied Driven Design builds on these boundaries to define a framework for human augmentation tools design and altering the body schema. These presented tools are aimed to maintain the essence of the embodiment as the human is the core and the purpose of the design.

Radical Bodies in general aims to further investigate the role of technology and its impact on our life and body. How to maintain transparency between our bodies and the computational/robotic means to enhance our well-being and efficiency. Radical Bodies does not mean to change the structure of our bodies or mutating them, but rather to use the tools as complementary to our bodies in a non-invasive manner.

The following topics outline future development and investigations for Radical Bodies:

Embedded & Embodied Communication

One of the major problems in human bodies augmentation is realizing new communication channels for sensing and feedback that can be customized and integrated with the tools.

For the side of the coin, is the sensing part. Two paradigms of embodied tracking can be used: the explicit and internal state. The first approach (explicit) to use our voluntary controlled muscles[1] as an input mechanism in the design process, or to use their spatial end position (posture of the arm for example), which have been addressed in the design chapter. Other approaches investigated this problem by measuring the internal state of the body and mapping it to some actions, such as BCI (Brain Computer Interface) research. BCI monitors brain cortices activities for a specific set of features and matched with pre-trained limited catalog of actions (such as move left, right, trigger on, off, ...etc), which can be used as an activation

---

[1]healthy body has over 650 voluntary skeletal muscles.

mechanism to external tools and devices. However, our embodied actions are rather continuous and do not fall into a discrete set of commands. The use of hybrid methods (mind-body driven approaches) however, could lead to a wider expansion of the toolset of continuous control. I intend to further explore this area for embodied interaction and design.

The other side of the coin is the feedback part. As has been discussed in Body Operation Modulation: MetaLimbs Section 5.5 , when we modulate the body operation using an external device, we would need to maintain the sensory feedback to the body. In that case, it is the proprioceptive feedback. The mechanism and type of stimulation used to the body play an important role in identifying whether the medium belongs to our body, or as being an external object. It enhances the subjective sensation of ownership toward the medium. This thesis addressed the feedback mapping for vision, auditory, and tactile modalities, however further exploration in spatial feedback is required. EMS is one of the approaches that can support this spatial transfer and manipulation to the proprioceptive state of the human body.

Beyond Tools

The cycle of human and robotics has been carried along many iterations which were developing a better understanding of the role of technology in our life. The concept of Radical Bodies stretches beyond the use of technology as tools, but rather as a way to reshape our bodies and augment us in a new spectrum of tasks and activities. Recent topic, Superhuman Sports [49], has addressed the role of technology in changing our bodily capabilities in sports or augmenting our sensory feedback. The further the technology is embodied into us, the less we are aware of it, and more engaged we become in the activity at hand. As has been demonstrated in this thesis, Embodied-Driven Design supports such an argument by designing activities and tools driven by our bodily capabilities, and amplifying them using digitized representations.

IV

BIBLIOGRAPHY

"Research is to see what everybody else has seen,
and to think what nobody else has thought."
ALBERT SZENT-GYORGYI

References

[1] Tennessee science standards. http://glencoe.mcgraw-hill.com/.

[2] Adalgeirsson, S. O., and Breazeal, C. Mebot: a robotic platform for socially embodied presence. In *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*, IEEE Press (2010), 15–22.

[3] Asai, Y., Ueda, Y., Enomoto, R., Iwai, D., and Sato, K. Extendedhand on wheelchair. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, ACM (2016), 147–148.

[4] Bach-y Rita, P. *Brain mechanisms in sensory substitution.* academic Press, 1972.

[5] Bach-y Rita, P. *Nonsynaptic diffusion neurotransmission and late brain reorganization.* Demos Medical Pub, 1995.

[6] Bdeir, A., and Ullrich, T. Electronics as material: littlebits. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, ACM (2011), 341–344.

[7] Berti, A., and Frassinetti, F. When far becomes near: Remapping of space by tool use. *Journal of cognitive neuroscience 12*, 3 (2000), 415–420, 2000.

[8] Biocca, F. The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer-Mediated Communication 3*, 2 (1997), 0–0, 1997.

[9] Blanke, O., and Metzinger, T. Full-body illusions and minimal phenomenal selfhood. *Trends in cognitive sciences 13*, 1 (2009), 7–13, 2009.

[10] Bonilla, B. L., and Asada, H. H. A robot on the shoulder: Coordinated human-wearable robot control using coloured petri nets and partial least squares predictions. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE (2014), 119–125.

[11] Botvinick, M., and Cohen, J. Rubber hands' feel'touch that eyes see. *Nature 391*, 6669 (1998), 756, 1998.

[12] Bouckley, T., Clive-Smith, M., Eun Kim, M., and Sugawara, Y. Eidos, 2012. http://www.timbouckley.com/eidos.html.

[13] Bracken, C. C., and Skalski, P. *Immersed in media: Telepresence in everyday life*. Routledge, 2010.

[14] Capelle, C., Trullemans, C., Arno, P., and Veraart, C. A real-time experimental prototype for enhancement of vision rehabilitation using auditory substitution. *IEEE Transactions on Biomedical Engineering 45*, 10 (1998), 1279–1293, 1998.

[15] Carlson, T. A., Alvarez, G., Wu, D.-a., and Verstraten, F. A. Rapid assimilation of external objects into the body schema. *Psychological Science 21*, 7 (2010), 1000–1005, 2010.

[16] Cohen, M., Bolhassan, N. A., and Fernando, O. N. N. A multiuser multiperspective stereographic qtvr browser complemented by java3d visualizer and emulator. *Presence: Teleoperators and Virtual Environments 16*, 4 (2007), 414–438, 2007.

[17] Cohen, M., and Fernando, O. N. N. Awareware: Narrowcasting attributes for selective attention, privacy, and multipresence. In *Awareness Systems*. Springer, 2009, 259–289.

[18] Compendium, D. Annual disability statistics compendium 2016, 2016. https://disabilitycompendium.org/.

[19] Dodgson, N. A., et al. Variation and extrema of human interpupillary distance. In *Proceedings of SPIE*, vol. 5291 (2004), 36–46.

[20] Drake, R. L., Vogl, W., and Tibbitts, A. W. Mitchell; illustrations by richard; richardson, paul (2005). gray's anatomy for students.

[21] Ducklings, i. Duckings japan. https://ducklings.jp/ (last accessed November 2017).

[22] Fan, K., Huber, J., Nanayakkara, S., and Inami, M. Spidervision: extending the human field of view for augmented awareness. In *Proceedings of the 5th Augmented Human International Conference*, ACM (2014), 49.

[23] Fernando, C. L. *Design and development of a Telexistence system to experience extended human body schema*. PhD thesis, Keio University, 2013.

161

[24] Fernando, C. L., Furukawa, M., Kurogi, T., Kamuro, S., Sato, K., Minamizawa, K., and Tachi, S. Design of telesar v for transferring bodily consciousness in telexistence. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (oct. 2012), 5112 –5118.

[25] Follmer, S., Leithinger, D., Olwal, A., Hogge, A., and Ishii, H. inform: dynamic physical affordances and constraints through shape and object actuation. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM (2013), 417–426.

[26] Gallagher, S. *How the body shapes the mind*. Clarendon Press, 2006.

[27] Gartner. Gartner's hype cycle for emerging technologies, August 2016. http://www.gartner.com/.

[28] Gibson, J. J. The senses considered as perceptual systems. 1966.

[29] Guenter, B., Finch, M., Drucker, S., Tan, D., and Snyder, J. Foveated 3d graphics. *ACM Transactions on Graphics (TOG) 31*, 6 (2012), 164, 2012.

[30] Guterstam, A., Petkova, V. I., and Ehrsson, H. H. The illusion of owning a third arm. *PloS one 6*, 2 (2011), e17208, 2011.

[31] Haggard, P., and Wolpert, D. M. Disorders of body scheme. In *In Freund, HJ, Jeannerod, M., Hallett, M., Leiguarda R.,(Eds.), Higher-Order Motor Disorders*, Citeseer (2005).

[32] Hanneton, S., Auvray, M., and Durette, B. The vibe: a versatile vision-to-audition sensory substitution device. *Applied Bionics and Biomechanics 7*, 4 (2010), 269–276, 2010.

[33] Harbisson, N. I listen to color. *TEDXTalk, http://www. ted. com/talks/neil_harbisson_i_listen_to_color* (2012), 2012.

[34] Hasunuma, H., Kagaya, H., Takatori, M., Fujimori, J., Mifune, F., Shikoda, S., Kobayasi, M., Itoko, T., and Tachi, S. Development of teleoperation master system with a kinesthetic sensation of presence. In *Proc. of the International Conference on Artificial reality and Telexistence, Waseda University* (1999).

[35] Heinlein, R. A. *Waldo*. Doubleday, 1942.

[36] Holmes, N. P., and Spence, C. The body schema and multisensory representation (s) of peripersonal space. *Cognitive processing 5*, 2 (2004), 94–105, 2004.

[37] inc, E. C. Auxdeco. http://www.eyeplus2.com/product-2/about-auxdeco (Japanese, last accessed November 2017).

[38] Ishii, H., and Kobayashi, M. Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (1992), 525–532.

[39] Itti, L. Automatic foveation for video compression using a neurobiological model of visual attention. *IEEE Transactions on Image Processing 13*, 10 (2004), 1304–1318, 2004.

[40] Kac, E. Foundation and development of robotic art. *Art Journal 56*, 3 (1997), 60–67, 1997.

[41] Kasahara, S., Ando, M., Suganuma, K., and Rekimoto, J. Parallel eyes: Exploring human capability and behaviors with paralleled first person view sharing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM (2016), 1561–1572.

[42] Kasahara, S., and Rekimoto, J. Jackin: integrating first-person view with out-of-body vision generation for human-human augmentation. In *Proceedings of the 5th Augmented Human International Conference*, ACM (2014), 46.

[43] Kawakami, N. Telesarphone-communication robot based on next generation telexistence technologies. In *Proceedings of 36th International Symposium on Robotics (ISR2005)* (2005).

[44] Kay, L. Sonic glasses for the blind: A progress report. *American Foundation for the Blind Research Bulletin 25* (1973), 25–58, 1973.

[45] Kazemitabaar, M., McPeak, J., Jiao, A., He, L., Outing, T., and Froehlich, J. E. Makerwear: A tangible approach to interactive wearable creation for children. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM (2017), 133–145.

[46] Kim, T. Effects of presence on memory and persuasion. *University of North Carolina, Chapel Hill, NC* (1996), 1996.

[47] Koizumi, N., Sugimoto, M., Nagaya, N., Inami, M., and Furukawa, M. Stop motion goggle: augmented visual perception by subtraction method using high speed liquid crystal. In *Proceedings of the 3rd Augmented Human International Conference*, ACM (2012), 14.

[48] Kunz, A., Nescher, T., and Kuchler, M. Collaboard: a novel interactive electronic whiteboard for remote collaboration with people on content. In *Cyberworlds (CW), 2010 International Conference on*, IEEE (2010), 430–437.

[49] Kunze, K., Minamizawa, K., Lukosch, S., Inami, M., and Rekimoto, J. Superhuman sports: Applying human augmentation to physical exercise. *IEEE Pervasive Computing 16*, 2 (2017), 14–17, 2017.

[50] Leithinger, D., Follmer, S., Olwal, A., and Ishii, H. Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM (2014), 461–470.

[51] Llorens-Bonilla, B., Parietti, F., and Asada, H. H. Demonstration-based control of supernumerary robotic limbs. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, IEEE (2012), 3936–3942.

[52] Lopes, P., Ion, A., Mueller, W., Hoffmann, D., Jonell, P., and Baudisch, P. Proprioceptive interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM (2015), 939–948.

[53] Lugrin, J.-L., Latt, J., and Latoschik, M. E. Anthropomorphism and illusion of virtual body ownership. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, Eurographics Association (2015), 1–8.

[54] Macaluso, E., and Maravita, A. The representation of space near the body through touch and vision. *Neuropsychologia 48*, 3 (2010), 782–795, 2010.

[55] MacLachlan, M. *Embodiment: Clinical, Critical and Cultural Perspectives on Health and Illness*. McGraw-Hill Education (UK), 2004.

[56] Maidenbaum, S., Hanassy, S., Abboud, S., Buchs, G., Chebat, D.-R., Levy-Tzedek, S., and Amedi, A. The eyecane, a new electronic travel aid for the blind: Technology, behavior & swift learning. *Restorative neurology and neuroscience 32*, 6 (2014), 813–824, 2014.

[57] Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. Changing bodies changes minds: owning another body affects social cognition. *Trends in cognitive sciences 19*, 1 (2015), 6–12, 2015.

[58] Malafouris, L. *How things shape the mind.* MIT Press, 2013.

[59] Maloney, J., Resnick, M., Rusk, N., Silverman, B., and Eastmond, E. The scratch programming language and environment. *ACM Transactions on Computing Education (TOCE) 10*, 4 (2010), 16, 2010.

[60] Maravita, A., and Iriki, A. Tools for the body (schema). *Trends in cognitive sciences 8*, 2 (2004), 79–86, 2004.

[61] Maravita, A., Spence, C., and Driver, J. Multisensory integration and the body schema: close to hand and within reach. *Current biology 13*, 13 (2003), R531–R539, 2003.

[62] Marcel, A. J., Eilan, N., et al. *The body and the self.* Mit Press, 1995.

[63] Meijer, P. B. An experimental system for auditory image representations. *IEEE transactions on biomedical engineering 39*, 2 (1992), 112–121, 1992.

[64] Melzack, R. Phantom limbs. *Scientific American 266*, 4 (1992), 120–127, 1992.

[65] MetaModel, i. The voice. http://www.metamodal.com/ (last accessed November 2017).

[66] Metzinger, T. Empirical perspectives from the self-model theory of subjectivity: a brief summary with examples. *Progress in brain research 168* (2007), 215–278, 2007.

[67] Minamizawa, K., Fukamachi, S., Kajimoto, H., Kawakami, N., and Tachi, S. Gravity grabber: wearable haptic display to present virtual mass sensation. In *ACM SIGGRAPH 2007 emerging technologies*, ACM (2007), 8.

[68] Minsky, M. Telepresence. 1980.

[69] Misawa, K., Ishiguro, Y., and Rekimoto, J. Livemask: A telepresence surrogate system with a face-shaped screen for supporting nonverbal communication. *Information and Media Technologies 8*, 2 (2013), 617–625, 2013.

[70] Morasso, P., Casadio, M., Mohan, V., Rea, F., and Zenzeri, J. Revisiting the body-schema concept in the context of whole-body postural-focal dynamics. *Frontiers in human neuroscience 9* (2015), 2015.

[71] Ogawa, N., Ban, Y., Sakurai, S., Narumi, T., Tanikawa, T., and Hirose, M. Metamorphosis hand: dynamically transforming hands. In *Proceedings of the 7th Augmented Human International Conference 2016*, ACM (2016), 51.

[72] Ogawa, S., Okahara, K., Iwai, D., and Sato, K. A reachable user interface by the graphically extended hand. In *Consumer Electronics (GCCE), 2012 IEEE 1st Global Conference on*, IEEE (2012), 210–211.

[73] Parietti, F., and Asada, H. Supernumerary robotic limbs for human body support. *IEEE Transactions on Robotics 32*, 2 (2016), 301–311, 2016.

[74] Parietti, F., Chan, K., and Asada, H. H. Bracing the human body with supernumerary robotic limbs for physical assistance and load reduction. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, IEEE (2014), 141–148.

[75] Patney, A., Kim, J., Salvi, M., Kaplanyan, A., Wyman, C., Benty, N., Lefohn, A., and Luebke, D. Perceptually-based foveated virtual reality. In *ACM SIGGRAPH 2016 Emerging Technologies*, ACM (2016), 17.

[76] Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition 22*, 3 (2013), 779–787, 2013.

[77] Pfeifer, R., and Bongard, J. *How the body shapes the way we think: a new view of intelligence.* MIT press, 2006.

[78] Puckette, M., and Zicarelli, D. Max/msp. *Cycling 74* (1990), 1990–2006, 1990.

[79] Puckette, M. S., et al. Pure data. In *ICMC* (1997).

[80] Ramachandran, V. S., Rogers-Ramachandran, D., and Cobb, S. Touching the phantom limb. *Nature 377*, 6549 (1995), 489, 1995.

[81] Rheingold, H. *Virtual reality: exploring the brave new technologies.* Simon & Schuster Adult Publishing Group, 1991.

[82] Riva, G. *From Communication to Presence: Cognition, Emotions, and Culture Towards the Ultimate Communicative Experience: Festschrift in Honor of Luigi Anolli*, vol. 9. IOS Press, 2006.

[83] Robotics, D. Doubletelepresence robot, 2015.

[84] Robotics, R. Kubi telepresence: First practical robotic platform for tablet telepresence, 2013. https://www.revolverobotics.com/.

[85] Rodman, J. Polycom (r). *White PaperThe Distributed Dawn: the Emergence of Versatile Performance in the Conferencing Environment* (2004), 1–7, 2004.

[86] Roessler, J., and Eilan, N. *Agency and self-awareness: issues in philosophy and psychology*, vol. 2. Oxford University Press, 2003.

[87] Rosenberg, R. S., Baughman, S. L., and Bailenson, J. N. Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. *PloS one 8*, 1 (2013), e55003, 2013.

[88] Ryoo, J., Yun, K., Samaras, D., Das, S. R., and Zelinsky, G. Design and evaluation of a foveated video streaming service for commodity client devices. In *Proceedings of the 7th International Conference on Multimedia Systems*, ACM (2016), 6.

[89] Saraiji, M. Y., Fernando, C. L., Minamizawa, K., and Tachi, S. Development of mutual telexistence system using virtual projection of operator's egocentric body images. In *ICAT-EGVE* (2015), 125–132.

[90] Saraiji, M. Y., Fernando, C. L., Minamizawa, K., and Tachi, S. Mutual hand representation for telexistence robots using projected virtual hands. In *Proceedings of the 6th Augmented Human International Conference*, ACM (2015), 221–222.

[91] Saraiji, M. Y., Fernando, C. L., Mizushina, Y., Kamiyama, Y., Minamizawa, K., and Tachi, S. Enforced telexistence: teleoperating using photorealistic virtual body and haptic feedback. In *SIGGRAPH Asia 2014 Emerging Technologies*, ACM (2014), 7.

[92] Saraiji, M. Y., Sugimoto, S., Fernando, C. L., Minamizawa, K., and Tachi, S. Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending. In *ACM SIGGRAPH 2016 Emerging Technologies*, ACM (2016), 20.

[93] Shapiro, L. *Embodied cognition*. Routledge, 2010.

[94] Sheridan, T. B. Musings on telepresence and virtual presence. *Presence: Teleoperators & Virtual Environments 1*, 1 (1992), 120–126, 1992.

[95] Shu, N. C. *Visual programming*. Van Nostrand Reinhold New York, 1988.

[96] Slater, M., and Usoh, M. Representations systems, perceptual position, and presence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments 2*, 3 (1993), 221–233, 1993.

[97] Stelarc. Third hand, 2015. http://stelarc.org/?catID=20265 (Last accessed November 2017).

[98] Suzuki, K., Wakisaka, S., and Fujii, N. Substitutional reality system: a novel experimental platform for experiencing alternative reality. *Scientific reports 2* (2012), srep00459, 2012.

[99] Tachi, S. Tele-existence-toward virtual existence in real and/or virtual worlds. In *Proceedings of the International Conference on Artificial Reality and Tele-existence (ICAT'91)* (1991), 85–94.

[100] Tachi, S., Arai, H., and Maeda, T. Tele-existence master-slave system for remote manipulation. ii. In *Decision and Control, 1990., Proceedings of the 29th IEEE Conference on*, IEEE (1990), 85–90.

[101] Tachi, S., Tanie, K., Komoriya, K., and Kaneko, M. Tele-existence (i): Design and evaluation of a visual display with sensation of presence. In *Theory and Practice of Robots and Manipulators*. Springer, 1985, 245–254.

[102] Tsui, K. M., Desai, M., Yanco, H. A., and Uhlik, C. Exploring use cases for telepresence robots. In *Proceedings of the 6th international conference on Human-robot interaction*, ACM (2011), 11–18.

[103] Wang, Z., and Bovik, A. C. Embedded foveation image coding. *IEEE Transactions on image processing 10*, 10 (2001), 1397–1410, 2001.

[104] Warwick, K. Cyborg 1.0. *Wired 8*, 2 (2000), 2000.

[105] Watanabe, K., Kawabuchi, I., Kawakami, N., Maeda, T., and Tachi, S. Torso: completion of egocentric telegnosis system. In *ACM SIGGRAPH 2007 emerging technologies*, ACM (2007), 21.

[106] Watanabe, K., Kawabuchi, I., Kawakami, N., Maeda, T., and Tachi, S. Torso: Development of a telexistence visual system using a 6-dof robot head. *Advanced Robotics 22*, 10 (2008), 1053–1073, 2008.

[107] Weiser, M. Ubiquitous computing. In *ACM Conference on Computer Science* (1994), 418.

[108] White, B. W., Saunders, F. A., Scadden, L., Bach-Y-Rita, P., and Collins, C. C. Seeing with the skin. *Perception & Psychophysics 7*, 1 (1970), 23–27, 1970.

[109] Wicab, i. Brainport. https://www.wicab.com/brainport-v100 (last accessed November 2017).

[110] Wu, F. Y., and Asada, H. Bio-artificial synergies for grasp posture control of supernumerary robotic fingers. 2014.

V

APPENDICES

Appendix A

EDD Related Publications

A.1 Conference Papers

A.1.1 International Conferences

1. **MHD Yamen Saraiji**, Charith Fernando, Masahiro Furukawa, Kouta Minarnizawa, and Susumu Tachi. *"Real-time egocentric superimposition of operator's own body on telexistence avatar in virtual environment."* In Artificial Reality and Telexistence (ICAT), 2013 23rd International Conference on, pp. 35-39. IEEE, 2013.

2. **MHD Yamen Saraiji**, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Development of Mutual Telexistence System using Virtual Projection of Operator's Egocentric Body Images."* In ICAT-EGVE, pp. 125-132. 2015.

3. **MHD Yamen Saraiji**, Shota Sugimoto, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending."* In ACM SIGGRAPH 2016 Posters, p. 20. ACM, 2016.

4. **MHD Yamen Saraiji**, Shota Sugimoto, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending."* In ACM SIGGRAPH 2016 Posters, p. 20. ACM, 2016.

5. **MHD Yamen Saraiji**, Yusuke Mizushina, Charith Fernando, Masahiro Furukawa, Youichi Kamiyama, Kouta Minamizawa, and Susumu Tachi. *"Enforced telexistence."* In ACM SIGGRAPH 2014 Posters, p. 49. ACM, 2014.

A.1.2 Domestic Conferences

1. **MHD Yamen Saraiji**, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Study on Telexistence LXXXV Layered Presence: Expanding Visual Presence using Simultaneously Operated Telexistence Avatars."* In Virtual Reality Society of Japan (2016).

2. **MHD Yamen Saraiji**, Kouta Minamizawa, and Susumu Tachi. *"Study on Telexistence LXXXVIII Foveated Streaming: Optimizing video streaming for Telexistence systems using eye-gaze based foveation."* In Virtual Reality Society of Japan (2017).

A.2 Demonstrations

1. **MHD Yamen Saraiji**, Charith Fernando, Yusuke Mizushina, Youichi Kamiyama, Kouta Minamizawa, and Susumu Tachi. *"Enforced telexistence: teleoperating using photorealistic virtual body and haptic feedback."* In SIGGRAPH Asia 2014 Emerging Technologies, p. 7. ACM, 2014.

2. **MHD Yamen Saraiji**, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Mutual hand representation for telexistence robots using projected virtual hands."* In Proceedings of the 6th Augmented Human International Conference, pp. 221-222. ACM, 2015.

3. **MHD Yamen Saraiji**, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Development of Mutual Telexistence System using Virtual Projection of Operator's Egocentric Body Images."* In ICAT-EGVE, pp. 125-132. 2015.

4. **MHD Yamen Saraiji**, Shota Sugimoto, Charith Fernando, Kouta Minamizawa, and Susumu Tachi. *"Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending."* In ACM SIGGRAPH 2016 Emerging Technologies, ACM, 2016.

5. Tomoya Sasaki, **MHD Yamen Saraiji**, Charith Fernando, Kouta Minamizawa, and Masahiko Inami. *"MetaLimbs: Metamorphosis for multiple arms interaction using artificial limbs."* In 44th International Conference on Computer Graphics and Interactive Techniques, ACM SIGGRAPH 2017. Association for Computing Machinery, Inc, 2017.

A.3   Awards

1. Best Demo Award, *"Mutual hand representation for telexistence robots using projected virtual hands."* In Proceedings of the 6th Augmented Human International Conference, pp. 221-222. ACM, 2015.

2. Honorable Mention & Best Demo Award, *"Development of Mutual Telexistence System using Virtual Projection of Operator's Egocentric Body Images."* In ICAT-EGVE, pp. 125-132. 2015.

3. Best in Show & Best Care and Welfare Project award, *"HUG Project."* In Pepper App Challenge 2015.

4. Best Research Presentation award, *"Study on Telexistence LXXXV Layered Presence: Expanding Visual Presence using Simultaneously Operated Telexistence Avatars."* In Virtual Reality Society of Japan (2016).

5. Student Research Competition Gold Prize, *"Layered telepresence: simultaneous multi presence experience using eye gaze based perceptual awareness blending."* In SIGGRAPH 2017

6. Best in Show award, *"MetaLimbs: Metamorphosis for multiple arms interaction using artificial limbs."* In SIGGRAPH 2017