The OctaVis:

A VR-Device for Rehabilitation & Diagnostics of Visuospatial Impairments

Dissertation

zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr. rer. nat.)
an der Technischen Fakultät der Universität Bielefeld

vorgelegt von
Eugen Dyck

Bielefeld 2016

Versicherung

Hiermit versichere ich,

- dass mir die geltende Promotionsordnung der Fakultät bekannt ist,
- dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte von Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle benutzten Hilfsmittel und Quellen in meiner Arbeit angegeben habe,
- dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Vermittlungstätigkeiten oder für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,
- dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe und
- dass ich keine gleiche, oder in wesentlichen Teilen ähnliche oder eine andere Abhandlung bei einer anderen Hochschule als Dissertation eingereicht habe.

Bielefeld, 2016

Printed on non-ageing paper – ISO 9706

Abstract

In an age of unprecedented technological impact on our cognition, this thesis presents interdisciplinary research on a virtual reality device meant to enhance traditional therapy of impairments in spatial thinking resulting from an accident, a stroke or some neuro-degenerative disease.

Analyzing common strategies and goals in neuro-psychological therapy, a stark resemblance with general tactics from game- and virtual environment design is disclosed (empowerment, interaction, personal meaning, flow, etc). Additionally, noticing the inability of traditional therapy to transfer training success to real-life situations, a grocery-shopping task tackling executive functions, spatial memory, orientation, and recognition is introduced as a promising rehabilitation scenario for visuospatial recovery.

Following the research results on presence and its effects on task-performance and skill-transfer, the development of a novel VR device – the OctaVis– is presented. It consists of eight touch screens surrounding the patient with a 360° panorama view. The patient is seated on an office chair equipped with a joystick in the armrest to move forward/backward and a rotary encoder in the shaft of the chair to determine the virtual moving direction. Being CE-certified, the device meets both, patient and hospital requirements (space & cost efficiency, low maintenance, easy to use by nontechnical staff). To render the virtual environment, while meeting project demands, a custom render framework for a single PC hosting three GPUs to drive eight monitors has been developed and optimized to render the supermarket (4M triangles) at more than 70 fps.

To evaluate the device, first, the extension of the horizontal field of view from 135° (three displays) to 360° (eight displays) is examined regarding its effect on spatial presence and on the accuracy in a pointing task. Second, driving the full eight screens, the effect of embodied self-rotation using the same measures is explored. In particular, navigation by rotating the world while the user is sitting stable is compared to a stable world and a self-rotating user. Finally, several clinical trials with patients confirm the OctaVis to provide stable learning, to generalize training success into real reality, and even to outperform standard diagnostic tests.

List of Used Acronyms

VR Virtual Reality

RR Real Reality

GPU Graphics Processing Unit

PTSD Post-Traumatic Stress Disorder

VT Virtual Therapy

SSM Spatial Situation Model

PERF Primary Egocentric Reference Frame

HMD Head Mounted Display

CVLT California Verbal Learning Test

VLMT Verbal Learning and Memory Test

VA Vertex Arrays

VBO Vertex Buffer Objects

UBO Uniform Buffer Object

MEC SPQ Measurement Effects Conditions Spatial Presence Questionnaire

FV Frontal View (Test Condition)

SV Surround View (Test Condition)

RU Rotating User (Test Condition)

MRTA Mental Rotation Test, A-Version

NET Paper-Pencil Neglect Test

CONTENTS

CHAPTER 1

Introduction

I fear the day when technology overlaps our humanity.
It will be then that the world will have
permanent ensuing generations of idiots.

— Albert Einstein

People have been concerned about the reach of technology into human life throughout history. While some cultures, up to date, despise engines and tires because they destroy the simple and pure life, others draw the line at organ transplantation or genetic engineering. Technology nowadays, however, not only assists humans in their everyday activity or with life itself. It indeed strives to surpass the very biology of mankind, opening a world beyond our can-do: sensors reveal inner tissue of living patients where no human eye can see, driver assistance systems anticipate dangerous traffic situations long before our conscience becomes aware of them, and the precision of robots in assembly lines overcomes human performance by magnitudes, to name a few. Though ethical review committees and politics regularly discuss the limits of autonomy and the future course of technology, which they most certainly should hold on to, there seems to be an inevitable ongoing progress in action. Step by step, technology advances towards devices or human enhancements to understand and act in a reality too far away, too subtle to be recognized, or too sensitive to be touched by humans.

Somewhat reversed, current technology also is advanced enough to trick the human mind into a reality that is not a reality at all, at least not in the common sense. Over the last decades we witnessed a steep development in the field of *virtual reality* (VR). Virtual environments in computer games, for example, create an illusion so strong that users actually feel forces or emotions in the real world though acting in a virtual one. In that view, it is astonishing to what degree some users immerse into a world made of pixels. While some concerns about such a technology, including the quote above, may be legit, this field of research offers great opportunities to deal with matters of the mind for good. VR, with big steps, walks the road from being a mere fun gaming experience to sincere virtual therapy. It does so combining the knowledge generated by the gaming industry, the results of numerous VR research groups around the world, and the experience of psychology departments and hospitals. This interdisciplinary approach, in concert, provides at least two compelling advancements for therapy:

1. *Patients* are allowed to experience an arbitrary artificial reality as real and therefore to transfer some of the virtually learned skills back into *real reality* (RR). This even includes experiences like walking, navigating, wayfinding, shopping or exposure to fearsome situations without actually leaving a room.

2. *Doctors* are allowed to observe patients in situations which otherwise would have stayed inaccessible to them. Additionally, these virtual situations are fully designable and controllable.

1.1 Motivation

Every year about 270.000 people in Germany suffer from a stroke. The immediate consequences range from loss of speech, affection of cognitive abilities like visuospatial thinking, attention, memory, executive brain functions to partial or full hemiparesis. Half of these patients never fully recover and remain disabled. Similar brain injuries are caused by accidents or other neuro-degenerative diseases (e.g. epilepsy). Taking into account that for example visuospatial cognition is needed for orientation, navigation from work to home, to understand a sign, or even to recognize someone beside you, it becomes clear how deep such injuries reach into everyday life and eventually disrupt it completely and forever.

Currently, training after such injuries involves mainly card games and paper-and-pencil tests, which target very isolated brain functions. While patients improve on these abstract training tasks, they often cannot sufficiently transfer their cognitive improvement to the broader and more complex everyday live challenges. Unfortunately, training that targets actual everyday routines in their corresponding physical settings (street crossing, shopping in a real supermarket, etc.) is infeasible due to cost and staff shortage in hospitals. Allowing for training in exactly such scenarios, this is where VR becomes a strong candidate for the treatment of patients with the above mentioned cognitive impairments. The downside of many current VR devices, however, is their unsuitability for an installation, use, and acceptance in a hospital as a rehabilitation device.

Funded through the NRW Ziel 2 program EFRE (Europäischer Fonds für regionale Entwicklung), this thesis is a partial result of the *Cognitive Interaction Technology for Medical Applications* (CITmed) project at Bielefeld's Center of Excellence CITEC. The major goal of the highly interdisciplinary team of computer scientists, psychologists, and medical partners was the development and employment of a novel VR training device overcoming the downside of current devices and enhancing training success for patients.

This thesis is devoted to all patients suffering from the mentioned cognitive constraints without the possibility to train in a safe and secure environment.

1.2 Contribution

Based on the stated motivation, the main contributions of this thesis are:

A novel VR device for clinical studies The OctaVis system is a virtual reality platform developed for diagnosis, rehabilitation, and training of patients with brain function disorders. To meet the special requirements of clinical studies, the system has been designed with ease of use, patient safety, ease of maintenance, minimal space occupation and cost efficiency in mind. Patients are sitting on a rotating office chair in the center of eight touch screen displays which are arranged in an octagon around them, thereby, providing a 360° horizontal panorama view. Navigation is intuitively controlled through chair rotation and a joystick in the armrest. A touch interface enables easy object selection as known from the real world.

A multi-GPU, multi-view rendering system Also, a high-performance system for multi-view rendering in the OctaVis setup is presented. In contrast to complex CAVE installations, which are typically driven by one render client per view, the OctaVis drives its eight displays by a single PC equipped with multiple graphics units (GPUs). An appropriate software setup, as well as the necessary low-level and high-level optimizations, to optimally exploit the parallelism of this multi-GPU multi-view VR system are introduced.

A study investigating spatial orientation Since the OctaVis aims at patients with disturbances in visuospatial thinking, a study was conducted to investigate potentially beneficial effects of a horizontal surround view and the user rotation in the real world on a spatial orientation task. The presented results qualify the OctaVis as a suitable VR device for virtual therapy concerned with spatial thinking.

An overview of clinical studies with the OctaVis Several co-authored studies with real patients are lined up to show that the OctaVis as a rehabilitation device provides not only *stable* visuospatial learning, but also *generalizes* that learning from the specific training task to other spatial challenges, and actually correlates with the corresponding *real-life* task.

1.3 Thesis Organisation

Virtuality The following chapter explains the fundamentals of VR and its use in therapy in general and in particular for neuro-degenerative diseases. The driving question here is: how, to what level, and with which results can our mind be convinced to assume virtual things and skills to be real? This is the groundwork and understanding why VR has such a big impact on its users.

The OctaVis Before providing an overview of the design and technology decisions that led to the development of the OctaVis system and an overview of the hardware setup itself, this chapter analyses current VR devices, their advantages/disadvantages, and their candidacy as a rehabilitation device in a hospital environment.

Software Setup Since there was no render-engine capable of driving the OctaVis according to its demands, this chapter describes the software that was developed to meet the derived requirements. It contains the overall software design including a multi-GPU render-engine and a modular framework to extend the package with custom experiment features.

Evaluation of Spatial Orientation After the description of the newly designed system in the chapters before, this chapter evaluates the capability of the OctaVis to enhance spatial learning in comparison to other common navigation metaphors and presentation models. This is a crucial evaluation since the device especially targets patients with visuospatial impairments.

Clinical Evaluation Finally, underlining the clinical relevance of the OctaVis for rehabilitation and diagnostics, several studies conducted in different hospitals are presented.

1.4 Publications

During the work summarized in this thesis, the following papers have been published.

E. Dyck, H. Schmidt, and M. Botsch, "OctaVis: A simple and efficient multi-view rendering system," in *Proceedings of GI VR/AR Workshop*, pp. 1–8, 2010.

E. Dyck, H. Schmidt, M. Piefke, and M. Botsch, "OctaVis: Optimization techniques for multi-GPU multi-view rendering," *Journal of Virtual Reality and Broadcasting*, vol. 9, no. 6, 2012.

E. Dyck, E. Zell, A. Kohsik, P. Grewe, Y. Winter, M. Piefke, and M. Botsch, "OctaVis: An easy-to-use VR-system for clinical studies," in *Proceedings of Virtual Reality Interaction and Physical Simulation (VRIPHYS)*, pp. 127–136, 2012.

E. Dyck, T. Pfeiffer, and M. Botsch, "Evaluation of surround-view and self-rotation in the OctaVis VR-system," in *Proceedings of the 5th Joint Virtual Reality Conference*, pp. 1–8, 2013.

P. Grewe, A. Kohsik, D. Flentge, E. Dyck, C. Bien, Y. Winter, M. Botsch, H. J. Markowitsch, and M. Piefke, "Learning real-life cognitive abilities in a novel 360°-virtual reality supermarket: A neuropsychological study of healthy participants and patients with epilepsy," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, p. 42, 2013.

E. Zell, E. Dyck, A. Kohsik, P. Grewe, D. Flentge, Y. Winter, M. Piefke, and M. Botsch, "OctaVis: A virtual reality system for clinical studies and rehabilitation," in *Eurographics 2013 – Dirk Bartz Prize*, pp. 9–12, 2013.

P. Grewe, D. Lahr, A. Kohsik, E. Dyck, H. J. Markowitsch, C. G. Bien, M. Botsch, and M. Piefke, "Real-life memory and spatial navigation in patients with focal epilepsy: Ecological validity of a virtual supermarket task," *Epilepsy & Bahavior*, vol. 31, pp. 57–66, 2014.

CHAPTER 2

Virtuality

Consciousness is an illusion that our internal experiences are reality, when in truth they are imperfect simulations of something we may never truly understand

— Jesse Shell

Jesse Shell is a game designer teaching at Carnegie Mellon University. Formerly he worked at Disney creating many different games, and among many other things before that, he was a magician and a professional juggler. The quote above has derived from both, personal experience based on a history of guiding the audience's attention as an entertainer, and his expertise as a game designer and professor. In his book *The Art of Game Design* [She08], which is considered as standard literature for future game designers, he emphasizes demographics, psychographics, player types, imagination, empathy and other psychological aspects rather than the technology and mechanics of a game. In fact, the game industry in general not only drives the development of rendering hardware and software, but also is a strong force generating insight into the human psyche in its relation to virtual worlds and virtual actions.

Today, before understanding the principles of gameplay, a specific technology to implement it, the user interface or artificial intelligence, a game designer is expected to know about:

1. Mental model creation

2. How to manipulate such an existing model in the player's mind

3. How to bind emotion to a mental model

4. The different categories of fun

5. How to harness imagination to spark curiosity

6. How to generate flow

This is by no means a complete list, but it already provides the importance the game industry applies to the understanding of the inner workings of the human mind to build exciting virtual worlds and engaging virtual missions. At the far end of this venture, excitement and engagement sometimes even turn to strong addiction, not just to gaming, but to the virtual world itself. Michael Highland, for instance, a game addict, made a compelling short movie about his personal experience called *As Real as Your Life*[1]. It was showcased at TED[2] by David Perry during his talk and demonstrates the power of the skills mentioned in the list above (Please, watch the video). Statements like "My real world crumbles" or "When I am driving down a road at sunset, all I can think is: This is almost as beautiful as my games are" give pause for thought. Leaving the personal drama and ethical questions aside, this self-portrait documents that there are people whose minds actually live in a virtual world, they are not playing a game anymore. Certainly not with this catastrophic end in mind, but rather with the ambition to make a good game, such convincing virtualities are designed by humans and become primary mental models in the minds of others.

[1] `https://www.youtube.com/watch?v=fxVsWY9wsHk`
[2] `http://www.ted.com/talks/david_perry_on_videogames#t-1147362`

2.1 Fiero and Flow

One obvious assistance to virtual world designers, or for that matter any kind of story-tellers, to create a mental model in the mind of the audience is our understanding of the real world itself. No human is born with a complete model of reality. We learn on the way what things mean and how they work. In an early stage, we believe all things have their own kind of minds, even chairs and tables. When children cry, because they hit their head on the table, it therefore sometimes helps when parents speak to the table in question and warn it not to hurt their children anymore. In the eyes of a child, justice has been served that way and it calms down. The child's mental model of reality certainly is false, but it lives and judges by the rules of that model. Unless new observations or experiences are made that contradict the current model, that child is going to stay in that reality. We are continuously refining our mental models of the universe, principles of social life and even the models of our true self. Though situated in a real world, we experience its reality only to the degree our current model allows for. Our ability for mental modeling empowers us to solve geometric problems or simulate life choices without hurting anybody, but it also opens the mind to any not true, but believable story, artificial world or experience, for bad and for good.

However, simply offering a virtual environment to play in is not enough to immerse a player or to hook him for a consecutive session. Hence, game designers try to engineer the sense of *fiero* in players navigating their mental worlds. Fiero is an Italian word for *intense satisfaction in achieving an aim.* Humans inevitably show it in epic-win situations by throwing their arms in the air, often shouting. We do not feel fiero winning a chess match against a week opponent. It has to be a nearly unsolvable challenge which we engaged really hard and long for, such that it feels like we earned it. The winning goal at the last minute of the finals of a world championship by a fouled player would be considered a proper candidate, or the sudden clarity of mind in an aha-moment after studying a particularly difficult topic for several hours. Biologically fiero correlates to very high dopamine levels in several reward centers of the brain at the same time, particularly in the mesocorticolimbic system. Besides playing a part in the creation of joy, this brain area also acts as an addiction circuitry for addictions resulting from extreme motivation [HWK+08].

In a video game, the path to such an experience leads through the *flow channel*. The concept of flow was initially investigated by Mihaly Csikszentmihalyi in his book *Bejond boredom and anxiety* [Csi75]. It is the delicate state of full engagement in a fun activity which completely absorbs the mind. Its characteristics are:

- Intense focus on the present moment
- Merging of action and awareness
- Loss of reflective self-consciousness
- A sense of personal control or agency over the situation or activity
- A distortion of temporal experience (The subjective experience of time is altered)
- Experience of the activity as intrinsically rewarding

Sometimes it is referred to as *the zone*. Flow, of course, is not limited to video games, it as well occurs in sports activities and even in wood carving encounters, but according to Csikszentmihalyi, it most reliably and efficiently is achieved in games.

Game designers try to build games with as much flow time as possible. One key circumstance to flow is the balance between the presented challenges and the skill-set of the person facing them. Challenges that feel too hard for the current skill-set cause anxiety of failure. So, we give up or do not engage at all. Challenges that are too weak for the personal skill-set cause boredom because we master everything with too much ease. But adjusting the skill to challenge ratio once is not enough. Since skills get enhanced over time, the challenges have to change accordingly to keep the player in the flow channel on his way to the epic win. Please see Figure 2.1 for an illustration of this channel climb.

Not every video game is designed with such strong emphasis on reward and dopamine release, but neither are games designed to just fill our leisure time. Game designers do not create a set of rules, a virtual world, some characters and a mission. They aim to create a *transforming experience* in

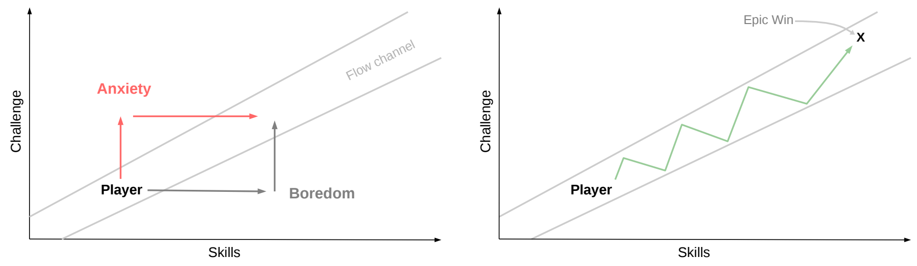

Figure 2.1: The flow channel. LEFT: The red path shows a too sudden increase of challenge complexity, which results in a feeling of anxiety. The only way to get back into the flow channel would be to train and enhance one's skills, but gamers could get too frustrated at that moment and leave the game. The gray path follows a gamer whose skills evolve faster than the challenge complexity until he also leaves the channel. His way back is a sudden increase of challenge complexity, but he may leave the game before that because of boredom. RIGHT: The challenge complexity is raised within channel boundaries, so that the user feels tension, but still can believe in being able to master it. After he masters the new challenge, he is given some time to relax, but also within channel boundaries. Before getting bored, a more complex challenge is presented (e.g. less time for the same task, more enemies, more skilled enemies, etc.). The ultimate, last challenge, if won, then creates the sense of fiero.

the mind of the player. Jesse Shell spends two chapters of his book to make that clear to the game designer in training.

2.2 Serious Games

Having seen the impact of a carefully designed virtuality on our emotions and on our perception of reality, we now ask whether this transforming experience can be extended to skills. If games are that transforming, maybe actively and consciously trained skills in an alternate world can also be transferred into real life.

In her book *Reality is Broken* [McG12] Jane McGonigal notices that an average young person will spend $10,000$ hours gaming by the age of twenty-one. This is approximately the same amount of hours spent in school when graduating from highschool. She then investigates the learned skills during this gaming and argues for society's engagement to harness these gaming skills for serious real-world problems. Depending on the game, gamers learn to organize meetings, to analyze human and system behavior, to solve logical and even mathematical problems, to plan ahead, to be resourceful, to work under stress, to communicate clearly, to interact with very different people compared to themselves, to think spatially, to write manuals, to program etc. Such skills, learned during and around gaming, seem so promising for businesses and educational organizations that in 2010 a new movement of *gamification* started [DKND11]. In order to induce such high levels of engagement in employees as seen in gamers, this community, with its own conference[3], tries to adapt game strategies to real learning and work environments. Often, an alternate reality with rules, titles, badges, and missions is created for that purpose. This reality then can either be visited with the help of a rendering device, or it has to be believed in in some other way.

Quest to Learn[4], for instance, is a school in New York where students identify themselves with great mathematicians, writers, evolutionary biologists etc. Instead of simply learning facts or techniques, they set out for certain missions, to crack mathematical codes, or to solve environmental problems. Some missions are secret and have to be found at first. Students level up on becoming a master of a specific skill etc. The whole curriculum and reward system has been reorganized with the help of game designers to spark engagement. It heavily uses technology. Another example, not only for schools, is the simulation game *World Without Oil*[5]. Here participants for a month develop personal strategies how to act in a time of a global oil crisis in real time. Catastrophic events are presented during their normal day activities and they have to decide how to act upon them. For this educational game, McGonical reports a true transfer of virtually learned skills to behavior in actual oil crises.

Gamification is not about introducing fun to otherwise boring or unpleasant tasks, it is about applying advantageous effects of virtuality and gaming to true reality, including true curiosity. Enabling learning for real life, it actually is about games for change. Unfortunately, change in habitual behavior,

[3] `http://gsummit.com`
[4] `http://q2l.org`
[5] `http://worldwithoutoil.org`

feeling or thinking is hard to achieve in general. To learn something new, an internal motivation barrier has to be overcome. But, since games seem to take the seriousness out of tedious tasks, learning the same thing in a game often feels not as hard as it would out of the game. In fact, it is the same problem, only well structured and clearly communicated as a series of achievements. We believe that if something is a game, it can be won. Since we know that games are designed, we believe that they are not too hard. Therefore, we build stronger confidence and are more likely to get to the goal of changing our knowledge or our faculties.

Change of thinking patterns is especially required for patients in neuro-psychological treatment. Studies like [Uni11, HJCBD09] support the application of games for such a cause and while healing syndromes of Post-Traumatic Stress Disorder (PTSD) by Tetris seems arbitrary, Matthew Fish, in his doctoral thesis [Fis14], proves that several casual games can even outperform pharmaceuticals as treatment for depression and anxiety. McGonigal, being a game designer, created the personal game *SuperBetter*[6] to heal herself from depression resulting from a concussion she had suffered before. Its effects were studied in [RJ13, RJR+15] and found to generalize across the population. These examples and many other clearly point at games as a serious instrument not just for education, but in the medical field, where change of thinking is required, as well.

2.3 Virtual Therapy

Unlike gamers, most neuro-psychological patients not as much strive for an in-game epic-win. Their disease is serious business to them, often making real life unbearable. This is not the time to play, many think, much less the time to believe in change by virtuality. Doctors suggesting a virtual therapy (VT) after a brain injury are frequently met with skepticism at least. After all, it is "just a game". As said in the beginning of this thesis, scepticism toward technology can root very deep. The current image of games in public does not benefit the situation either. VT still is at its beginning and its healing capacities, though well documented, have not reached the public

[6] http://www.superbetter.com

awareness yet. This section summarizes some arguments for VR in therapy compared to traditional approaches.

Although VR has been called a technology in search for an application [Wan96], according to [RWB98], successful treatments were reported for different kinds of phobias, psychopathological dysfunctions, pain reduction, stress reduction, eating disorders, erectile dysfunction, rehabilitation of memory, attention, spatial skill, executive functions and everyday-life training before 1999. These trials encompass children, young people, some elderly, patients with physical disabilities or mental deficiencies, autistic persons, burn victims, and terminal cancer patients. A repeated review of the literature by Riva in 2005 [Riv05] revealed no new application fields for VT, but confirmed the findings from before. In more recent years VT success includes motor and balance functions, for instance, in stroke [STM+10] or Parkinson [LRZ+12] patients, and even serves as occupational therapy [Hal08] in rehabilitation centers.

Certainly VR is not a good fit for all psychological or neurological issues. Instead, as pointed out by Rizzo [RWB98], it surmounts conservative therapy in specific areas where the following attributes are desired for treatment.

Exposure Anxiety or fear of any kind can be reduced by exposing the patient to a provoking stimulus. Gradual exploration and a controlled presentation of the feared stimuli lead to change. Traditional approaches like desensitization, implosion therapy, or flooding internally are based on an exposure model. PTSD or Obsessive Compulsive disorders can be treated this way. VR can expose patients to situations gradually and in a safe self-controlled environment.

Active distraction For pain management, a patient can be involved in some interactive VR that draws his attention away from his perception of pain. Being heavily focused on a task in VR, wounds of burn patients can be dressed less painfully for example.

3-D visual stimuli To assess cognitive functions like attention or visuospatial thinking, a dynamic 3-D environment can be presented to the patient. Rotation tasks and object combination exercises can be designed very easily.

Involvement in interactive activity Training of executive functions, which consist of planning, sequencing, sustaining attention, resisting interference, coordinating simultaneous activities, cognitive flexibility, and the ability to deal with novelty, is a strong candidate for VR. Real-world environments for such exercises cannot be controlled with the same certainty as in VR.

These advantages are convincing from the doctor's point of view or for patients aware of their situation and of the need to actively create change about it. But in reality, many patients are so desperate about their situation that they surrender themselves to the circumstances and become unwilling for change. Traditionally, therapists then try to trigger the cognitive process of *empowerment* to reactivate self-motivated engagement [Men99]. Empowerment is a multi-faceted construct emerging from the overall effect of three dimensions to which, according to [Riv03], VR as well contributes beyond conservative approaches:

1. **Perceived control** includes beliefs about authority, decision-making skills, availability of resources, autonomy in the scheduling and performance of work, etc.
2. **Perceived competence** reflects role-mastery, which requires the skillful accomplishment of one or more assigned tasks and successful coping with non-routine role-related situations.
3. **Goal internalization** captures the energizing property of a worthy cause or an exciting vision (like an inspiring mission in a game).

Surprisingly similar to the process of learning, neuro-psychological change is achieved by an intense focus on a particular set of facts or an experience [Wol02]. An elaborate exploration from different perspectives, for instance, allows the patient to relive situations in his imagination or reinvestigate mindsets making his internal convictions available for reorganization. The common goal of therapy is to identify internal representations of reality that do not match actual reality and help to adapt them. Particularly the English word *imagination* emphasizes the role of images in contrast to mere thought for this kind of treatments. Since the early beginning of clinical psychology, therapists base their techniques on the analysis and modification of mental images. Freud, Jung and Adler regarded them as one major means to access the unconscious. As surveyed in [VM98], mental images hold even the capacity for a strong somatic impact. They can increase heart rate, modify pupil size, increase the level of glucose in the blood, modify muscular tension, alter skin temperature, etc. The survey also mentions that the processing of sensor functions excites the same cerebral area as pure imagination does. Further, it reveals that also in regards to quality of experience imagination and true perception cannot be distinguished with absolute certainty. Finally, the imaginative system is found to be the core of our psyche, which mechanics of recovery, elaboration and response to external stimuli are based on. VR presents itself as a perfect fit, feeding a patient's imagination during such treatments. In comparison to mental imagination techniques, VR offers a precisely controllable, revisable, content-rich environment with the opportunity of guiding attention to every little detail of the presented images throughout the entire therapeutic session. Especially the provided interaction intensifies the experience of reliving a situation in one's imagination in order for change.

From the neuro-scientific point of view, interacting with the surrounding environment has its own fundamental effect on the brain. In [RABJ98] Rose et al. elaborate on rehabilitation from brain injuries caused by traumatic incidents, stroke or neuro-degenerative conditions. They highlight the ramifications *environmental enrichment* has on the functioning of the brain. Healthy rats held in an enriched environment (housing including wooden blocks, plastic tubes, table tennis balls etc.) superseded their fellows held in smaller housings without any "toys" to play with in both *complex learning* and *problem solving*. This superiority as a result of enrichment holds for rats during rehabilitation after brain damage as well. In a maze-learning task, rats postoperatively housed in a non-enriched environment performed significantly worse compared to their control individuals. Similar experiments on birds, dogs and primates indicate no principal change of results. Rose et al. then extend these findings to humans with brain injuries, but without experimental proof. The benefits of enriching social, physical and mental activity is commonly acknowledged by clinicians, but not efficiently employed. Recent publications like [JAB+12] still criticize the vast amount of time stroke patients spent disengaged or inactive during their stay in a clinic. In [Raj13] Rajkumar argues for the pressing need of more environmental enrichment in medical facilities in general to overcome patients' isolation and passiveness. Clearly, there is a professionally founded demand for more interaction. Virtual reality can meet exactly these demands, even for patients restricted in their mobility. As a matter of fact, interaction is the major feature of VR. Same as for imagination, Rose et al. claim that virtual and real interaction trigger broadly similar cerebral areas, such that general transfer of interaction effects should occur. However, since transfer is a complex process, some questions remain unanswered and a definite statement about interaction transfer cannot be made.

Training any cognitive abilities in a therapy, of course, has the immediate goal of regaining attention, memory, spatial orientation, problem-solving abilities, motor skills or the like, but the guiding overall motivation is always to enable or improve on the quality of everyday life. Virtual simulations of corresponding *meaningful* situations are very well documented to result in effective skill transfer to real life. Apart from therapy, the army, for instance, employs flight-, naval-, and groundbattle-simulations to train soldiers for real combat. Businesses train their factory employees to virtually operate the machines they are going to handle in reality later. Virtual therapy settings for impaired persons easily encompass meaningful situations like road-crossing, visiting a bakery, a post office, a supermarket, moving within a house, driving in a virtual city with traffic, using public transportation, navigating according to a map or path description (see [MPL98] for the corresponding trials). Mastery of such "trivial" situations represent independence and self-sufficiency, a feeling very strongly contributing to the perceived quality of life. Tasks with personal importance energize us, no matter how immersive the environment is. Our focus intensifies and we are more likely to enter the flow channel. If not, we are at least more eager to learn and remember the new knowledge or skill for the long run. Interestingly, Shell [She08] also mentions *personal meaning* as a key concept to drive motivation in games.

Carefully observing the presented concepts of game design, including the advantages of VR, and comparing them to the basic strategies of neuropsychological therapy, we notice a strong resemblance. Both enterprises strive for a deep impact on the mind and, therefore, share the same tactics and insights into the human psyche. Please see Table 2.1 for a summary of

VR/games	**Therapy**
Loss of time *awareness* and *selfconsciousness* during flow	Active *distraction*, e.g., for pain management
Players' psyche is *transformed* through games (even to addiction)	Goal to *change* mental abilities
Personal *control* over activity or situations and *solvable* challenges during flow	Perceived *control* & *competence* during empowerment
Clear goals, exciting missions	*Goal internalization* for empowerment
Create an experience in the mind of a player by building a *mental model* in it and even induce emotion	*Imagination* as strategy to identify internal misrepresentations of reality and to reorganize ones convictions
Interaction as the major feature of VR and video games	Environmental *enrichment* enhancing physical, social and mental *interaction* in order to accelerate rehabilitation
Intrinsic reward as a major issue to generate flow and the addition of personal meaning to improve on impact	*Personal, meaningful* situations to amplify willingness for training

Table 2.1: Comparison of basic aspects of games and VR to basic elements of neuro-psychological therapy. Therapists and game designers face surprisingly similar problems with the inner-working of the human mind and therefore can contribute strongly to each other's fields.

the prominent relations rendering VR as a seemingly perfect support tool for therapy.

Obviously, a pure in-game or in-VR epic-win situation is neither a patient's goal nor that of his or her therapist. The great enthusiasm seen in gamers may not emerge in a virtual rehabilitation atmosphere to the same degree, but the striking resemblance of therapy goals and game effects nevertheless promises a true-life epic win by therapy success.

2.4 Entering Virtuality

So far, we have seen the concept of flow as the major tool for guiding people from a real place to a virtual one and keeping them there. Flow depends on the balance of challenges to personal skills and requires constant activity on the user side. It is driven by tension and relaxation of action. In the scientific context of VR, however, we read about *immersion* and *presence* instead of flow to achieve the same goal.

Researchers from the computer sciences often refer to the state of "being there", meaning in a virtual place, as "immersion". Representatives of the psychology departments, on the other hand, tend to use the word "presence" and differentiate it with respect to immersion. Literature defining presence mentions immersion as one of many ingredients to create presence.

Presence is described as a spatial immersion into a mediated environment [Sla03]. It means the state of mind, the subjective sensation, and the feeling of being in one place though being in another. To get psychologically immersed into that place, as many real world sensations as possible have to be believably replaced by technology. The objective extent to which this is done by a VR device is then referred to as immersion.

In contrast to flow, in presence it is not the activity which lets the user leave the real world, but the mediation of the environment and place itself. Weibel and Wissmath [WW11] compared the two phenomena and found that flow and presence are clearly distinct psychological states, but correlate to each other. They share common properties like attention or immersive components but aim at different targets. Flow focuses on task characteristics and is the result of the immersion into an activity, whereas presence focuses more on media characteristics and is the result of the immersion into a place. Weibel and Wissmath find that task performance in VR depends on both, flow and presence. In general, it is believed that the experience of the however generated sense of "being there" strongly contributes to learning success and skill transfer.

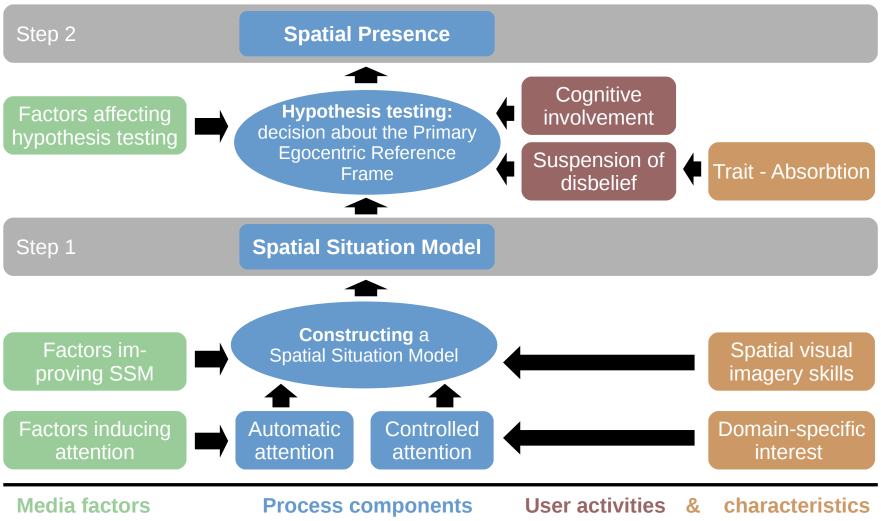

Figure 2.2: Two-step presence process by [WHB+07]. The blue path depicts the formation of spatial presence. The media factors to the left and the user activities and characteristics to the right, both affect the creation of the Spatial Situation Model (Step 1) and the following decision about the Primary Egocentric Reference Frame (Step 2).

The crucial insight from this section is that there are different means to leave the physical world and to enter the realm of authored imagination. To create a holistic and transforming experience, several factors have to be considered for the mediating technology, the task, and the participating individual's mind.

2.5 Spatial Presence

In the past, there has been public disagreement on the understanding of the concepts of presence [SUS94, WS98, WHB+07, SB11]. (See [Sla99] for an explicitly harsh example of opposition). As a result, in her report [You03], Youngblut lists 32 different questionnaires used to measure presence in research. The different theories in psychology supporting these questionnaires are summarized in [LBF+15]. One of the major and more recent approaches is a two-level process to presence by Wirth and colleagues. See [WHB+07] for the original paper.

Their two-step process, as summarized in Figure 2.2, is a more general approach to presence in media. It encloses reading of books, watching movies and interaction with a VR. It combines the situative concept of traditional presence with the activity focus of flow. The first step starts by paying attention to the mediated environment. Attention has an *automatic* part sensitive to media factors and a *controlled* part sensitive to the participant's psychological characteristics. From the acquired information, then, a *Spatial Situation Model* (SSM) is constructed in the imagination. It is the mental representation of the depicted scenery, including distances, size and texture of objects, their spatial relation to each other, etc., which is continuously updated throughout the exposure time. In the second step *Spatial Presence* not simply emerges from the SSM, but rather is a result of a decision whether the *Primary Egocentric Reference Frame* (PERF) is the virtual place or the physical. This decision guides further perception to or from virtuality and is a constant struggle throughout the whole experience. A book, for instance, can provide sufficient descriptions of a location to immerse the reader into the specific scenery, but if she or he does not accept this logic as her or his PERF, the book's content will be interpreted secondarily and the book itself will be located as an object in the physical world, which is then the current PERF.

Attention, construction of the SSM, and the decision about the PERF, all are thought to be affected by three types of factors:

Media factors The more spatial cues the media utilizes, the more they correspond to reality, and the more they induce attention to the content, the more stable and detailed the SSM is thought to become. The contributing factors include spatial sound, field of view, means of navigation, shadows, occlusions, and the like. During the second step, missing, contradictory, or bad cues can interrupt the process and break presence. Lag-time during an activity in VR or the missing thunder after a lightning, for instance, are possible causes for such a break.

User characteristics Personal interest in the mediated content, the story, or the task provides internal motivation, which strengthens the attention and enhances the construction of the SSM. Further, individuals with high cognitive abilities like spatial-visual imagery build stronger SSMs. For them, it is easy to fill the gaps eventually missing from the media representation. For the second step, the personal tendency to be absorbed by a task or an imagination highly affects the decision about the PERF.

User activities High cognitive involvement in a task is meant to benefit the hypothesis testing about the PERF. Intense focus helps to overrule distracting media factors, and arising disbeliefs are thought to be deliberately suppressible.

The beauty of the two-level model is the inclusion of personal traits and its hierarchical structure. In combination with the clear separation of traits, activities, and media factors it is an elegant way to account for the diversity of presence experiences. Depending on the individual, media imperfection may be compensated by mental, imagery supplement, or personal interest. Vice versa, weak imagery skills and no interests in the domain can be overcome by a comprehensive mediation through technology building a persuasive SSM. In the second level, even a deficient SSM can be actively chosen as the PERF, which often happens for VR devices with small screens, low resolutions, unrealistic graphics, or for absolutely abstract content as long as the task at hand is involving enough. Reversely, a static virtual place, requiring no involvement from an observer at all, can immerse a person completely because of the realism of the presentation.

For the aspects of VR, as listed in Table 2.1, that promote therapy, the two-level model provides a comprehensive explanation of their construction with respect to technology, personality, and activity. Empirical evidence for this approach to presence can be found in [HLCJ12, HWK+12, VWGO04].

2.6 Measuring Presence

Knowing about the emergence of presence and its possible relations to therapy, we now look into its measurements. In general, two different types are proposed in the literature [You03]: *subjective* and *objective*. Since, at the time being, there is no definition of the concept of presence which is commonly agreed upon, measurement techniques are under a constant development as well.

2.6.1 Subjective Measurements

Subjective measurements require some sort of self-report from the participant. Mostly, this is a questionnaire filled out after the VR experience. Such measuring assumes that every participant has a trustworthy self-awareness and memory of it. Otherwise, results could not be compared. This, of course, cannot be guaranteed and poses a source of measuring error. Additionally, as already mentioned, Youngblut finds 32 different questionnaires due to different understandings of presence and contributing factors. Using two different questionnaires on the same VR experience, thus, can result in different results for the degree of perceived presence. This was demonstrated even for the two most widespread questionnaires, the Presence Questionnaire by Witmer & Singer [WS98] and the Slater-Usoh-Steed Questionnaire [SUS94]. Another pitfall was found by Usoh et al. [UCAS00]. They conducted an experiment for two separate groups. The same task was performed in a real room by the first group and in a virtual room by the second. The reported presence levels should favor the real room condition, but did not. Their conclusion from this insight is that questionnaires can only discriminate levels of presence in within-subject studies, where participants experience all different environments or devices in question and rate only in comparison.

An interview, where the questions are not standardized and the examiner has the opportunity to derive his next question from the last given answer, is most suitable to investigate a VR experience. The risk of misunderstanding a question is minimized this way. Also, the participant may have noticed aspects that she or he can voluntarily report. Unfortunately, due to the uniqueness of every interview, no true comparison can be calculated between the reports and a quantitative statement cannot be concluded. However, this type of measure particularly qualifies for initial studies to understand the general appeal of a specific VR device and to identify its weaknesses.

The *Virtual Presence Counter* by Slater and Steed [SS00] tries to compensate for the disadvantage of questionnaires being filled out after the VR experience and depending on correct recall. Here, indications about breaks in presence have to be made online during the trial. Every time the participant recognizes or reacts to a stimulus from the real world, she or he has to announce so. These announcements are counted and related to presence. While overcoming the memory problem, this approach introduces a potential bias on the participant toward awareness of the real world.

Despite the inaccuracy, the dependency on memory, the relying on a correct understanding of questions, and the restriction to a relative rating, almost every study, known to the author, evaluating a new VR device employs some kind of questionnaire. Though quantitative measures may be questionable, qualitative assessment of presence can be assured at least for the popular questionnaires. Also, some VR properties (e.g., mental models, fun, felt realism, intuitiveness of control) simply cannot be directly measured in an objective manner.

2.6.2 Objective Measurements

Noticing the disadvantages of subjective measurements, several objective physiological measurements were proposed. Youngblut lists the following:

1. Heart Rate
2. Skin Conductance
3. Skin Temperature
4. Respiration Rate
5. Mental Workload
6. Muscle Tension
7. Eye Scanning
8. Eye Blinking Rate

These properties are monitored online, as aimed at in the Virtual Presence Counter approach. Moreover, it is a continuous real-time measurement. Analyzing this data and relating it to different virtual events during a single VR experience is thought to reveal presence behavior on a timeline. Aside being an online measurement, physiological quantities are independent of any participant and examiner bias. Heart rate and skin conductance seem to be the most reliable indicators associated with presence, though some inconclusive results are reported for them as well.

Another line of objective measurements focuses on the participant's behavior. Facing a fearful situation in VR, for instance, only causes corresponding reactions if the situation is believed to be real in the sense of presence. True signs of fear in facial expressions or other bodily reactions are therefore interpreted as data points. A further example is to ask the participant to point to a certain item when that item is represented in the virtual environment and in the real room as well. Depending on his choice his PERF becomes clear. The drawback of this technique is its sensitivity to tasks. Not every activity in every virtual environment generates significant behavior that can be recognized by an observer and unambiguously mapped to presence.

2.7 How Much Presence is Enough?

Obviously, we wish for a precise measurement of presence to establish an in-depth comprehensive relationship to learning and skill transfer. Unfortunately, since every presented measurement suffers from at least inaccuracy, a safe formulation of learning and skill transfer as a function of presence has not been found yet.

Instead of understanding the role of presence in all its nuances, some researchers focus on practical output. Research in task performance, which is traditionally thought of as a function of presence, has produced inconclusive results. Youngblut's literature review about task performance and presence [You03] reveals that 50% of 82 findings during different studies show no correlation between presence and task performance. She argues that tasks can become automatic over time or because of their simplicity. In such cases, presence becomes irrelevant with regard to task efficiency.

Bowman and McMahan [BM07] found that task performance depends only on specific task fitting immersive components. In contrast to the traditional view, they propose a multidimensional approach to the effect of presence on task performance (Figure 2.3). Traditionally, it is thought that any

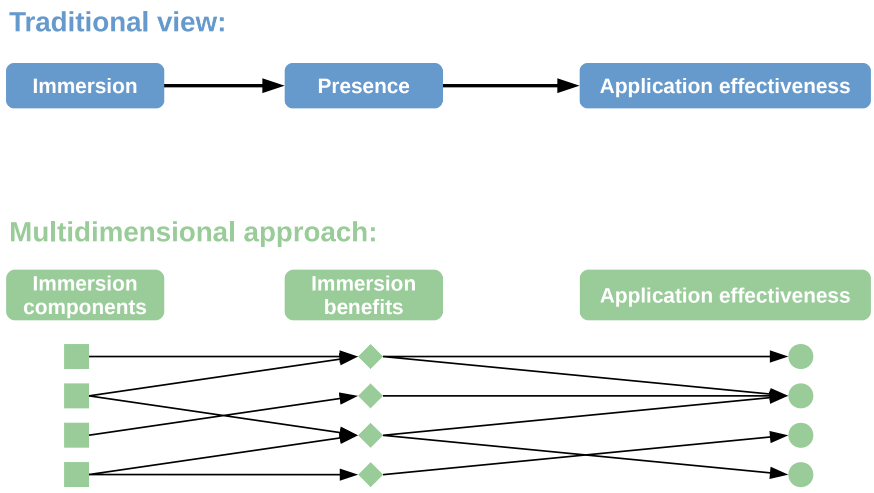

Figure 2.3: Benefits of immersion according to [BM07]. At the top, the traditional view is depicted, where good task performance results from high presence which builds from broad immersion. At the bottom, a specific task performance enhances because of a subset of fitting immersion benefits which themselves result from a few responsible immersive components.

task benefits directly from presence, which itself is a result of immersion and some further factors. Bowman points out that immersion, besides enabling presence, generates task specific benefits like *spatial understanding* or *peripheral awareness*. These benefits emerge from only a subset of the available immersive components. Spatial understanding, for instance, is supported by stereoscopic rendering, but not so much by the realism of the scene. In some cases, scenes are even rendered in an abstract manner to enforce spatial understanding. Spatial understanding, in turn, is not required by every application to the same perfection. Escaping from a maze or path-planning for an oil well certainly benefit more from spatial understanding than hitting a baseball does. The multidimensional approach explains why certain VR tasks can be performed perfectly despite the absence of a full sense of presence. Nonetheless, Bowman notices a higher relevancy of presence for applications concerned with training transfer. If the goal is to train actions or reactions for a real environment or for certain real circumstances, the sensory stimuli of the VR device should match the real world as closely as possible.

Furthermore, relaxing the importance of permanent presence, Mantovari and Castelnuovo [MC03] suggest that efficient learning and mental change in VR is even better achieved through reoccurring dynamic shifts between a strong immersion and a reflection of recent, virtual activities from a different and external point of view. Comparing to learning in real life the same pattern can be recognized. Having focused on a hard topic or having practiced a new skill for some time, it helps when we emerge from it and revisit the task from an outside perspective.

For a success of virtual therapy, presence is only one ingredient. Even if a VR device is capable of delivering the highest sense of presence, the patient's willingness and trust are the decisive factors. Building for a certain type of therapy, a VR setup has to balance technology, personal traits and deficiencies, and the task in a smart cooperating way.

2.8 Outlook

After describing the strong meaning that game designers attach to human psychology during their design process, this chapter elaborated on the impact of VR on the human mind. We have seen that virtuality has the power to change a human mind for bad and for good. Virtual therapy can benefit from many aspects VR offers. The resemblance between traditional therapy techniques (imagination, control & competence, personal meaning, etc.) and VR/gaming features (mental experience, solvable challenges, intrinsic reward, etc.) supports a successful application of VR in therapy. Finally, we investigated the concept of guiding an individual from the real world into the virtual one and focused on the sense of presence, its measurements, and needed amplitude for task performance and training transfer. The following chapters will present a novel VR device for virtual therapy in this light.

Chapter 3

The OctaVis

I open my eyes,
and I see pixels all around me...
— Samantha Ackerman

The OctaVis system is a novel virtual reality platform, which was developed in the highly interdisciplinary medical project *CITmed: Cognitive Interaction Technologies for Medical Applications.* Its major goal was the diagnosis and rehabilitation of patients with brain function disorders, as they might result from stroke, cerebral traumas caused by accidents, neurological or psychiatric diseases. In particular, the emphasis was on higher cognitive brain functions such as memory, spatial orientation and navigation, as well as executive functions like path planning.

In traditional neuro-rehabilitation, the process of re-learning such cognitive abilities is done by standardized paper-pencil or card tasks. Unfortunately, the improvements gained with these methods do not sufficiently transfer to real-life scenarios. The reasons for that are: (1) the cognitive functions are trained isolated and (2) the tasks are rather abstract and far from the problems in daily routine. To compensate for these two obstacles, therapy to be successful has to face the patients with training tasks that (1) are utilizing several cognitive abilities simultaneously and (2) are as close as possible to the daily real-life routine of the individual patient.

The medical institutions, with decreasing resources in staff members, treatment time, and labor costs, often cannot realize such an individualized training. For patients with navigation problems, for instance, to relearn and memorize the way from home to the hairdresser, accounting for attention at road crossings etc., a one-on-one supervision would be necessary for several consecutive days. No public hospital can provide such intense care.

However, in the previous chapter, we have seen that VR devices can be facilitated to create persuasive complex real-life situations in our imagination. Knowing this advantage over the traditional paper-pencil approach and with the working therapy tactics in mind (Section 2.3), we turn towards a proper virtual task fitting the two therapy criteria.

In accordance with the importance of a *meaningful task*, the VR scenario chosen for training the mentioned cognitive abilities is grocery shopping in a virtual supermarket. It is a real-life scenario contributing very strongly to the perception of self-sufficiency. Therefore, its mastery is a natural motivation for most patients, especially elder ones. Being involved in such a training patients get distracted from sorrows or loneliness and become active in the sense of *environmental enrichment*. Furthermore, training in a virtual environment, thereby excluding eventual public embarrassment (e.g. getting lost, taking too much time) or life-threatening consequences of failure (e.g. due to a fall), feeds internal motivation in the sense of control and competence, as aimed for by the process of *empowerment*.

In particular, to train memory, spatial thinking, and the executive function of path planning, patients have to memorize a list of items, have to navigate through the supermarket in order to find and buy them, and should improve their path through the supermarket over multiple training sessions.

3.1 Requirements

As much as externally possible by a task, in addition to the compliance with therapeutic strategies, the grocery shopping task should considerably activate the *user activity* segment and the *user characteristics* segment of the two-step model of presence described in Figure 2.2. Looking at the *media* segment, especially for real-life situations highly immersive VR technology has been shown to considerably improve the transfer of training success to real reality (see, e.g., [RBR05]). A high level of technical immersion, feeding as many queues as possible that are needed for the training of grocery shopping, therefore was a major requirement to be met in this project.

On the one hand, the components providing immersion had to be tailored for a smooth performance of exactly this task, but on the other hand, they should assure a more general use for diagnostic purposes and further investigations of spatial thinking, memory, and executive functions as well.

Although many VR systems provide a sufficient level of immersion for the kind of rehabilitation training as thought of in this project, they do not qualify for clinical studies since they do not meet the following crucial requirements:

Ease of use: Most existing VR systems are used by VR experts in academic or industrial research labs only. In contrast, the OctaVis has to be used by patients, typically being elderly people suffering from a stroke and without any prior computer or VR experience.

Maintenance: Since the system has to be operated by clinical staff, which typically does not have a strong technical background, it should be as easy as possible to operate and maintain. This is in strong contrast to complex VR systems like CAVE installations, which are typically driven by a high-performance network of distributed render clients and therefore require experienced specialists.

Cost efficiency: To perform clinical studies in several hospitals or to provide the VR system to rehabilitation clinics, the system has to be reasonably cost efficient.

Figure 3.1: Two photographs of the OctaVis system. Eight screens, arranged in an octagon, provide a 360° panorama visualization of the virtual environment. Two door segments can be opened (right). Navigation in the VR is performed through a modified office chair, whose orientation determines the movement direction, and a "throttle joystick" in the armrest. Easy and natural interaction with objects is enabled through a simple touch screen interface.

Space efficiency: Similar to financial budgets, space or rooms are typically also limited, requiring a VR system with an as compact as possible spatial footprint.

Medical requirements: Patient safety is the highest goal in any clinical study. Since many stroke patients also suffer from hemiparesis to a certain degree, they cannot be expected to stand or walk without support. Consequently, robust and save chairs must be employed. Moreover, the clinical staff has to be able to monitor and supervise the experiments, as well as to intervene at any moment.

Meeting the above requirements and then even being successfully deployed to four hospitals for clinical studies, in the following the OctaVis system (Figure 3.1) is described in its details. Aiming at a diagnostic and rehabilitative purpose of the device, the emphasis of the description lies on the overall-view of the system (Section 3.3) as well as on the evaluation of its level of provided general sense of presence (Section 3.4). In-depth investigations of the rendering software, spatial thinking support, and task performance follow in subsequent chapters.

3.2 Related Work

This section discusses related work on VR systems in terms of hardware architectures. A high level of immersion is the major goal of any VR platform. As a consequence, many studies have been conducted in order to determine the various technical factors contributing to the sense of presence.

Although there are concurrent theories about the very nature and the measurements of presence in VR (Section 2.5), this line of research agrees upon the two major technical causalities: (1) the *presentation* of the VR to the user and (2) the scope and quality of *interaction* in the VR. The presentation is perceived more realistic as more senses are stimulated in a consistent manner as real reality does (e.g., unlimited field of view, high resolution, surround sound, sufficiently detailed models and realistic rendering). User interaction is perceived as natural if variety and physical movements mimic actual reality without abstraction layers like controllers (e.g., real walking, turning, gestures, touching).

We have already seen some objections against the permanent run for the increased sense of presence in general (Section 2.7). Bowman and McMahan [BM07] particularly criticize a blind run for realism in presentation and interaction and argue for an appropriate subset of rendering features and interaction metaphors involved in the task the user is supposed to perform in VR. This insight allows us to find successful compromises between as realistic as possible presentation and interaction on the one hand and ease of use and patient safety on the other hand.

The following sections discuss existing hardware architectures that provide an omnidirectional field of view, before looking at navigation and interaction interfaces suitable for a virtual supermarket scenario.

3.2.1 VR Presentation

In order to optimally trick the visual sense in VR, the presentation device has to provide a high-resolution omnidirectional view of the virtual environment. Many VR systems employ head-mounted displays (HMDs) since they provide a seamless horizontal and vertical 360° view. Oculus Rift [LLC16] and HTC Vive [Cor16] are currently the most sophisticated HMDs on the market aiming at such a total visual immersion. Unfortunately, HMDs lack of visual self-perception since body limbs remain occluded. It is possible to counteract such occlusions, for instance, by adding a virtual avatar, but such setups require expensive motion capturing hardware and may easily inflict distraction in patients. Additionally, depending on the condition, some patients are sensitive to or not able at all to wear a helmet-like device. Polcar and Horejsi [PH15] also have shown HMDs to induce more serious cyber-sickness symptoms to inexperienced VR users.

CAVE installations [CNSD+92] are known for providing a very high level of immersion, especially due to their large field of view, but they disqualify for the CITmed project because of high cost, occupied space, and maintenance effort. To improve space and cost efficiency, several MiniCAVE systems have been proposed [WSV+99, Sch08], which basically are CAVE-like systems with smaller projection areas and less space consumption. However, the space requirements of these systems are still too large for most clinical facilities.

Microsoft's MiniDome [BW10] avoids the typical sharp edges and corners of a CAVE system by employing a hemispherical projection. Unfortunately, the projectors are placed inside the dome and therefore cast shadows as soon as the user moves between them and the dome surface. While in their project this was the desired interaction pattern, it is not suitable for our project. Taking the dome idea one step further, cybersphere systems [FRE03,Bar16] project the virtual environment onto a seamless sphere. This is ideal in terms of field of view but does not fulfill the mentioned requirements on cost, space efficiency, and maintenance because these systems rely on custom-tailored projection surfaces. This is also true for circular systems like [HJHL08].

Google's Liquid Galaxy project [Inc16] presents the Google Earth data on eight flat screens arranged in a circle around the user. The use of displays (instead of back-projections) allows for a smaller spatial footprint. Their display circle is not closed in order to have an open entrance, such that the horizontal surround view is broken. Using ClusterGL [NHM11], each of their displays is driven by a dedicated render client, which increases cost and maintenance efforts.

The OctaVis system also uses flat screens, but these are arranged in a closed circle (octagon) centered around the user in order to provide a full 360° horizontal panorama view of the virtual environment (see Figure 3.1). In contrast to the Liquid Galaxy approach, the OctaVis is driven by a single PC, which effectively reduces cost and maintenance effort, as described in Section 3.3.1.

3.2.2 VR Navigation and Interaction

Besides the presentation of the VR, interaction with the virtual environment is the other critical factor for immersion. In this project interaction means navigating through the virtual supermarket and selecting/buying products.

Interestingly, a very similar VR supermarket scenario has been analyzed by Renner et al. [RDS+10], who conducted a series of experiments investigating different navigation metaphors (path-drawing, lean-based velocity, walking-in-place, world-in-miniature, scene-in-hand) and interaction techniques (ray-casting, image plane) in a CAVE installation. They found that *novice users failed* to complete the given tasks in reasonable time and therefore had to evaluate their techniques with participants having sufficient VR experience. In contrast, in clinical studies, no VR experience at all can be expected from patients. Typically these are elderly people with cognitive disabilities, e.g., due to a stroke. Hence, we have to find simpler and more intuitive methods for navigating in and interacting with the VR.

Small-movement navigation metaphors like mouse, keyboard, game pad, or wand are not suitable for novice VR users, because they introduce an additional abstraction layer for navigation, complicating the VR task to be performed and not being immersive.

The most natural way to navigate in a VR obviously is walking. The different metaphors like real walking [RL09], walking-in-place [SUS95] or within a rotating sphere [FRE03, MFW08] are known to feed both the vestibular and the proprioceptive queues, thereby generating high immersion. Unfortunately, these systems contradict the requirement for a small footprint. Moreover, patients with hemiparesis might not be able to walk or even stand without support.

Emulating the reality of a shopping experience Whitney et al. [WSB+07] employ a real shopping cart in front of a treadmill. Coming perfectly close to reality, this approach has to be discarded for the OctaVis for the same reasons as walking-in-place and for its overspecialization.

Allowing users to sit, ChairIO [BBH05] is a navigation idea similar to that of the OctaVis. Sitting on a sweeper chair one navigates through the VR by rotating the chair (controlling orientation) and leaning forward/backward (controlling translation). The VR presentation, however, is carried out only for a single frontal view, such that the virtual world has to be rotated around the user, thereby failing to stimulate the vestibular and proprioceptive queues in a natural way. Apart from the neat concept, the flexible spring balance attached to the sitting platform causes safety problems with patients.

In the OctaVis system, therefore, a robust rotating office chair is employed, equipped with a throttle joystick in the armrest (Section 3.3.2). The walking direction is intuitively controlled by rotating the chair into the desired direction and the joystick controls forward/backward movement (Figure 3.2).

While for object selection several interaction metaphors exist, e.g. using a wand, joystick, game pad, or pointing through finger tracking [RDS+10], they are typically either not easy to use or not immersive. Since in the OctaVis system the user is surrounded by flat screens anyway, it comes naturally to use touch screens, such that object selection can be performed very easily by a simple touch on the screen (Figure 3.2).

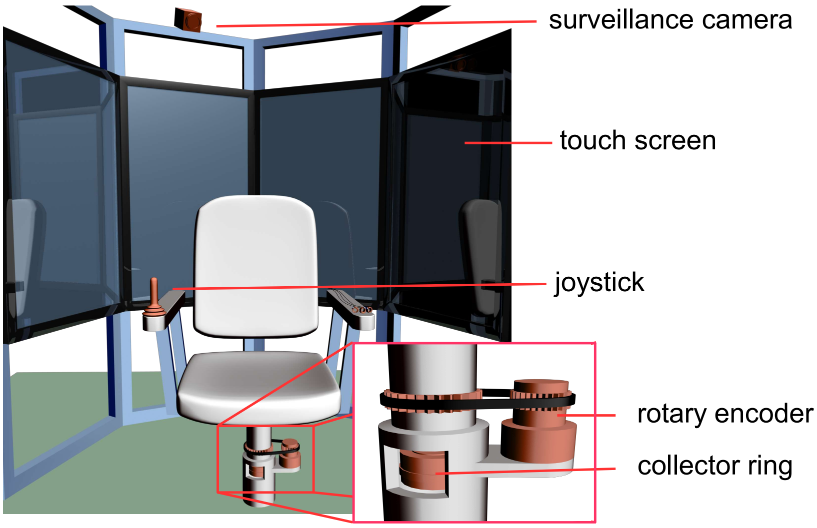

Figure 3.2: Illustration of the navigation and interaction components. The orientation of the chair, determining the movement direction, is measured by a rotary encoder and a "throttle joystick" in the armrest adjusts the speed of movement. Easy and natural interaction with objects is enabled through a simple touch screen interface.

3.3 The OctaVis VR-System

The discussions of the previous section showed that existing VR systems do not meet the requirements for the clinical studies listed in Section 3.1. Navigation and interaction in the VR have to be intuitive, but also feasible for patients of different age and disability level. In addition, setup times should be short and the VR-training has to be safe for the patient at any moment. Since the studies are supervised mainly by medical staff, the system should be easy to operate and maintain even without technical background.

The OctaVis system has been designed with these special requirements in mind. Similar in structure to the related work discussion, the following section describes the presentation of the OctaVis and then the navigation and interaction.

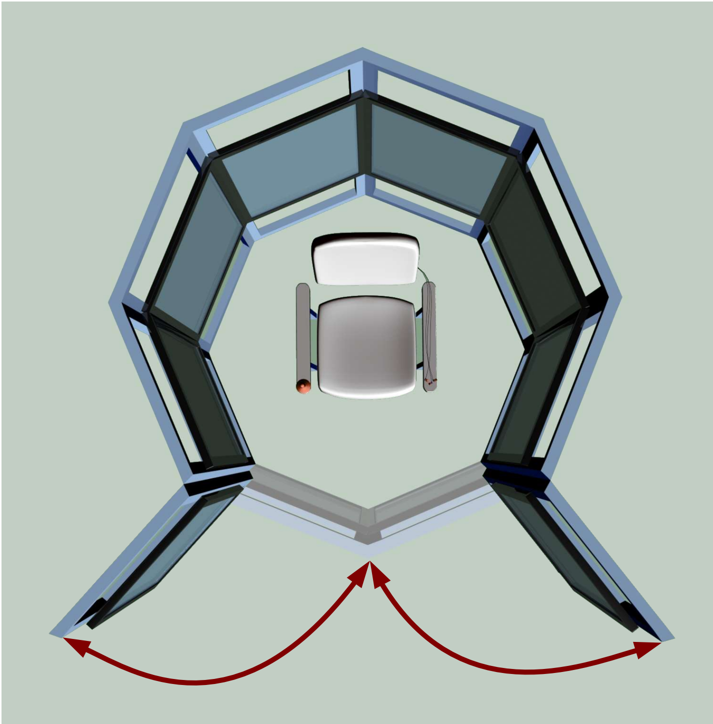

Figure 3.3: The OctaVis viewed from the top. Two displays act as door segments and allow to easily enter the system.

3.3.1 VR Presentation

In order to appropriately stimulate the visual sense, the OctaVis system consists of eight standard displays arranged upright in an octagon around the patient. This setup provides a full 360° horizontal view of the virtual environment. Each screen is mounted on an aluminum profile segment about an arm-length away from the patient, who is sitting on a rotating office chair in the center of the octagon. Two of the octagon segments are assembled as doors, providing an easy and safe entrance and exit for patients. See Figures 3.1, 3.2 and 3.3.

In order to enable easy selection of objects in the VR, touch screen displays are employed (*EloTouch 2639L* 26"). The eight touch modules are connected via USB (see Figure 3.4) and trigger standard mouse events. Hence, they are easily integrated into the VR framework (Chapter 4). While the touch option considerably simplifies user interaction, it comes at the price of bigger frames around the displays. However, it has been shown that seams between adjacent views do not influence performance in virtual real-

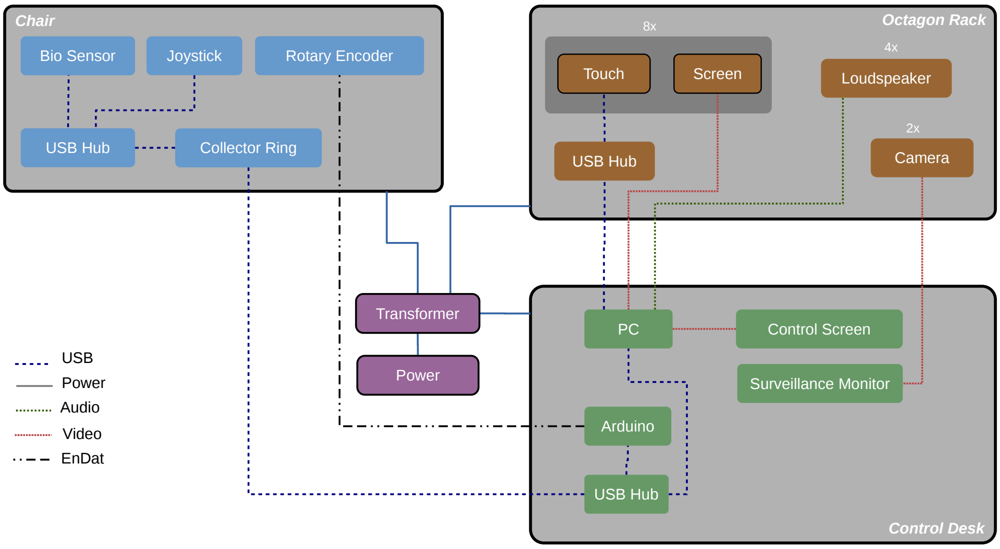

Figure 3.4: Hardware wiring plan, showing interconnections of the different modules. (1) The Chair module, containing the bio-sensors, the joystick, and the rotary encoder. (2) The Octagon Rack module, containing eight screens and four loudspeakers for presenting the VR. Also, two surveillance cameras for patient monitoring are mounted here. (3) The Control Desk module, containing the PC, the operator display for controlling the application, and a surveillance monitor showing the video signals of the surveillance cameras.

ity systems [MPS11], which is also confirmed by the study results in Section 3.4.

In contrast to most other VR systems, which typically follow a distributed rendering approach using one render client per view/screen, the OctaVis is driven by a single PC. As depicted in Figure 3.5, this workstation is equipped with three graphics cards, each of which is connected to three screens, resulting in nine screens in total: eight for VR presentation and one operator display. This single-PC multi-GPU approach considerably reduces hardware costs and maintenance effort. It also simplifies implementation of graphics-related functionality and reduces latency since no network synchronization is required like it would be for distributed rendering. Finally, since the system is just one (powerful) PC, it can be operated easily, not demanding any special technical background.

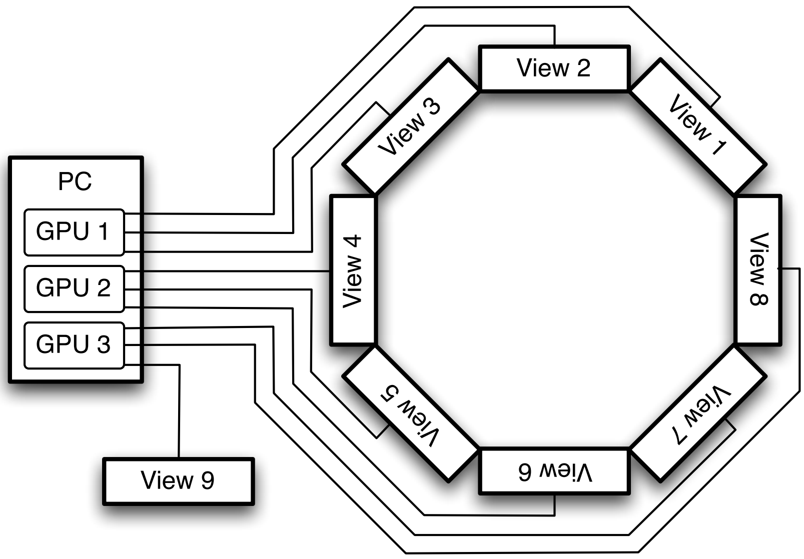

Figure 3.5: Schematic view of the rendering setup. One PC, equipped with three GPUs, drives the eight VR displays and an operator display.

In terms of graphics hardware, ATI consumer-level graphics cards (*ATI HD Radeon 5850*) are employed. In comparison, a NVIDIA solution turned out to be either much more expensive (professional QuadroPlex systems or Quadro cards) or to be considerably slower, because the consumer-level cards do not allow for efficient multi-view multi-GPU rendering. To exploit this efficiency and visualize complex scenes in real time, a custom-tailored rendering architecture had to be developed. Due to driver limitations of the employed bio-sensor (Section 3.3.3) the use of Microsoft Windows 7 as the operating system is mandatory, which requires certain low-level optimizations to fully exploit the parallel performance of the multi-GPU system, as described in detail later in Chapter 4. It enables real-time rendering (≈ 70 fps) of the detailed supermarket model on the eight OctaVis displays.

Besides the visual sense, the auditory stimulus is another important factor to trigger presence at the sensory scale. Therefore, four loudspeakers are installed above the display arrangement to increase sensory richness and fidelity.

3.3.2 VR Navigation and Interaction

Many intuitive navigation metaphors, like walking-in-place or ChairIO, are not suitable for physically handicapped patients, who might not be capable of sitting on a tilting chair or walking without additional help. Hence, the navigation and interaction techniques for the OctaVis has to be custom-tailored towards patients sitting on a robust chair placed in the center of the display arrangement.

For the navigation an interaction metaphor similar to an electronic wheelchair is used: The movement direction is intuitively controlled by rotating the chair into the desired walking direction. Movement speed (forward/backward) is controlled by a joystick in the armrest. When standing in front of a shelf, left/right movement of the joystick translates to sidewards movement along that shelf.

An advantage of this system is that rotating oneself on the chair—instead of rotating the virtual world around oneself using a joystick—matches the physical motion to the virtual action. This correctly stimulates the proprioceptive and vestibular queues, which has been shown to improve immersion and to better support learning of spatial configurations [KLB+98]. For the translatory movement, no physical stimulation is triggered. However, in the clinical context with partially handicapped people sitting on a chair, only the rotation can be handled in a physically consistent manner. To compensate for unilateral leg impairment, an optional footrest can be attached to the chair.

The chair orientation is measured using a rotary encoder (*Heidenhain ERN120*) placed around the shaft of the chair and connected to the PC through an Arduino board. The encoder has an accuracy of 0.35° at a sampling frequency of 300 kHz. To allow for unlimited chair rotation without any cable entanglements, a collector ring (*A-Drive Technology SRH50120*) is integrated into the chair such that the connected cables remain hidden from the user under a metal plate (Figure 3.2).

The throttle joystick is a *Metallux MJ-3K MTP* mounted in the armrests. It is connected via a USB hub attached to the chair, which itself is connected to the PC through the collector ring (see Figure 3.4). The joystick is an

Figure 3.6: Photograph of a user navigating in the virtual supermarket and buying items via a touch interface.

analog device, allowing the user to continuously control the movement speed up to a maximum speed, which has been empirically determined to minimize cyber-sickness.

Regarding interaction in the VR, objects are selected/bought by touching them on the screens, i.e., by natural and intuitive arm movements (Figure 3.6). Accidental selection of objects is avoided by a distance constraint, ensuring objects to be within arm-reach of the user in order to be bought. Combining the touch interface with the rotating chair metaphor an intuitive whole-body involvement is achieved for interacting with the virtual environment.

3.3.3 Clinical Requirements

Further requirements that have to be fulfilled for the use of VR systems in medical experiments are the recording of patients' physiological reaction and the supervision of the experiment.

In order to measure the heart rate and stress level (skin conductance) of the patient, two bio-sensors (*ProComp Infiniti*) are incorporated into the chair's

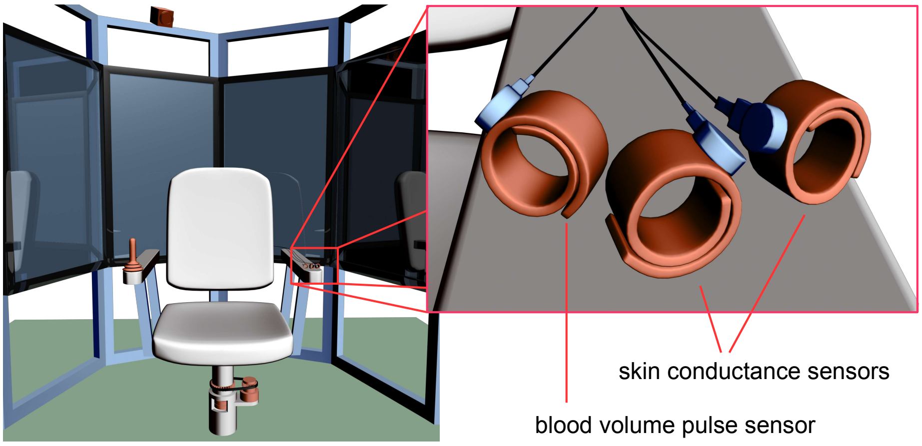

Figure 3.7: Illustration of the OctaVis system with a close-up of the skin conductance and blood volume pulse sensors used for patient monitoring.

armrest (Figure 3.7) and connected to the PC via the USB hub attached to the chair. The experiment is operated and controlled by the examiner using an operator display (view 9 in Figure 3.5). Two surveillance video cameras, mounted at the top of the OctaVis system, give the operator detailed information about the patient's action in the OctaVis system. Finally a galvanic separation (*Noratel IMEDe 2000*, Figure 3.4) is incorporated in order to secure patients from potential electric shocks. Apart from the medical use, the bio-sensors and the surveillance video cameras can be used to measure presence as proposed in Section 2.6.2

After fulfilling these special medical requirements, the system has been successfully CE-certified as a Class 1 medical device in Germany, which allows conducting clinical studies with patients.

3.4 Evaluation

The OctaVis system was developed in a highly interdisciplinary effort by a team of computer scientists, technicians, psychologists, and medical researchers. From the criteria listed in the beginning of this chapter, it meets the space requirements ($< 150cm$ in diameter, see Figure 3.1). Thanks to the single PC architecture, the system is easy to operate and maintain even without a technical background, and it was well accepted by the hospital staff. Since only standard consumer-level hardware components are used, the system is affordable (< 20k Euro material costs). Being CE-certified, the OctaVis system was in use in four different hospitals: the psychiatry (Prof. Dr. Martin Driessen), neurology (Prof. Dr. Wolf-Rüdiger Schäbitz), and epilepsy (Prof. Dr. med. Christian G. Bien) clinics in Bethel and the Marcus-Klinik (Dr. Thomas Brand) for neuro-rehabilitation in Bad Driburg. In all hospitals, VR-training was performed with admitted patients.

During the next two sections, we look into an initial pilot study reflecting the exploration of the hardware setup and the supermarket model in a general *search task*, and second, we examine the sense of presence during the actual *grocery shopping* training in a different study. Both studies investigate a general sense of presence opposed to spatial presence which will be examined in Chapter 5.

Search Task

The psychologists of the CITmed team conducted an empirical within-subject study [PK10]. Comparing the platform to a usual single-screen setup with mouse-keyboard interaction, they measured the subject's sense of presence during several consecutive search tasks. To this end, following the most common measurement approach, the participants filled out the Witmer and Singer questionnaire [WS98] after the experiment. This questionnaire is well established in the presence research community and widely used across the field.

The exact procedure was as follows. After a short time of customization to the navigation concept the subject was asked to perform the following six actions:

1. "Go to the meat counter and count the sausages"
2. "Go to the cheese counter and count the cheese"
3. "Go to the cereal shelf"
4. "Position yourself in front of the egg shelf"
5. "Go to the whiskey shelf"

6. "Leave the supermarket"

After each subtask, the participant stopped and reported her or his success to the advisor before (s)he was allowed to proceed. These assignments had to be accomplished in both conditions. The order in which the two arrangements were tested by the same person was randomized.

27 subjects in the age of 18–35 were tested in total. The results show a significantly higher sense of presence in the OctaVis-condition ($z = -4.37$, $p < .0001$). The mean values for the presence score are $m = 130.52$ ($sd = 17.36$) in the OctaVis-condition and $m = 104.26$ ($sd = 15.2$) in the single-screen condition.

Due to the use of touch screen displays, seamless displays could not be employed. This results in clearly visible seams between the eight screens (Figure 3.6). Particularly asked about the effect of these seams in an additional custom questionnaire, 21 subjects stated them to be not disturbing at all, 3 were disturbed slightly, and 3 were confused by them.

Grocery Shopping

Moving on to the actually chosen scenario for rehabilitation, we now evaluate the level of immersion and the resulting sense of presence in a user study with 19 healthy but elderly participants (4 male, 15 female) within the age range of 32–94 ($m = 65.42$, $sd = 15.57$). This group acts as a control group for a study with 10 stroke patients (7 male, 3 female) being 34–79 years old ($m = 59.4$, $sd = 17.09$).

VR for therapy research has been criticized for its lack of standardization. Study results cannot be compared due to the diversity of used devices, simulated environments, and tasks. Our training paradigm, therefore, is based on the rationale of classic neuro-psychological tests of verbal learning and memory, such as the California Verbal Learning Test (CVLT) [NSTOW08] and the Verbal Learning and Memory Test (VLMT) [HLL01]: In the experiments participants are first introduced to the system and perform a simple training course in a virtual office in order to become familiar with the OctaVis system. In the following 8 trials they perform a grocery shopping task in the virtual supermarket, for which a list of 20 products has to

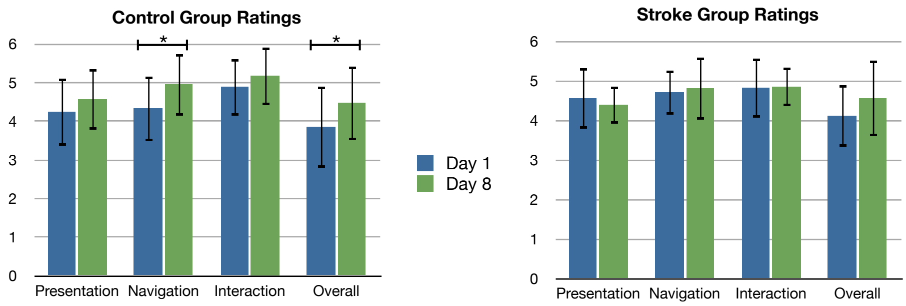

Figure 3.8: Questionnaire results for healthy elderly people (left) and stroke patients (right), showing scores above average on every scale. Stars mark significances. Error bars depict standard deviation.

be memorized and bought. In every trial, the same 20 items have to be bought, except for the seventh trial that introduces a different, distractive list. In the last trial, the initial product list has to be bought again, but this time without any new presentation of this list. The whole training takes at most 14 days. The maximum time interval between each of the first six trials is 48 h and 24 h between trials six to eight.

After finishing this training on the first and eighth day, a questionnaire inspired by Witmer and Singer [WS98] was filled out to investigate the quality of the VR in terms of *presentation*, *navigation*, and *interaction* as responsible media factors to trigger presence. Also, the *overall* impression of the VR system was asked about. In total the questionnaire consists of 31 items on a 0–6 scale (0: very bad, 6: very good). In order to analyze and test the data for significance, we calculate descriptive statistics and perform the non-parametric Wilcoxon test to compare results of day 1 to day 8. Figure 3.8 depicts the results for the control and the stroke group. The individual evaluations are discussed below.

- The *presentation* scale yields six questions examining different aspects of model quality, rendering quality, and the contribution of the display system to the sense of diving into a different place. Compared to the other scales, the ratings for the presentation component are lower but still significantly above average (score "3").

- Nine questions ask about realism and intuition of the control paradigm and the perceived movement to evaluate the *navigation* metaphor. Also, system response and eventual difficulties have to be rated. The results are high above average. For the control group, the navigation score even improves significantly ($z = -3.127$, $p = 0.002$) over time.

- Five questions make up the *interaction* scale, which measures the perceived realism, intuition, and the system's response to touch-actions. From both groups on both days, this metaphor gets scores around "5".
- The last scale represents the *overall* impression and coherence of the OctaVis experience. Here, six questions directly ask for the perceived presence, the ability to focus on the task, and how convincing the general feeling is. A significantly ($z = -2.665$, $p = 0.008$) higher score is found for the control group on day 8 compared to day 1.

In total, the results of all scales clearly demonstrate the OctaVis system to be a very immersive setup causing a real sense of presence in the virtual scene. The most noticeable facts of the questionnaire are (1) the increased rating over time by the control group and (2) the high initial scores for all scales by the stroke patients. Both statements prove that in the OctaVis users do not loose appreciation after the initial excitement phase, which VR systems often suffer from. Furthermore, all participants (being non-experts!) succeeded in the virtual shopping experiment, whereas in the CAVE-based study of Renner et al. [RDS+10] novice users failed to perform a very similar task.

In addition to the positive questionnaire evaluation, participants often asked for the location of the exit door in the display arrangement after finishing the experiment. This is another indicator that the participants lost their orientation in real reality and primarily located themselves in the virtual supermarket, again hinting at a high level of presence generated by the OctaVis system (see PERF in the two-step model of presence in Figure 2.2).

3.5 Summary & Outlook

In this chapter, we have seen the novel OctaVis VR-system, where users are sitting on a rotating chair in the center of eight displays arranged around them. The eight displays provide a 360° horizontal full-panorama view of the virtual world, through which the user can easily navigate via chair rotation and a throttle joystick. Object selection is naturally performed through an intuitive touch-screen interface.

Using just a single PC to drive the eight VR-displays, the OctaVis system is easy to maintain, cost efficient ($<$ 20k Euro hardware cost), and has a small spatial footprint ($<$ 150cm in diameter). The system was designed for the use in clinical studies, with patient safety and patient abilities in mind. It has been deployed to four hospitals. The first experimental studies confirm that the OctaVis system is easy to use even for elderly people after a stroke without any VR experience and that the combination of a task in a virtual supermarket environment and the chosen hardware components generates a high level of presence.

Before a more thorough investigation of the contribution of self-rotation and surround view to presence and task performance in Chapter 5, first, we will have a detailed look at the software system driving the OctaVis.

CHAPTER 4

Software Framework

Programming today is a race between software engineers striving to build bigger and better idiot-proof programs, and the Universe trying to produce bigger and better idiots. So far, the Universe is winning.

— Douglas Adams

Obviously, the requirements of *ease of use* and *maintenance* from Chapter 3 hold for the software framework as well. Being operated by non-technical staff, the handling of an experiment should be straightforward. Extending the framework for new types of experiments should be supported as well. Besides these general requirements, a sharp focus is demanded on the performance of the rendering component because shortcomings in visualization can easily break the sense of presence, especially in a surround view.

Following the decision to render eight views, plus one operator display, with a single PC equipped with three graphics units, this chapter emphasizes the rendering architecture, the parallelization efforts, and the low- and high-level performance optimizations that eventually enable the real-time rendering of complex VR environments on such an arrangement.

4.1 Related Work

Virtual reality software as a middle layer shields application developers from hardware and rendering details. At least it handles a variety of input devices (wands, gloves, steering wheels) and different output devices (HMD, CAVE, power-walls, projection systems). Often packages also provide scripting, (stereo) rendering, cluster support, editors, art pipeline, tracking, and configuration systems to ease application development and add flexibility. Virtools [Das16], Vizard [Wor16], and Instant Reality [Fra16] all represent such extensive frameworks but disqualify for this project due to their commercial license and price.

Since the visual stimulus is of great importance, VR software often relies on third-party scene graph libraries. Instant Reality, for example, is built on top of OpenSG, Vizard on top of OpenSceneGraph. Using such a scene graph library, however, is not possible for the OctaVis, since most existing libraries are neither designed nor optimized for single-PC multi-GPU rendering systems.

Favoring rendering quality, another approach is to extend existing game engines. Jacobson and Lewis [JL05], for instance, experimented with the Unreal [1] engine, Juarez et al. [JSB10] with the CRYENGINE[2], and MiddleVR [Mid16] is an add-on to the Unity[3] engine. Unfortunately, being designed and optimized for single-screen applications, these game engines all lack native multi-GPU support and the mentioned extensions target distributed rendering setups only. When setting up a local cluster on a single machine, these extensions slow down the performance of the underlying engine and increase latency. Also, Unity and MiddleVR are closed source projects not allowing for fine tuning of data storage and distribution.

[1] `www.unreal.com`
[2] `www.cryengine.com`
[3] `www.unity.com`

The most important question concerning rendering, therefore, is how to design the rendering architecture such that the parallel performance of a multi-GPU system is exploited in an optimal manner.

The major graphics vendors already provide solutions for combining several GPUs in order to increase rendering performance (NVIDIA's SLI, ATI's CrossFire). Note, however, that these techniques only support a single graphics output, and hence are not applicable in a multi-view setup. NVIDIA's QuadroPlex is a multi-GPU solution that combines up to four Quadro GPUs in an external case. Two QuadroPlex boxes can, therefore, be used to drive eight displays, but such a system comes at a price of more than $20,000. ATI's EyeFinity is a technology for driving several displays by one graphics board, and hence is rather a multi-view approach, but not a multi-GPU solution.

To distribute rendering to the different GPUs, initial experiments with higher-level APIs such as the distributed scene graph OpenSG and the multi-GPU-aware scenegraph OpenSceneGraph turned out not to be flexible enough to give (easy) control over the crucial details affecting multi-GPU performance (as discussed in Section 4.6).

A distribution of low-level OpenGL commands based on Chromium [HHN+02] was done by Rabe et al. [RFL07], who built a system similar to the OctaVis. However, their performance results were rather disappointing, mainly due to the overhead induced by the Chromium layer. An elegant abstraction layer for parallel rendering is provided by the Equalizer framework [EMP09], which allows for distributed OpenGL rendering using render clusters, multi-GPU setups, or any combination thereof.

With the design goal of keeping the rendering software as simple and efficient as possible, none of the mentioned distribution packages were employed. Rather, our custom-tailored rendering engine handles each view and each GPU individually and allows for a low-level OpenGL solution for the target system.

VR Juggler [BJH+01] is an open source VR package supporting standard VR tasks while allowing for custom graphics programming interfaces. In particular, it supports OpenGL, OpenGL Performer, OpenSG, and Open-

SceneGraph. VR Juggler itself allows multi-pipe setups but handles just the window management. Since the window management and the low-level rendering turned out to be strongly interconnected for multi-pipe optimizations, this package also was dismissed for this project.

Instead, hoping to beat the universe at the race (compare quote in the beginning of the chapter), a slim, multi-pipe optimized VR architecture is developed, which offers ease of implementation and simple extensibility.

4.2 Rendering Architecture

Since the rendering architecture is not built on top of a high-level framework, manual care has to be taken of the distribution of render commands to the different GPUs (Section 4.2.1) and the data management (Section 4.2.2). This low-level control enables the employment of crucial performance optimization techniques (Sections 4.3 and 4.4).

4.2.1 Distributing OpenGL Commands

The multi-GPU multi-view architecture consists of three GPUs with three outputs each, which drive the eight VR displays of the OctaVis system. We, therefore, have to be able to address OpenGL commands to a particular display attached to a particular GPU.

At application start-up, a single full-screen window for each of the eight views is generated, with each window having its own OpenGL context. For Unix operating systems, distributing render commands is easy, since an individual X-server can be explicitly assigned to each view. In the CITmed project, however, external constraints require MS Windows as operating system, which does not provide this explicit control.

The Windows Display Driver Model 1.1 used in Windows 7 provides improved support for multi-GPU applications, but it turned out that the actual performance strongly depends on the GPU driver. Initial experiments revealed that the NVIDIA driver dispatches *all* OpenGL render commands to *all* available GPUs. This obviously prevents efficient paralleliza-

GPUs	Views/GPU	Views	FPS NVIDIA	FPS ATI
1	1	1	58.3	273.6
2	1	2	27.5	253.8
3	1	3	19.3	258.8
1	2	2	30	129.7
2	2	4	14.6	121.5
3	2	6	9.8	121.5
1	3	3	–	82.4
2	3	6	–	79.4
3	3	9	–	78.1

Table 4.1: Comparing frame rates for a 870k triangle model using NVIDIA GeForce 9800 GX2 and ATI Radeon HD 5770 cards, respectively, in several multi-GPU and multi-view setups.

tion. In order to address a specific NVIDIA GPU the OpenGL extension `WGL_NV_gpu_affinity` has to be used, but this extension is only available for (expensive) Quadro-GPUs.

In contrast, the ATI driver dispatches render commands to just the one GPU responsible for the current window, which is the GPU attached to the display the window was created on. It therefore turned out to be crucial to create the eight windows at the correct initial position on the respective view. Creating all windows on the first view and moving them to the proper position afterward does not work. When taking this subtle information into account, the ATI driver allows for efficient parallelization of several GPUs.

Table 4.1 compares the performance scalability of NVIDIA GeForce 9800 GX2 GPUs (two outputs each) and ATI Radeon HD 5770 GPUs (three outputs each) with varying numbers of GPUs and varying numbers of views per GPU. All experiments were performed on a standard PC with an Intel Core i7 930 CPU running Windows 7.

It is clearly visible that the NVIDIA system scales inversely proportional to the number of views: No matter whether two views are driven by one GPU or by two GPUs, the performance drops by a factor of about two. This is an immediate consequence of the NVIDIA driver sending OpenGL commands to all available GPUs, as also described in [EMP09].

The ATI system, instead, scales almost perfectly. Rendering one view per GPU gives the same performance on one, two, or three GPUs. When keeping the number of GPUs fixed, the performance decreases almost linearly with the number of attached views per GPU. Hence, the ATI system can fully exploit the parallelization of multiple GPUs. Because of these reasons, ATI Radeon HD 5770 GPUs were chosen for the OctaVis rendering system, which subsequently, to further boost performance, were replaced by ATI Radeon HD 5850 cards.

4.2.2 Data Management

Thanks to the simple single-PC architecture, distribution of the scene data and the render states over the network to individual render clients are avoided. However, in order to render the scene in a multi-view application each OpenGL context (i.e., each view) needs access to the scene data, such as e.g., geometry, textures, and shaders. In order to minimize memory consumption, scene data is not duplicated for each view. Instead, one copy only is stored per GPU, which is then shared by all views attached to this GPU.

This behavior can easily be implemented by shared OpenGL contexts. The OctaVis rendering engine incorporates this functionality into a very simplistic scene graph with standard nodes for shaders, shader uniforms, and triangle meshes. Shader nodes, for instance, then store a reference to a shader object only. The shader object in turn stores the shader-data on each available GPU (but *not* for each individual view). See Figure 4.1 for an illustration. In such a case, where each GPU drives up to three views, the memory consumption is reduced by a factor of three.

4.3 Low-Level Optimizations

In order to optimize rendering performance, one has to identify and eliminate the typical performance bottlenecks: CPU load, data transfer from CPU to GPU, and GPU load. In a multi-view multi-GPU environment, even more attention has to be paid to these issues.

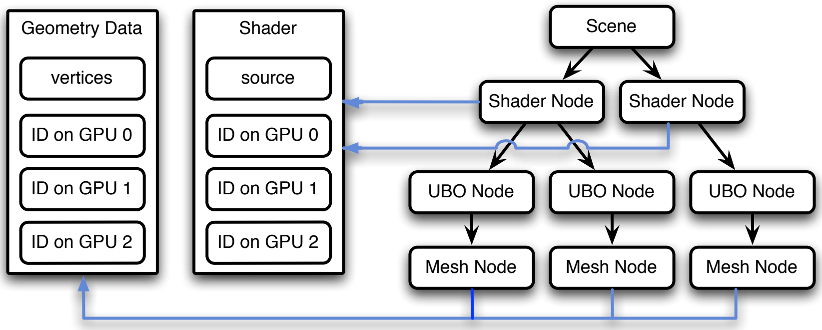

Figure 4.1: Scenegraph nodes share data through shared OpenGL contexts to reduce memory consumption.

Rendering geometry in *immediate mode* quickly makes the application CPU-bound due to the massive amount of `glVertex()` function calls. Therefore, vertex positions and triangle indices are stored in *vertex arrays* (VA), which allows to render meshes with a single function call and thereby eliminates the CPU bottleneck. However, in a multi-view setup the data has to be transferred from main memory to the GPU eight times (for each view) in each frame, such that data transfer immediately becomes the bottleneck.

Data transfer can be eliminated by storing the vertex arrays in *vertex buffer objects* (VBO) on the GPU. This turned out to be absolutely crucial in a multi-view setting. The respective data storage follows the same shared context paradigm as described in the previous section. With VBOs, the bottleneck is no longer data transfer, but the per-vertex computations of the GPU. Because of the general applicability of these optimizations, since OpenGl 3.x immediate mode and vertex arrays are declared deprecated.

Furthermore, the GPU load can be reduced by caching computations performed for individual vertices. If a triangle is rendered and one of its vertices has been processed before and is still in the cache, these computations can be re-used. To maximize cache-hits the individual vertices and triangles of the mesh are re-ordered using the method described in [YLPM05], which (depending on the model) yields a significant performance gain.

Finally, in order to optimally exploit all available GPUs, the render traversal is parallelized: Each GPU is served by a dedicated render thread that pro-

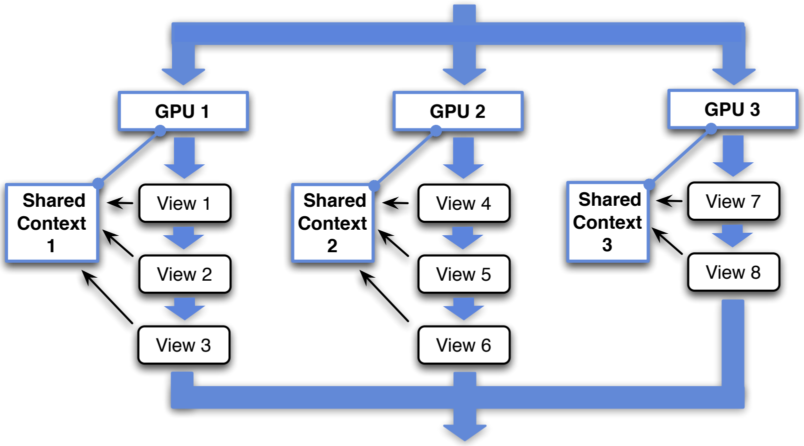

Figure 4.2: All views attached to the same GPU share a render context and are drawn consecutively, while the rendering is parallelized over the available GPUs.

cesses all views attached to that GPU in a serial manner. Parallelizing over the (two or three) views on the same GPU did not increase performance. Figure 4.2 gives an overview of the rendering process.

Table 4.2 compares the different optimization techniques using three ATI Radeon HD 5770 GPUs. This experiment uses an early version of our VR supermarket, consisting of about 1.7M triangles. In this scene, all instances of a particular object (typically several copies of one shopping item in a shelf) are merged into a single triangle mesh. This results in 75 meshes in total, which are stored either in vertex arrays or vertex buffer objects, respectively.

When the geometry is stored in vertex arrays, but not in VBOs, the geometry data is transferred to the GPU for each active view. Consequently, the performance drops with each additional view, even when they are attached to different GPUs. Storing the geometry in VBOs eliminates the transfer costs, which then yields the convincing scaling behavior discussed above. After re-ordering vertices and triangles for each individual VBO, the rendering architecture is able to visualize the 1.7M triangle scene on nine displays at a rate of more than 100 frames per second.

GPUs	Views/GPU	Views	FPS VA	FPS VBO	FPS RO
1	1	1	38.8	214.3	344.2
2	1	2	22.3	214.6	344.8
3	1	3	14.2	214.7	344.7
1	2	2	19.2	107.7	171.3
2	2	4	10.2	107.5	171.2
3	2	6	7.0	107.5	171.1
1	3	3	12.7	71.6	113.6
2	3	6	7.2	71.4	113.5
3	3	9	4.7	71.3	113.3

Table 4.2: Frame rates for different optimizations (vertex arrays, vertex buffer objects, cache-friendly reordering), using ATI Radeon HD 5770 GPUs.

4.4 High-Level Optimizations

The performance optimizations described in the previous section were sufficient for the virtual supermarket consisting of 1.7M triangles. However, in order to further increase the realism and the immersion in the CITmed project, a more detailed VR supermarket was created. The new model was constructed from 680 individual objects (furniture, shopping items), from which 94k instances were created, eventually resulting in 4M triangles (see Figure 4.3).

Furthermore, it should be possible to interact with each individual object instance, for example in order to buy individual products by simply picking them. Note that this is not possible when storing all instances of a particular product within a single VBO, since then individual instances cannot be made invisible after being bought. Representing each object instance by a separate scene graph node, i.e., splitting the scene from 680 objects into 94k objects, allows for simple interaction with object instances, but reduces the performance to about 1 fps due to CPU load.

At this point, three high-level optimizations are applied to achieve the goals: Geometry instancing allows naturally to handle scenes with many duplicated objects (Section 4.4.1), view-frustum culling (Section 4.4.2) and multi-GPU load balancing (Section 4.4.3) further increase performance. All

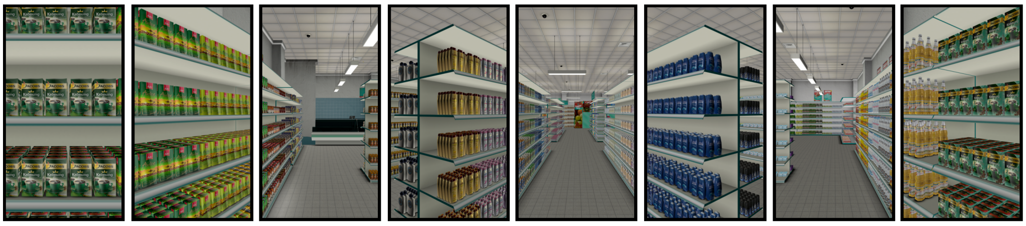

Figure 4.3: Eight views of the VR supermarket (4M triangles), corresponding to an "unfolded octagon". This model was used for the optimizations described in Sections 4.4.2 and 4.4.3.

experiments shown in this section have been performed using three ATI Radeon HD 5850 GPUs.

4.4.1 Geometry Instancing

Since version 3.3, OpenGL provides hardware-accelerated geometry instancing for VBOs, which allows rendering multiple instances of a single object with a single function call. This technique needs two data streams, one supplying the object data (geometry, textures, shaders) and the other providing the positions/orientations of the individual instances. For each instance, a shader transforms the geometry according to the given instance transformation. This way, geometry and texture data have to be stored once per object only, instead of once per instance.

For maximum performance, both the object data and the instance data have to be stored on the GPU, since otherwise the data transfer immediately becomes the bottleneck. The engine, therefore, stores all instance transformations within a single *Uniform Buffer Object* (UBO) on the GPU. Another UBO holds binary on/off information for each object instance, which is used for turning off objects once they have been bought.

Geometry instancing allows for conceptually clean and efficient handling of object instances. In terms of performance, however, geometry instancing does not make a difference for the high-resolution supermarket of 4M triangles. Storing all object instances in 680 VBOs or employing geometry instancing both yields 25 fps. The main advantages of geometry instancing are therefore (1) the ability to interact with individual object instances (pick or buy each single product) and (2) the largely reduced consumption of GPU memory. Geometry instancing reduces GPU storage from 243 MB per GPU for single VBOs down to 8 MB per GPU (2 MB mesh data, 6 MB instance data). This significant reduction is mainly due to the highly repetitive supermarket scene with a ratio of 680 objects to 94k object instances.

4.4.2 View-Frustum Culling

To achieve real-time performance even for the more complex supermarket model of 4M triangles, the framework exploits the fact that all object instances are arranged spatially close to each other, namely in the same product shelf. This allows to compute a bounding box around each group of instances and to perform view-frustum culling during scene graph traversal for each of the eight views. Comparing CPU-based view-frustum culling to GPU-based occlusion culling for our ATI GPUs, the former turned out to be slightly more efficient than the latter.

For the 4M-triangle supermarket, the CPU-based view-frustum culling increases performance from 68 fps to 205 fps for single-view single-GPU rendering, and from 25 fps to 56 fps for eight views rendered using three GPUs.

4.4.3 Load-Balancing

View-frustum culling can lead to a considerable performance gain – depending on how much of the scene is outside the frustum and therefore is culled. Since this proportion varies with camera position and viewing direction, frustum culling inevitably leads to an unbalanced load distribution for the eight individual views of our multi-view setup. For instance, at a location close to a wall in our supermarket some views mainly have to render the wall, while others have to process almost the whole scene. Since due to view synchronization, the slowest GPU determines the frame rate of the overall rendering, some kind of load balancing should be employed.

Since each GPU has to render two or three views, one could try to counterbalance the varying load per view on the level of GPUs. The most straight-

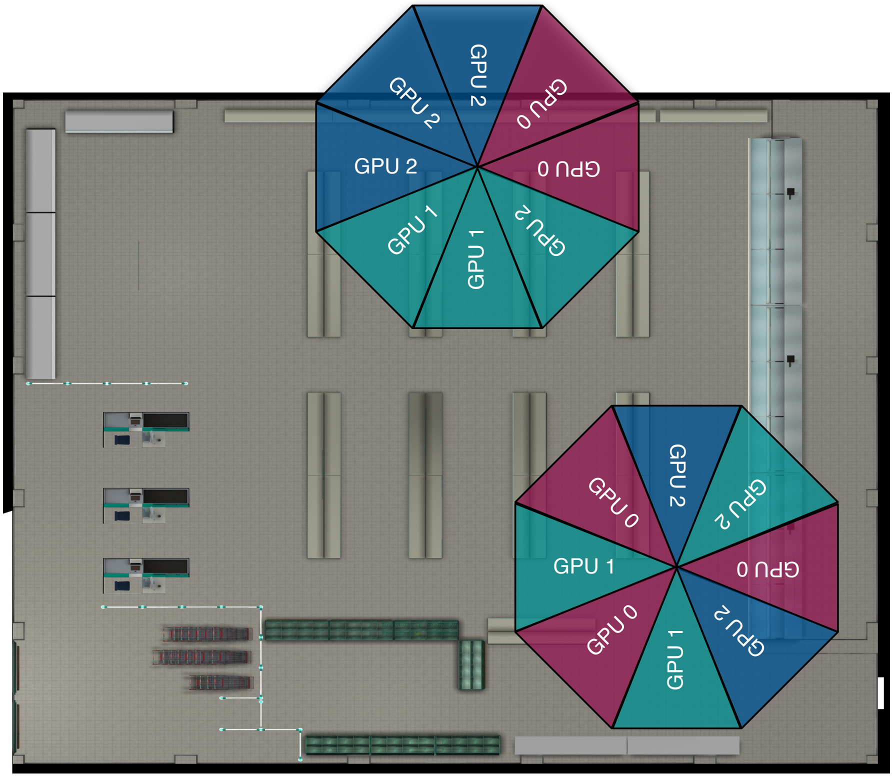

Figure 4.4: Top-view of the supermarket scene, showing *neighboring* view configuration (top) and *interleaved* view configuration (bottom).

forward assignment of views to GPUs is the one where each GPU is responsible for a set of *neighboring* views. Assigning the views in an *interleaved* manner (Figure 4.4), however, results (in general) in a much better load balancing.

Table 4.3 compares the rendering performance for the two view configurations shown in Figure 4.4, with the camera being randomly positioned in the scene. The timings are taken separately for each of the three GPUs, where the slowest GPU (in this case GPU 1) determines the overall frame rate. This experiment clearly shows the interleaved view configuration to yield better load balancing, resulting in a performance gain from 56 fps to 74 fps. Note that this result is almost independent of the location in the supermarket, leading to a general performance improvement.

View Configuration	GPU 0	GPU 1	GPU 2
Neighboring	205	56	109
Interleaved	91	74	154

Table 4.3: Frame rates of each individual GPU for different GPU/view configurations.

4.5 Global Architecture

After understanding the rendering part of the software, we now shortly zoom out to a more global view on the framework. Flexibility was a major goal for the project, for instance in order to be able to connect different input devices or to perform different types of experiments. Spatial thinking experiments, in particular, range from shopping tasks in a detailed virtual supermarket to orientation tasks in abstract virtual environments (e.g., Morris water navigation task). Because of that, each experiment is encapsulated in a separate class and derived from the abstract class *Experiment*, thereby providing a persistent interface for the main application. A similar method has been applied to all attached devices, which derive from a common parent class *Module*. Figure 4.5 depicts this software architecture principle.

This class hierarchy turned out to be flexible and beneficial during the development stage. It provides a simple way for the integration of different devices for navigation (game pad, joystick, mouse, and keyboard) or different variations of a certain device. Even the VR representation and visualization is implemented as such a module. Optional features such as different experiments or different input devices or sensors can be controlled either via simple configuration files or via a graphical user interface.

4.6 Discussion

Creating own software despite the existence of similar products comes with the general advantage of total access to every single system component. Creating an own, extensible VR software middle layer, despite the existence of VR Juggler, Instant Reality, etc., proved to be especially conve-

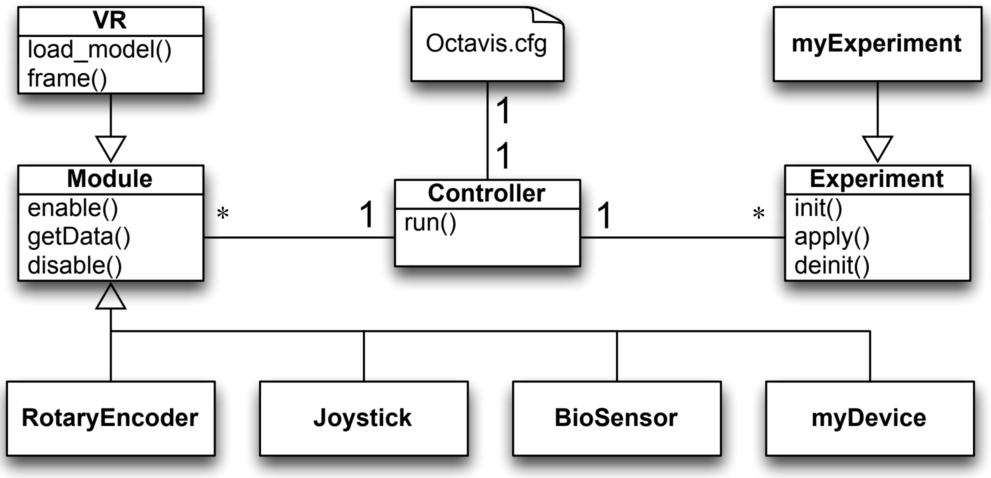

Figure 4.5: UML diagram of the main software architecture.

nient. Midway through the project, when ideas about new types of experiments emerged from the department of psychology, it was easily possible to thoroughly account for new in-game measurements, rule systems, additional input device types, even bio-sensors, eye-trackers, and cameras for head tracking. The mentioned third-party software does not allow for such simple and seamless inclusion of devices with custom device drivers (e.g. rotary encoder and bio-sensor).

Successful extensions to this general architecture were implemented for the famous Morris water maze task, a grocery shopping task in front of a single desktop monitor with a mere joystick control, a combined mouse-keyboard control, and a gamepad control. Also, the different technical conditions required for the experiment in the next chapter were easily added due to this flexible software setup.

In terms of the rendering component, the applied low-level and high-level optimizations result in a multi-view rendering system that is capable of visualizing a complex 4M triangle scene on eight displays at a rate of 74 fps, which corresponds to 2.4G triangles/second. At the same time, the system allows for fast interactions on an object instance level. Figure 4.6 shows some impressions of the current model.

Figure 4.6: Some example views of the high-detail supermarket, consisting of about 94k object instances and 4M triangles. Each row shows several views for different viewing positions in the scene.

As mentioned in Section 4.1, experimentation with the two popular scene graph libraries OpenSceneGraph and OpenSG did not achieve a comparable performance, mainly due to the following reasons:

OpenSG is specialized to distributed rendering, therefore, for a multi-view implementation like the OctaVis, eight local render server processes have to be used to drive the eight displays. Since each render server requires a full scene copy, this consumes significantly more memory than sharing contexts as proposed in Figure 4.1, and therefore does not allow for highly complex scenes. Moreover, due to sub-optimal window creation and setup (see Section 4.2.1) OpenGL commands are dispatched to all GPUs, which slows down the rendering to about 1 fps.

OpenSceneGraph correctly creates and initializes windows, supports for shared OpenGL contexts, and allows for precise control of VBOs. However, its rendering performance is still only about 70% of the one proposed in this chapter. This is probably due to the higher overall complexity of the scene graph system.

4.7 Summary & Outlook

In this chapter, we have seen the inner workings of the rendering system driving the OctaVis platform. It clearly demonstrates the potential of a multi-view rendering system based on a single PC equipped with multiple GPUs. However, the results also indicate that the multi-GPU performance can crucially depend on a few seemingly minor implementation details, on optimization techniques, and on GPU drivers and operating systems. The developed rendering architecture scales almost perfectly thanks to such few but carefully done performance adjustments.

In terms of the global software framework, it is easy to extend for future experiments by both, new hardware and new tasks in VR. In all four clinics the OctaVis has been deployed to, the experiments are successfully handled by non-technical staff after a simple short introduction to the software system.

After this detailed examination of the software framework establishing a realistic rendering of the therapeutic, virtual environment, we now, in similar depth, turn to the patient's perception of the virtual space. The following chapter investigates the contribution of self-rotation and a surrounding view to the mental creation of the Spatial Situation Model.

CHAPTER 5

Spatial Evaluation

A map says to you.
Read me carefully, follow me closely, doubt me not...
I am the earth in the palm of your hand.

— Beryl Markham

"What did you do, memorize a map of the city for fun?" says Christina.
"Yes," says Will, looking puzzled. "Didn't you?"

— Veronica Roth (Divergent)

In Section 2.3 we have already seen the importance of everyday life tasks for virtual therapy. Managing a "trivial" task (e.g. grocery shopping) strongly contributes to the feeling of independence and thus quality of life. Many of such everyday-life activities involve or even heavily rely on spatial understanding. In fact, visuospatial abilities are a key attribute for managing everyday life. No matter whether we travel by plane to a conference, by train to visit a relative, or by car to see the ocean, we need to locate ourselves in space to decide which direction to go to arrive at a fixed destination. Also, to find our way around in a city or in a building (supermarket, retirement home, school), we rely on our spatial understanding. We do so even when other people describe specific locations to us, when we interpret pictographic signs, or when we read any kind of spatial map.

Since spatial thinking is one of the seven primary mental abilities in the theory of intelligence, and also since sports, design, mechanical construction, and even perception of music strongly depend on a functioning spatial cognition, we are very limited when the corresponding brain areas get injured.

Luckily, the human property of visuospatial thinking is highly accustomed to fast learning and relearning. Every time we visit a new place, we learn its spatial layout. Even someone who is used to an orthogonal Manhattan layout of space, after just a few days in the wild forest, begins to distinguish the former "equal" trees, the shapes of branches, the subtle color changes in the leafs, and the minimal irregularities in the slopes to find some landmarks or bearings to build a mental representation of his or her surroundings. Therefore, even after a heavy injury, for instance through stroke, neurodegenerative disease, or an accident, visuospatial thinking often can be fully recovered through specific training.

This chapter is concerned with the effect of the OctaVis on the learning of space. Namely, we analyze which effect on spatial presence and spatial orientation skill (1) the eight-screen surround-view and (2) the user's physical self-rotation have. We will do so using the Measurement Effects Conditions Spatial Presence Questionnaire (MEC SPQ) [VWGO04], which measures not only the feeling of presence in general, but focuses on *spatial presence*, as described in Section 2.5, and a custom pointing task.

5.1 Related Work

It has been shown that spatial knowledge does not transfer easily from virtual reality to the real world [DM95, Pso95, GM98]. We want to test the effect of the OctaVis on the ability to orient oneself and to navigate through the virtual environment, the so-called *way-finding*. In the initial definition of way-finding by Kevin Lynch [Lyn60] the four components of way-finding are: orientation, route decisions, mental mapping, and closure. An overview can be found in [RE98]. Mental maps, also called cognitive maps, are the concepts agents create in their minds to enable them to plan their activities. They are an abstraction of the environment based on the sensory input of the agents. Mental maps are (sometimes) helpful in way-finding, but way-finding does not necessarily have to rely on mental maps.

In their work on child development, Piaget and Inhelder [PI67] differentiate between perceptual space and conceptual space. The perceptual space is created and inherently tied to sensori-motor activities that lead to corresponding perceptions. Its development precedes that of the conceptual space in early childhood. The orientation aspect of way-finding can be linked to the concept of perceptual space, whereas mental mapping is clearly linked to conceptual space. So there are two different cognitive mechanisms that are at work in way-finding. At least the first in development, the perceptual space, is strongly grounded in sensori-motor activities and thus, some effort has to be taken to provide sufficient sensori-motor cues. A VR interface for navigation thus has to provide sufficient cues for both levels to fully support natural way-finding, e.g., by providing appropriate visual cues and a suitable locomotion technique [DP02]. As certainly already noticed, the OctaVis addresses both levels: a 360° surround-view to maximize visual cues and a chair allowing physical self-rotation for the locomotion part.

5.1.1 Related Work on Locomotion

There are effects of the choice of the locomotion technique on way-finding and in particular on orientation. It has been shown that locomotion interfaces that provide proprioceptive and vestibular feedback have a positive effect on the user's navigation capabilities [DS93, CGBL98, RL09]. According to expectation, interfaces that come close to the real walking experience provide an excellent feedback and support navigation performance of humans.

Interfaces that come close to natural walking are redirected walking [RKW01], motion compression [NHS04], seven-league-boots [IRA07], and scaled-translational-gain [WNM+06]. However, all of these interfaces have been designed to allow for navigation in larger-than-real spaces (i.e., the size of the virtual world is larger than the real estate) and thus feature some kind of compression or scaling [WNR+07]. It is unclear how this affects navigation performance. In fact, providing the correct proprioceptive cues for real walking but showing a boosted walking performance in VR could make way-finding even more problematic. Also, some of these techniques require a method for re-orienting the user in the virtual world whenever he or she approaches a barrier in the real world, which might reduce presence [RKW01, PFW09]. On the other side, interfaces that only allow for a partial approximation of natural walking, such as walking-in-place or treadmills, yield only suboptimal performance [Hol02, DCC97].

Currently, there is a disagreement about whether real walking simulation is the key factor [RL09] or if it is sufficient to make the users do physical rotations to alter their orientation [RSPA+06]. This is where the OctaVis-approach is positioned, where the user is required to physically change his or her orientation, although in a seated position.

There is not much work on navigation through virtual reality in a seated position. A recent exception is the work on redirected driving by Bruder et al. [BIPS12]. They use electric wheelchairs that drive in the real world to provide sufficient proprioceptive feedback to the driving user. However, they are using HMDs to provide a full visual immersion into virtual reality. They find that people are a little less sensitive to a redirection of their orientation when driving the wheelchair than when walking. This could imply that users are less sensitive to rotations of the wheelchair. In the OctaVis, however, the users have full control of their rotation, while in the study of Bruder and colleagues a joystick was used to guide the motor-driven wheelchair. A promising insight of their study is that they found no significant difference in presence, thus driving through the virtual world while being seated was sufficient to establish a similar presence as natural locomotion.

5.1.2 Related Work on Field of View

In early work on the effect of the field of view on presence, Hendrix and Barfield [HB95] already showed that increasing the field of view increases presence. However, their restricted technical setup only allowed a comparison in a range of 10° to 90°. Interestingly, they did not find a significant improvement between 50° and 90°.

A decrease of the field of view was observed to come along with a decrease in cognitive map building performance [AM90]. McCreary and Williges [MW98] reported that the field of view did not have an effect on landmark knowledge, but had a significant effect in a pointing task, where participants had to point to occluded objects. They also found a correlation between computer experience and performance in the pointing task.

In a desktop-based setup with up to 180° field of view, Seay et al. [SKHR01] showed that an increase of the field of view to 180° increases the feel of presence, but also increases nausea, especially for participants not actively navigating in the virtual world. Lin et al. [LDP+02] tested displays with fields of view of 60°, 100°, 140° and 180°. They report an increase of presence with an increase of field of view, albeit the increase seemed to approximate a plateau between 140° and 180°. In their memory test, the results also correlated positively with the perceived presence.

However, there is a general effect of underestimating traveled distance in virtual reality [WK98], which might affect way-finding. It is unclear whether orientation is affected in the same way—the fact that techniques like redirected walking (see above) can make use of a human tolerance for a tinkered-with rotation might hint that there is. Knapp and Loomis [KL04], however, could not find a significant effect of a limited field of view on distance estimation.

Under the assumption that an increased field of view does increase presence, which in turn increases memory capabilities regarding scenes in the virtual world and thus supports cognitive map building, the OctaVis was designed to support the maximum field of view independent of the viewing direction of the user. This does not hold for most desktop- or projection-based studies mentioned above and the effects remain to be evaluated.

5.2 Method

To investigate spatial presence and orientation in the OctaVis a two-step experiment was designed. For the first step, the real field of view was altered while during the second step the rotation aspect of the navigation metaphor was changed. For these two steps the following three experiment conditions were compared:

Frontal View (FV): In the first condition (Figure 5.1, top) three displays rendered the virtual maze. This resulted in a wide, but still limited, frontal view of 135°. To navigate the world the user operated a joystick with two axes: one to move forward/backward and the other to rotate the virtual scene in order to change the traveling direction. This rotating world principle is a common metaphor used in most first person desktop games.

Surround View (SV): For the second condition (Figure 5.1, middle) the rotating world metaphor was kept, but the field of view was extended to a 360° horizontal surround view employing all eight displays for rendering the maze. To prevent confusion about the front-direction one monitor was explicitly marked as front-monitor.

Rotating User (RU): The third condition (Figure 5.1, bottom) also used all eight monitors for surround-view rendering, but the control metaphor was different. Instead of a rotating virtual world with the user sitting stable, now the user physically rotated herself to control the traveling direction. This way, "front" was where the user turned the chair to, which was measured by the rotary encoder in the office chair (Figure 3.2). The joystick had just one axis enabled used to move forward/backward.

To enable a comparison of the three conditions in a within-subject manner, three mazes have been designed. According to the findings in the field of space syntax and research about dementia-friendly architecture [MS09, DE00, BWÖ04], orientation in a building benefits from straight circular corridor systems, right-angled turns, architectural symmetry, direct sight, large rooms, and different textures and geometries per room.

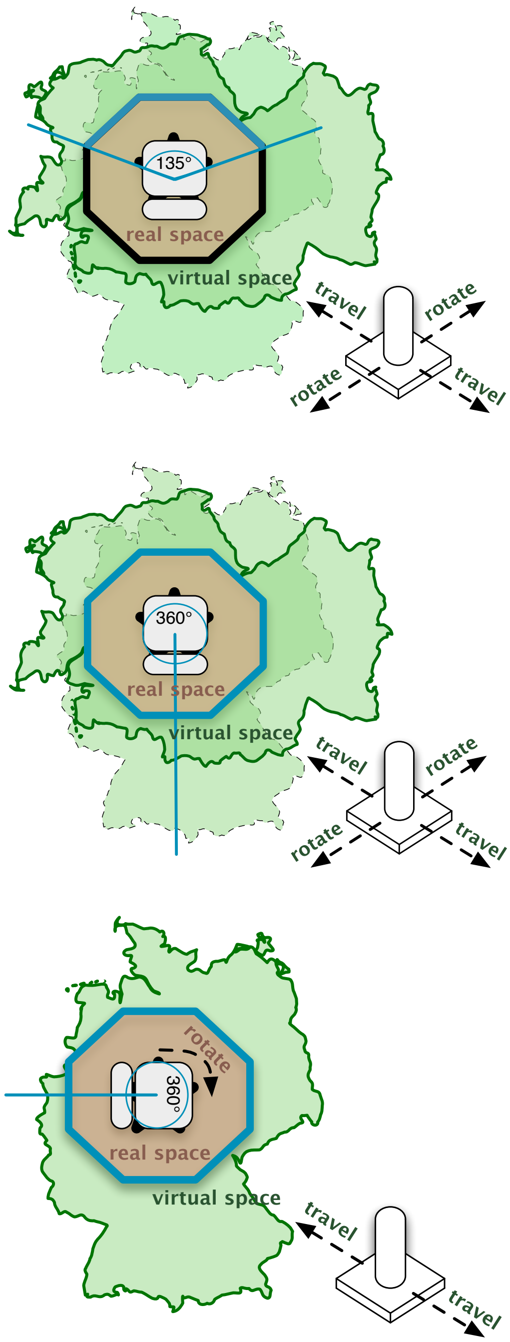

Figure 5.1: The three conditions designed for the study. Top: The Frontal-View condition (FV) has a 135° field of view (3 displays) and rotates/travels the virtual space through a joystick. Center: The Surround-View condition (SV) employs a 360° field of view and again navigates the VR by joystick. Bottom: The Rotating-User condition (RU) has a 360° field of view, but the user rotates in real space to control walking direction in the virtual space; the joystick is used to navigate forward/backward only.

Since this study was aimed at challenging orientation, the mazes had to be hard with respect to the above criteria. To allow for adjustable complexity, a modular system of 30 equally textured, small, octagonal cells with different entrance-exit-combinations was developed. These cells can be configured on a rectangular grid to create a maze, a corridor arrangement, or an architectural room-assembly. A pre-study with different cell combinations ensured both a sufficient and comparable complexity between mazes. Based on the experience and results from this pre-study, the three corridor-arrangements shown in Figure 5.2 were chosen to evaluate the OctaVis. Each arrangement consists of three asymmetrically coiled corridors, some of which feature unusual turns of 45°.

Semantically, each maze acted as a museum hosting four images of the same topic. The first one showed how different creatures eat (humans, bears, sharks, eagles), the second one how they move (humans, dogs, cheetahs, bears), and the last one was about styles of government (monarchy, Obama, Putin, Merkel). The images were displayed in the central room joining the three wings and also in the last room at the end of each corridor (Figure 5.2). An impression of an inside view is given by Figure 5.3.

In the actual experiment, every participant had to perform three consecutive sub-experiments, each in a different maze. All participants started in Maze 1 and finished with Maze 3, but the three conditions (FV, SV, RU) to be used for the mazes were randomized, such that 16 participants performed the experiment in the order FV-SV-RU, 17 in the order SV-RU-FV, and further 16 in the order RU-FV-SV. In total 49 subjects (24 Female, 25 Male, mean age=23.8, sd=7.1) with different education backgrounds (pupils, university students, craftspeople, office employees) participated. All experiments were conducted with healthy attendees instead of patients with brain function disorders.

After a short tutorial in a simple virtual building, where participants learned the respective navigation metaphor and got accustomed to the current viewing setup, the participants performed the following actions per experiment:

1. Be spawned in the central room and memorize its picture.

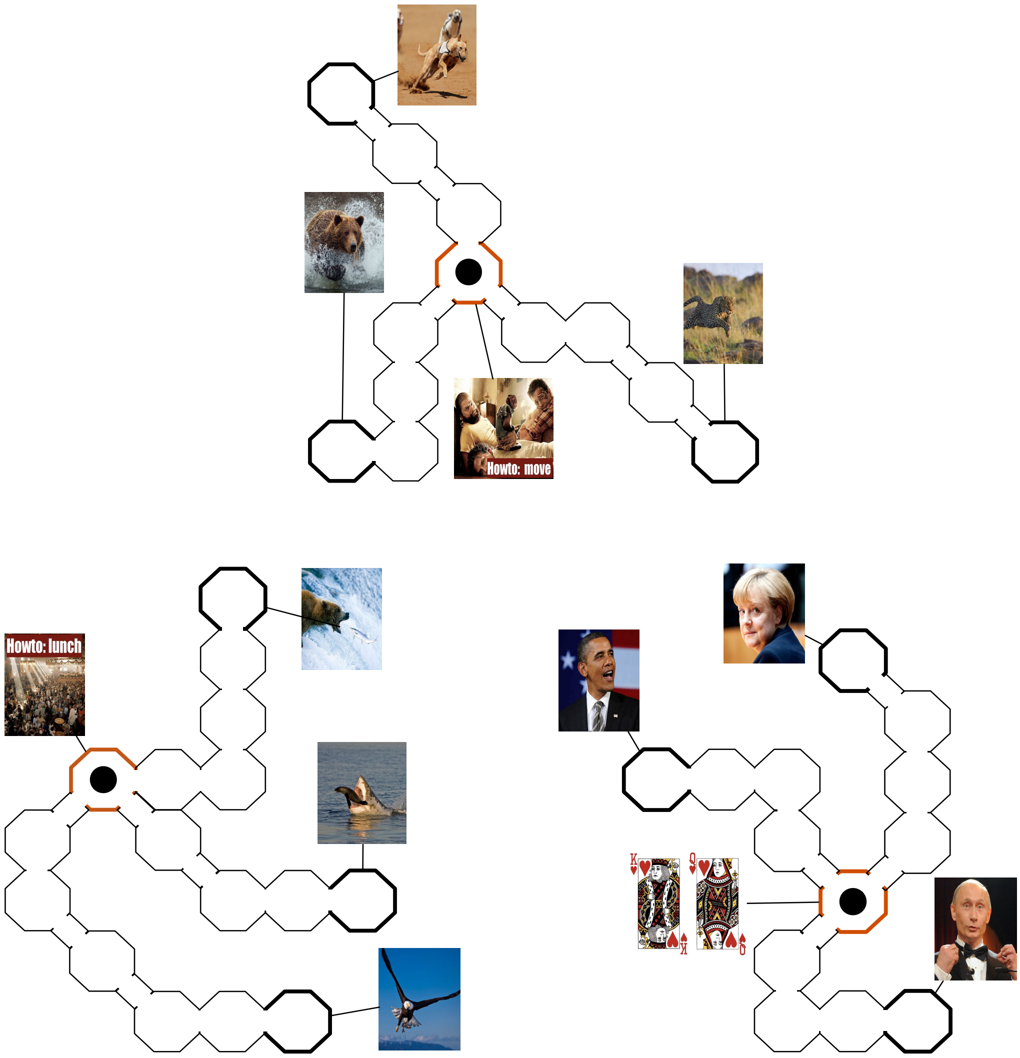

Figure 5.2: Maps of the three museum-like mazes each participant saw during consecutive runs. Participants always started at the central room (black dot). Here and in the rooms at the end of each corridor an image was displayed. These four images were to be remembered in their spatial relationship to each other and pointed to during the pointing task.

Figure 5.3: Screenshot of a maze as participants saw it.

2. Walk to the end of each corridor, memorize the picture in the last room, and walk back to the central room.
3. Point from the central room to each of the pictures in the corridors in the order visited.
4. Again, walk each corridor to its end, point back to the image in the central room and to the images in the other corridors, and travel back to the central room.
5. Point from the central room to each of the pictures in the corridors in the visited order.

After completing the set of three experiments, the participants had to fill out a questionnaire. Finally, the participants' mental rotation capabilities were tested with the paper-and-pencil *Mental-Rotation-Test* by Vandenberg [VK78] in its A-version (MRTA).

For each condition, two types of measurement were employed to quantify spatial presence and spatial orientation skill:

Pointing Task: First, a pointing task was performed within the virtual maze [CGBL98]. Pointing to an object was done by touching a touch-screen, which issued a ray from the viewer's position in VR through the touched pixel of the respective screen/view. The horizontal angular difference between this ray and the exact direction to the object in question was used as the angular error.

Questionnaire: Second, a combined questionnaire was filled out. The four-items version of the MEC SPQ with a 1–5 scale (1="not at all", 5="very much") was extended such that for each question the participants had to give three answers, one for each condition. Also, five questions rating general aspects of the system were added, again for the different conditions: fun, tiredness, cyber-sickness, intuitiveness of control metaphor, and realism of control metaphor.

5.3 Results for Frontal-View vs. Surround-View

First, we compare the questionnaire results and pointing task accuracies of the frontal view (FV) condition to the surround-view (SV) condition.

Since the data from the MEC SPQ was not normally distributed and a log-normal transformation did not solve the problem, the results were calculated with the non-parametric Mann-Whitney-U test. These computations discovered no significant effects on the sub-scales *Attention*, *Involvement*, and *Suspension-of-Disbelief*, but very clear ones on the following sub-scales:

- *Spatial Situation Model* ($W = 26011$, $p = 0.004$)
- *Spatial Presence Self Location* ($W = 8217$, $p = 0.000$)
- *Spatial Presence Possible Actions* ($W = 25031$, $p = 0.001$)

Figure 5.4, left, shows the mean values of all sub-scales. Note that all significant sub-scales directly reflect spatial presence. In contrast, attention, involvement, and suspension-of-disbelief are more related to general presence, not spatial presence in particular. This suggests a distinguished impact of the additional five screens on the participants' subjectively experienced location.

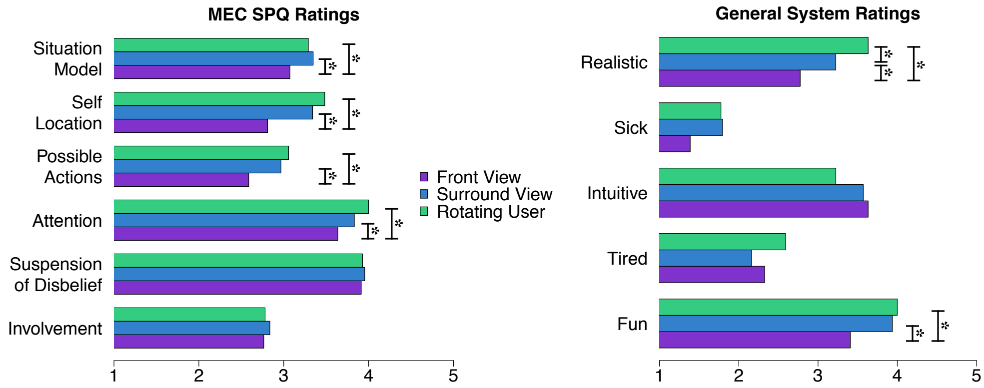

Figure 5.4: Left: Ratings of MEC SPQ comparing its sub-scales for all three conditions ($*$ = significant). Right: Ratings for the general system aspects ($*$ = significant). In both charts the rating 1 means "not at all", 5 means "very much".

Analyzing the general system ratings (Figure 5.4, right) with the Mann-Whitney-U test, the following two effects were found to be significant:

- *Realism* ($W = 923.5$, $p = 0.021$)
- *Fun* ($W = 807.5$, $p = 0.002$)

The differences between the ratings for *Cyber-sickness*, *Intuitiveness of control*, and *Tiredness* were not significant. On the one hand, this is somewhat contrary to the literature claiming a wider field of view causing more cyber-sickness [LaV00, SKHR01]. On the other hand, 135° is already a wide field of view to begin with. Since the control metaphor stayed the same for both conditions, it is no surprise that significances for the questions related to its intuitiveness did not manifest.

Since the pointing tasks differed in complexity, the following groups of tasks had to be analyzed individually:

- from every picture to every other,
- from a specific corridor-picture to every other picture,
- from a specific corridor-picture to another specific one,
- from a specific corridor-picture to the center-picture,
- from the center (1st time) to a specific corridor-picture,

- from the center (2nd time) to a specific corridor-picture.

To avoid distortions because of the different sub-experiment order per participant, for this analysis only the data from the first sub-experiments (Maze 1, respective first condition) was considered. This data also was not normally distributed, so again the Mann-Whitney-U test had to be used for comparison. Contrary to the expectations, no significant differences between any groups of pointing tasks in the FV condition compared to the SV condition could be noticed. Figure 5.5, left, compares the pointing errors between conditions. Though the spatial presence was found to be significantly higher with a surround view, it seems to have no effect on the actual orientation in the same virtual environment.

5.4 Results of Rotating World vs. Rotating User

After showing the effects of extending the field of view, we now look into the effects of user embodiment, comparing the results of the surround-view (SV) condition (stable user, rotating virtual space) to the rotating-user (RU) condition (stable virtual world, self-rotating user).

The MEC SPQ ratings (Figure 5.4, left) revealed no significant differences except for the *Attention* sub-scale ($W = 17398$, $p = 0.045$). In a dialog after the experiment, many participants stated that they were very focused on the control mechanism in the RU condition. Hence, the questions regarding the *Attention* sub-scale may have been misunderstood in the sense that they were answered with respect to the attention on the hardware controls and not on the task in the virtual maze. According to the MEC SPQ results, disabling virtual world rotation and instead performing direction changes by embodied self-rotations does not have a significant effect on spatial presence.

For the general system questions the participants significantly ($W = 904.5$, $p = 0.014$) distinguished their ratings only on the *Realism* scale in favor of the RU condition (Figure 5.4, right). This is an important result suggesting the removal of the abstraction layer introduced by common control

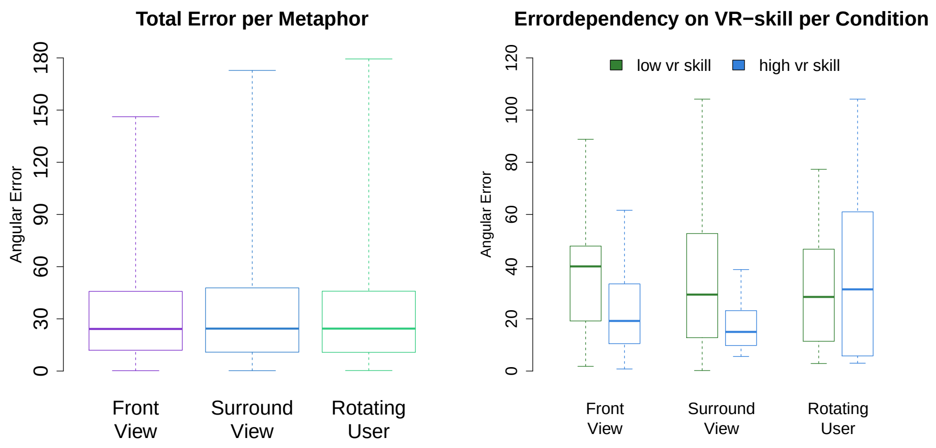

Figure 5.5: The angular pointing errors. Left: Angular error considering every pointing action in the first maze, showing very similar medians and distributions for all three conditions. Right: The angular error of pointing tasks for corridors just traveled decreases with higher VR-skill for the FV and SV conditions, but is independent of it in the UR condition. The lines in the boxplots represent median values.

metaphors, such as joysticks or other devices used to rotate the world, instead of having an embodied rotation like we are used to from the real world. However, the results for *Intuition of control* are not significant. One can argue that this is due to the fact that most participants are young, and even if they reported a low VR-experience, they are already somewhat familiar with the rotating world paradigm (e.g., through computer games), but not with self-rotation. The attention to the unknown control metaphor may have had a negative influence on the *Intuition* ratings.

Similar to the step from FV to SV, the step from SV to RU yields no significant differences in pointing errors between the conditions for the first maze. In fact, Figure 5.5, left, shows that the error distributions are nearly the same for all three conditions if looked at *in total*.

A closer examination of the specific pointing tasks, like in Section 5.3, this time revealed a difference in correlations between pointing error and VR-skill between the two conditions. Namely, for pointing from the end of a corridor back to the central room a Spearman correlation factor of $\rho \approx -0.35$ was found for both the FV and the SV condition. On the contrary, in the RU condition, the correlation factor was nearly zero. Figure 5.5, right, shows the dependency of pointing error on VR-skill for all three conditions. Since self-ratings of VR-skill are error prone and no participant stated himself a VR-skill of 5, participants were classified into low VR-skill (1–2) and high VR-skill (3–4). Though statistically not significant, the chart clearly shows that in the conditions where the world was virtually rotated (FV and SV) the error decreases with increased VR-skill, meaning the user benefited from former VR-experience. In the RU condition, however, VR-skill had no influence on the pointing accuracy. Here participants with low VR-skill produced nearly the same angular error as participants with high VR-skill.

Since the average VR-skill per group was similar (FV=2.5, SV=2.1, RU=2.1) and also the means of the MRTA (FV=6.3, SV=6.3, RU=6.5) did not differ much, this observation truly seems to depend on the metaphor only.

Because this effect was not true for arbitrary pointing tasks, user rotation appears not to be beneficial for virtual map learning in global (conceptual) space. However, since it was true for the specific pointing task at the end of a corridor just traveled, we can suspect it to contribute to orientation and the ability to locate oneself in the immediate (perceptual) space.

5.5 Discussion

The questionnaire results clearly indicate that expanding the user's view to a full 360° surround-view enhances the feeling of spatial presence in the virtual scene. It also improves realism and apparently is more fun. In addition to the recorded statements many participants in the Surround-View condition looked behind themselves before a pointing action, although they could also have rotated the virtual world in front of them. This furthermore supports the actual use of the displays behind the user. However, the quantitative measures of angular error in the pointing task do not show any improvement of accuracy. But this is due to the fact that the involved senses and their stimuli are nearly the same for both the FV and SV conditions. The difference in the *peripheral* sight of 22.5° (=(180°-135°)/2) for each side may be neglected.

Changing the rotation metaphor from a stable user and a rotating virtual world to a stable virtual world and a self-rotating user further improves significantly on realism. Though otherwise the questionnaire results for the Rotating-User condition do not significantly vary from the Surround-View condition, they do so compared to the Frontal-View condition for spatial presence *(Situation Model, Self-Location, Possible Actions)* and *Fun*. This means that the new control metaphor in RU does not undo the enhancements gained by the transition from FV to SV. For pointing actions along corridors just visited the independence of the pointing error from the VR-skill in the RU condition suggests an advantage over the FV and the SV condition for immediate self-location, not for map learning in general.

Since the participants were rather young and all somewhat accustomed to virtual environments and computer controls, we may encounter different results for an elderly group without any such knowledge. Only a single participant of this study fulfills this criterion. At the age of 56, his median angular errors for the pointing task were 113.3° (FV), 83.4° (SV) and 49.4° (RU). Within-subject differences between conditions this big were not observed for any younger participants.

5.6 Summary & Outlook

In this chapter, we evaluated the effect of the full horizontal surround view and the embodied self-rotation of the OctaVis system on spatial presence and spatial orientation. The results clearly indicate that both features are indeed beneficial in the context the OctaVis was designed for: training and rehabilitation of spatial cognition for patients with brain function disorders, e.g., due to a stroke. This target group typically consists of elderly people with no prior experience in computer games or virtual reality. As a consequence, they benefit from the surround view as well as from the rotating-user control metaphor.

A carefully designed hardware providing a high sense of presence (Chapter 3), an optimized software smoothly rendering and easily operable by non-technical staff (Chapter 4), and general benefits for spatial learning (this chapter), of course, do not entirely satisfy the requirements posed on a new VR device build for virtual therapy. Even though appreciated by patients and examiners, a therapy device has to actually produce therapy success for real patients with real limitations. Therefore, in the next chapter, some results of different clinical trials will be presented.

CHAPTER 6

Clinical Evaluation

Prove all things; hold fast that which is good.

— Apostle Paul

In the introduction chapter of this thesis, we set out to develop a novel device and therapy environment benefiting both, examiners and patients. The OctaVis with its technical arrangement and CE-certificate (Class 1 medical device in Germany) obviously suits the hospital and the patient side from a technical perspective. The more prominent aspect, which builds on this technical base, however, is a measurable improvement in visuospatial, memory and executive function performance in patients training in the supermarket chosen as the virtual therapy environment.

This chapter first presents the essence of several studies investigating the success of learning in the OctaVis and, second, exemplifies the diagnostic capability of the device.

6.1 Training & Rehabilitation

In a clinical context, it is very important to emphasize that the grocery shopping task was not arbitrary chosen. As already mentioned in Section 3.4, the paradigm applied in the following studies is based on the rationale of classic neuro-psychological tests of verbal learning and memory, e.g., the CVLT [NSTOW08] and the VLMT [HLL01]. In contrast to these tests, learning in the VR shopping task is not an isolated training of mere memory, but rather a multi-modal encounter more similar to real-life learning.

The basic idea, as mentioned earlier, is to listen to an auditory presentation of a list of 20 grocery products (*List A*) to be bought, then to find their location in the supermarket, and to buy all remembered items. This procedure is repeated for several trials (*training*), once per trial. The trial after the training (*interference*), items from a distraction list (*List B*) have to be bought. This list contains different but easily confused products compared to List A (e.g., champagne instead of wine). To test the stability of learning List A, the participants have to buy items according to List A without hearing it again (*free recall*) after the interference trial.

The time periods between trials and the specific amount of total trials vary for the different experiments presented in this section. Hence, they are pointed out for every study individually.

All studies presented in this section were conducted by the psychology team of the CITmed project. The data for the epilepsy patients was gathered in cooperation with the epilepsy center in Bethel, the experiments with the stroke patients were performed in the stroke unit of the neurology clinic in Bethel.

6.1.1 Stable Learning

The first study investigated and measured training success for 19 healthy university students (5 male, 14 female) of age 19–28 years ($m = 23$, $sd = 3.45$), and a small group of epilepsy patients (4 male, 1 female) of age 25–47 ($m = 35.04$, $sd = 8.08$). The task was exactly the same as described above, but, due to faster learning, for the healthy students limited to a five-day period, where on the fourth day the distractive list was presented (Figure 6.1, top). In case of the healthy participants the study revealed a stable learning effect for the number of correctly bought products ($F = 18.82$, $p < .001$), where the distractive list on day 4 has almost no influence

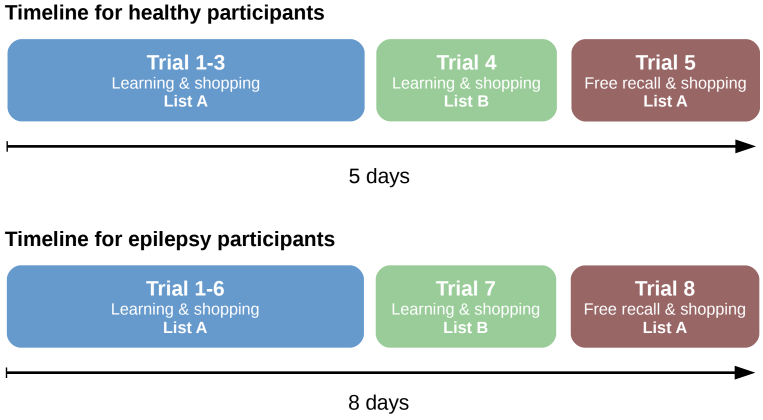

Figure 6.1: Schematic of the course of trials. *Top:* Healthy participants trained for 3 consecutive days with List A. The interference trial was performed at day 4, and the free recall trial the day after. *Bottom:* Epilepsy participants trained for 6 consecutive days with List A. The interference trial was performed on day 7, and the free recall trial the day after.

on the performance on day 5 (Figure 6.2, left). For the epilepsy group, which performed an 8-day treatment (Figure 6.1, bottom), the results are qualitatively very similar, but for a quantitative evaluation more patients are needed (Figure 6.2, right). The negligible influence of the distractive list is in contrast to the standard verbal learning tests [NSTOW08, HLL01] and proves the efficacy and stability of the multi-modal learning in the OctaVis setup. More detailed results from this study can be found in [GKF+13].

6.1.2 Generalized Learning

Having established a stable learning effect for healthy, young university students and a small group of epilepsy patients, a different study extended the focus to: 13 people with focal epilepsy, being 19–51 years old ($m = 32.3$, $sd = 10.0$), 11 stroke patients within the age range of 34–76 ($m = 61.0$, $sd = 15.2$), and 13 healthy seniors of age 61–94 ($m = 71.4$, $sd = 10.8$). In addition to the memory performance in 8 trials during the coarse of at most 14 days, here, spatial thinking and its generalization beyond the supermarket task was investigated. Each of 6 trials of the training phase was performed

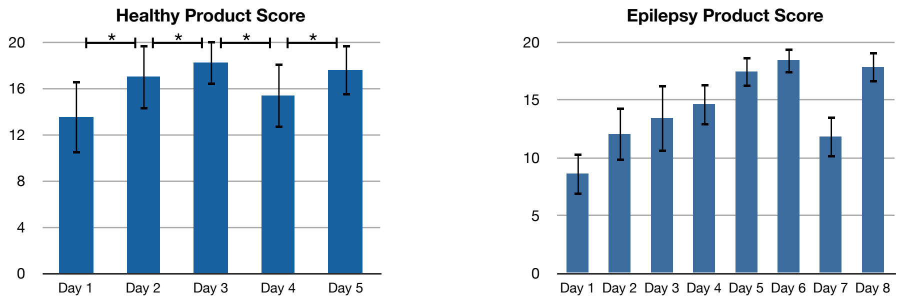

Figure 6.2: Product scores of grocery shopping in the virtual supermarket. Stars mark significances. Error bars depict standard deviation. *Left:* Healthy university students with five days of shopping. Day four holds the results for the distractive shopping list. *Right:* Epilepsy patients with eight days in the supermarket. Day seven holds the results for the distractive shopping list. (Significances were not calculated due to the small sample of 5 epilepsy patients.)

within a time interval of 48h after the former. To properly measure the effect of the interference trial, trials 6,7 and 8 were allowed a maximum in-between-time of 24h. Trial 7 consisted of two runs: The first run was the interference trial with List B, and the second, which was performed immediately afterwards, a free recall and shopping of List A. A visual aid can be found in Figure 6.3.

Figure 6.4 shows the number of correctly bought items per trial. In general, one can observe that, same as in the former study, the numbers increase during the training phase, and that the distractive trial on day 7 has only a very small effect on the two subsequent trials with free recall of List A. A repeated measures ANOVA proves statistical significance of these improvements for seniors ($p = 0.003$), stroke patients ($p = 0.024$) and epilepsy patients ($p < 0.001$). Similar to the improvement in product score, the length of the walked trajectory and required time decreased during the eight days (by about 30% and 40%, respectively), which shows an improvement of spatial orientation, map learning, and executive functions like path planning.

To analyze the general improvement of visuospatial performance, participants carried out the Rey Figure [Rey41] test before the first trial and the comparable Taylor Figure [Tay69] test after the last trial. Concerning

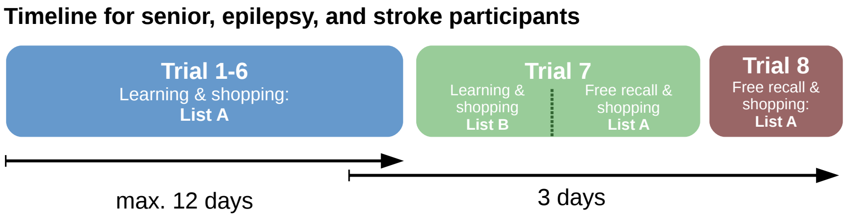

Figure 6.3: Schematic of the course of trials. The training phase consisted of 6 trials. These 6 trials had to be performed within a period of 48h between each trial. Trial 6,7, and 8 were conducted on consecutive days. Trial 7 was divided into two runs: The interference run with List B was immediately followed by a free recall and shopping of List A. The experiment finished with a further free recall and shopping trial on the day after the interference trial. The maximum total amount of days adds up to 14.

the visuo-constructive ability to assemble a figure from its components, all groups improved in average (seniors: 3.7%; stroke: 13.3%; epilepsy: 3.6%), but according to the Wilcoxon calculus only the patients showed statistical significance (stroke: $z = -2.052$, $p = 0.040$; epilepsy: $z = -2.455$, $p = 0.014$). For the visuospatial memory, participants had to remember and draw the figure after 30 minutes. Again each group improved (seniors: 2.0%; stroke: 18.8%; epilepsy: 15.7%), but only the patients did so significantly (stroke: $z = -1.956$, $p = 0.050$; epilepsy: $z = -2.276$, $p = 0.023$). It is plausible that the seniors did not improve significantly since they had no brain injuries to recover from. These results clearly show that the OctaVis contributes to a *general* improvement in visuospatial cognition in patients.

Finally, a questionnaire concerning the perceived efficacy of the training was filled out after day 8 (Figure 6.5). On a 0–6 scale (0=not at all, 3=average, 6=very much), participants rated whether the OctaVis experience was *interesting*, *motivating*, and *useful*. All groups rated these items very high ($> 4.5 \in [0, 6]$), which is an important result since we know that internal motivation is crucial for treatment success. Participants then rated whether they think they improved on *memory*, *orientation*, and their grocery *shopping* performance. While both patient groups rated all learning areas above average, the seniors did so only for the memory gain. This is

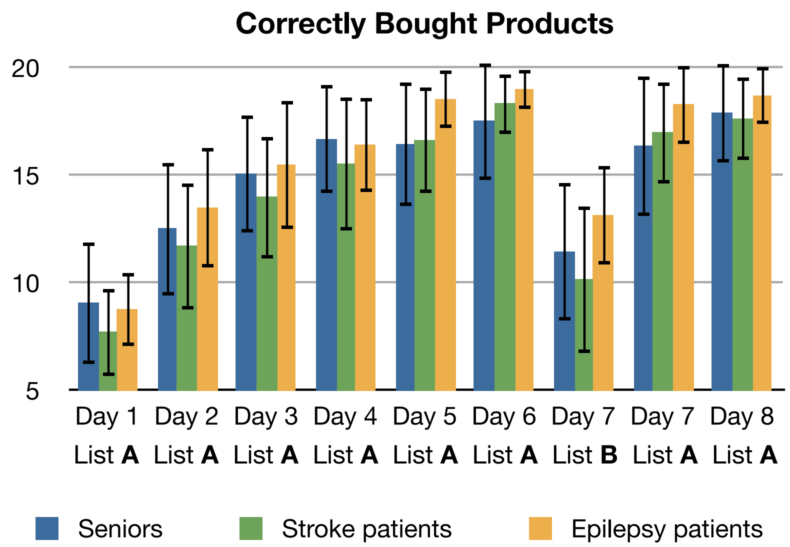

Figure 6.4: Results of OctaVis training for healthy seniors, stroke and epilepsy patients depicting the number of correctly bought items for each trial.

again plausible, since the healthy seniors did not need help with orientation or shopping.

To compare the generalization effect of the VR training to a neuro-psychological standard therapy, in a further study [GRL$^+$13], 15 patients with a focal epilepsy trained (according to the course described in Figure 6.3) in the grocery shopping setup in the OctaVis while 13 other epilepsy patients participated in the standard neuro-psychological therapy for the same period of time, acting as a control group. Before and after the training, an extensive neuro-psychological examination was performed to measure several cognitive traits. The results yield an improvement of visual memory for both groups. The VR group significantly improved on visuospatial, spatial-constructive and visual-mnestic abilities. Only the VR group significantly improved on the Bergen-Left-Right-Discrimination Test [Oft02], and on the visual-constructive aspect during sketching of geometric figures. Patients in total rated the VR training more suitable and motivating for learning for everyday life challenges.

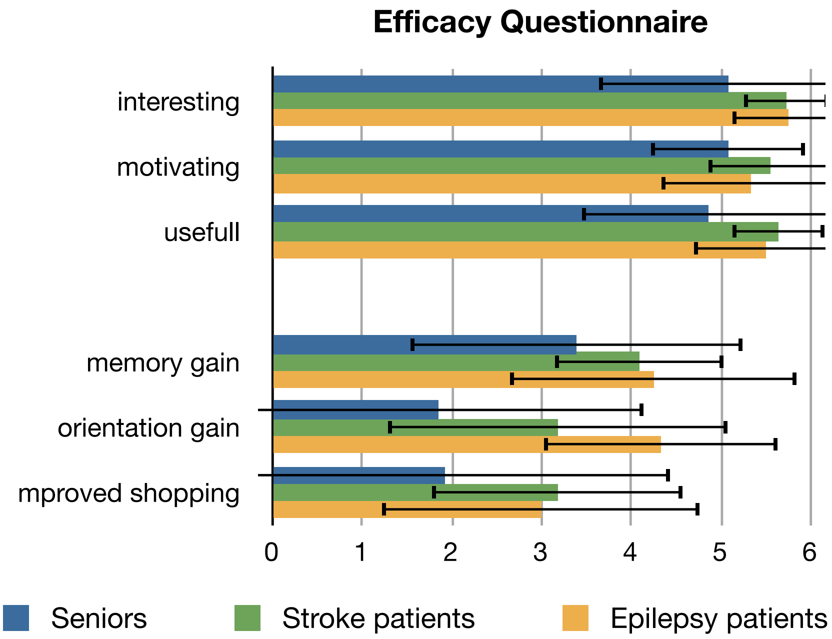

Figure 6.5: Results of OctaVis training for healthy seniors, stroke and epilepsy patients depicting the ratings of training efficacy, concerning general appreciation (top) and perceived learning effects (bottom).

6.1.3 Real-Life Correlation

A subgroup (7 patients with focal epilepsy) of another study, which dealt with the ecological validity of the supermarket task, went shopping in a real supermarket before training in the virtual one. Participants had to memorize the items from a list written on paper. They were given 2 minutes for that task, to proper encode the list. Afterward, participants had to find the remembered items and to "buy" them by tipping them. A real shopping cart with a special wheel for measuring distance was used during the whole shopping task, which also was timed with a stopwatch.

Since the virtual supermarket does not match the floor layout of the real supermarket and offers a different set of products, direct comparison of results is not possible, but correlations were found for (1) *correct products per time* across all VR learning days with correct products per time in the real supermarket ($p = 0.050$) and for (2) *correct products per distance* across all VR learning days with correct products per distance in the real supermarket ($p = 0.033$).

This data supports a generalization of the results achieved in VR not only to a visuospatial paper-pencil test (see former section), but also to real-life visuospatial challenges. The in-depth analysis of this study can be found in [GLK+14].

6.2 Diagnostics

In hospitals, to decide for the proper treatment, the correct and precise diagnosis is a crucial part. Collaboration partners of the CITmed project at the NeuroCure Cluster of Excellence in Berlin conducted a study to test such a diagnostic application of the OctaVis [UWM+13]. Further pressing on the comparison to standard techniques, they looked into the diagnosis of visual neglect by the usual paper-pencil Neglect-Test (NET) opposed to a diagnosis in the OctaVis. The surround view of the OctaVis allows for a spatially broader assessment of the field of view, which is corrupted by the neglect. A group of 10 elderly patients performed four subtests (Star Cancellation, Line Bisection, Dice Task, and Puzzle Test) of the NET and in the OctaVis. A control group of healthy participants with a matching range of age ($m = 60, sd = 8$) did the same. As expected, the control group performed significantly better than the patient group in both conditions. Together with higher ratings for enjoyability, the OctaVis, therefore, presents itself as an at least equally capable alternative to conventional diagnostics. An assignment of all the participants to the corresponding group based on their performances in the two tests further revealed a more precise success rate for the OctaVis condition (90% hit rate) than for the NET (70% hit rate).

6.3 Summary

From the rehabilitation perspective, the conducted clinical studies render the OctaVis system at least competitive with traditional treatment for limited memory, spatial thinking, and executive functions.

The VR device promotes comfortable multi-modal learning as required by the supermarket task, which, in contrast to the standard CVLT, results in more stable learning for healthy young and senior individuals as well as for epilepsy and stroke patients.

Further, training in the OctaVis significantly improves *general* spatial thinking, beyond the specialized shopping task. Comparing this generalization to that of a standard neuro-psychological therapy, only VR training results in a significant improvement in right-left discrimination and visual-constructive abilities. The generalization even seems to hold for shopping in a real supermarket.

On a personal level, different patient groups rated the VR training very interesting, motivating and useful for (re)learning everyday life tasks, which is very important for the psychosomatic aspects of treatment.

Finally, the OctaVis was proved to also contribute to a more reliable diagnosis of neglect patients, benefiting examiners and patients.

Chapter 7

Conclusion

In the end, people are persuaded not by what we say,
but by what they understand.

— John C. Maxwell

Possible psyche changing effects of virtual reality, a make-believe world, of course, can be seen critically, for instance in the context of excessive gaming etc. In this thesis, however, we have witnessed that the power to change can be employed for successful treatment of impairments in spatial thinking, memory, and executive functions as well if channeled carefully. In fact, improvement in these areas strongly relies on proper immersion into a make-believe world or task.

In the OctaVis, the immersion into this virtuality, according to the two-step model of spatial presence, is achieved by an adequate balance of *technology*, *task*, and patient's *personal goals* or limitations.

In terms of *technology*, targeting improvement of spatial cognition, the virtual environment for therapy is presented in a 360° surround view on eight displays powered by a single PC (with three GPUs) running a custom-tailored render-engine. The system moreover provides a real-world rotation of the patient for orientation and navigation. In total, these technology decisions have been proven to enhance spatial understanding and the sense of presence.

In terms of activity or *task*, the grocery shopping is seamlessly integrated with the technology. Buying items by touching them on the touch-screens and navigating the aisles by a wheel-chair metaphor is perceived as natural. The task in general, requiring memorization of the shopping list, spatial attention to recognize the items in the shelves, map-learning, and path-planing, fully engages the patient's cognitive involvement and is also strongly supported by the surround view and the real user rotation.

In terms of *personal goals*, patients are highly motivated by the choice of a basic multi-modal real-life task. Being enabled to shop on their own, which strengthens the feeling of self-sufficiency, patients are pulled into participating in the training. Also, the technology itself promotes a motivating fun factor.

Apart from the immersive capabilities, for virtual therapy in a clinical setting, a VR device has to meet several general requirements. The OctaVis system is easy to use by patients and by staff, easy to maintain, cost and space efficient, and it meets the requirement of patient safety and supervision. It is an officially CE-certified Class 1 medical device in Germany and has been deployed to four different hospitals for trials with epilepsy, aphasia, depression, stroke and visual neglect patients. Both examiners and patients appreciate the device, not only for its simplicity but also for its suitability in their respective patient context.

Applying the OctaVis as therapy device in a clinical environment demonstrated its superiority over traditional approaches for both, stability and generalization of learning.

In conclusion, the OctaVis is ready for clinical application, namely to diagnose and rehabilitate limitations in at least visual-spatial cognition. It probably should not replace every existing treatment and diagnostic method. Instead, it can act as a supplement for patients with low motivation for training or for patients generally affine to or curious about technology. Also, it can complement traditional diagnostics with a more detailed investigation of the field of view or where rotation is involved.

Further work on the presented system could involve more displays. A second row of monitors, for example, would allow for a higher vertical field of view. A horizontally mounted monitor in front of the chair could represent the shopping cart showing already bought items. Though not significant in the study results, displays with smaller frames or none at all could contribute to a smoother experience.

Instead of an HMD, which clouds the vision of the patient's hand during the buy action, a mixed-reality approach may yield interesting results but comes with the disadvantage of wearing a helmet. Glasses, on the other side, may provide an acceptable solution. Microsoft's *Holo Lense* or Google's *Google Glasses*, if allowed for custom content, seem to be interesting candidates. Combined with a force feedback data glove, the buy-experience could be enhanced to a full grasp-action, opposed to the touch of a monitor.

A promising project to enhance our render engine was launched on February 2016 by Khronos. Their Vulkan specification targets low level GPU control removing the high-level driver abstraction of OpenGL. One interesting feature for our setup seems to be *multi-threaded render command uploads*. In contrast to the one-thread-per-context philosophy in OpenGL, this allows for a draw call parallelization of even the three displays driven by the same GPU. Since our render tasks are not CPU bound, this technique will not necessarily result in a big performance gain, but nevertheless offers further optimization opportunities. Another feature of Vulkan is *memory control* allowing applications to map virtual addresses to physical memory pages. Copying resources to the GPU can be done asynchronously in a special direct memory access (DMA) queue. Such a fine grained memory control may be utilized to enhance data management between the three GPUs. Unfortunately, Vulkan 1.0 does not support resource sharing between GPUs. However, it allows the use of, for instance, intel's on-chip GPU for a computation task and allows the results to be send to a GPU of a different vendor. In total, this opens the access to more GPU power. A third interesting development is a new *driver model*. With the introduction of Vulkan, GPU vendors are forced to provide special thin-layer Vulkan drivers to allow for the mentioned low level driver access. These new drivers together with Vulkan's novel extension for window system integration (WSI) enable more control over data distribution. Since in our project we had to exclude

NVIDIA GPUs due to their strict distribution policy between windows, this problem may be overcome by Vulcan.

Additionally, the software suite housing the actual therapy environment could be extended to a toolkit. Such a toolkit could provide a simple level editor to change the complexity and the layout of the supermarket or to create new buildings. Applying the theory of space-syntax, from just a few parameters, random buildings with complexities specific to the impairment to be treated could be generated procedurally and fine tweaked afterward by hand. Clinicians, with such an intelligent and easy to use level designer at hand, could design experiments almost without the help of technical personnel.

Regarding the insights of game design, it would be interesting to experiment with dynamic challenge changes according to the current skill level of the patients. Products could be rearranged or even the shelf locations could be altered between trials. Some products could be on sale and therefore at a different location etc. The mission could be changed to buy items while minimizing costs (e.g., by avoiding expensive brands). Also, elements of surprise could be added (e.g., tipping over a shelf on collision). Typical high-score boards or in-game rewards could have positive effects on the participants' motivation. For further motivation, the virtually bought items could make up an actual recipe of the patient's choice, which after the training would be delivered from an actual supermarket and cooked instead of the usual hospital food. Of course, this would result in additional work for the nursing staff. However, according to TECHINASIA[1] Homeplus[2] (a grocery store in Korea) already offers the home delivery of virtually bought groceries. Their customers can scan a QR-code of a product displayed anywhere in the city and purchase it within a mobile application.

In the end, this project has proven its potential for truly meaningful applications. It has survived its research- and initial clinical testphase, but whether it will live up to its ultimate purpose heavily depends on the measures taken to move it out of the research community into the pragmatic realm of clinical practice.

[1] `www.techinasia.com/homeplus-virtual-store-south-korea`
[2] `www.homeplus.co.kr/`

Appendix A
Statistical Testing

The studies mentioned in this thesis use several different tests to calculate significant distinctions between participant groups or between experiment conditions. This appendix shortly describes these tests and the meaning of their characteristic variables according to IBM's SPSS Software Guide[3] and the R language manual[4].

To decide for a specific test, the following two aspects of a measurement have to be identified:

Dependent vs. Independent A data sample is independent of another if both come from distinct populations (a group of samples). These populations can originate from the same participant group, but must not affect each other. The appreciation ratings for a navigation metaphor by different participant groups, for instance, are considered independent. The number of correctly bought items in a supermarket on consecutive days, on the other hand, is considered dependent because each following day profits from the gathered knowledge of the former.

Parametric vs. Non-parametric Data samples distribute around a mean value. Unfortunately, the function to describe this distribution varies depending on the experiment. Parametric distributions are those which can be described by a finite set of parameters, e.g., the normal distribution. Here σ and μ are two parameters to define a bell curve around a mean value. If a distribution cannot be described by a finite set of parameters, so-called non-parametric descriptions are employed. Such distribution functions are assumed to have an infinite number of parameters or one parameter with infinite dimensions. Here, the knowledge of the distribution function is allowed to grow with each

[3] `statistics.laerd.com`
[4] `cran.r-project.org`

examined sample. Parametric distribution functions stay the same independent of the underlying amount of samples.

Having established the data dependency and the type of distribution of a measurement, there is a variety of tests to choose from to calculate significant deviations of one sample group from another:

Z-Test Assuming a normal distribution around a certain value, this test detects how far the mean of a population is apart from that value. The Z-score measures this distance in units of standard deviation.

Mann-Whitney-U / Wilcoxon Test The *Mann-Whitney-U Test* is a non-parametric test for independent samples. It compares two populations without any assumption about their underlying distribution function. The samples can be either on a continuous or on an ordinal scale. The basic idea of the test is to rank every observation and then for every sample of the first population to count its wins over every sample of the second population by comparing ranks. The calculated result value is called U, or W if we use the R language as statistics tool, and corresponds to the sum of ranks that won. The dependent version of this test is often referred to as the *Wilcoxon Test*. Often, a Z-score is reported for these tests since for increasing sample sizes the distribution of W approximates the normal distribution. Instead of measuring the mean distance of the actual samples from the expected value, Z then measures the distance of the calculated W from its expected value under the null-hypothesis. This is the hypothesis that there is no significant difference between populations.

Repeated Measures ANOVA The purpose of this parametric analysis of variance (ANOVA) is to detect differences in means of at least three related/dependent consecutive measures. For instance, we could measure the loss of weight during a three-month period while being on a carbon diet. If the measures are taken every week, this test could help to determine when the biggest effects occur etc. The calculated result value F is the ratio of two variances. The variance between groups is divided by the variance within the groups. A group in the dieting example would be one of the weekly measuring times.

For all mentioned tests a p-value describes the probability of finding the observation in question. Results are commonly assumed to be significant for $p < 0.05$.

Bibliography

[AM90] Patricia L. Alfano and George F. Michel. Restricting the field of view: Perceptual and performance effects. *Perceptual and motor skills*, 70(1):35–45, 1990.

[Bar16] Barco. Barco Dome Displays. `www.barco.com/en/Products/Proprietary-all-dome-digital-projection-system.aspx`, June 2016.

[BBH05] Steffi Beckhaus, Kristopher J. Blom, and Matthias Haringer. Intuitive, hands-free travel interfaces for virtual environments. In *New Directions in 3D User Interfaces, Workshop of IEEE VR 2005*, pages 57–60. Shaker Verlag, 2005.

[BIPS12] Gerd Bruder, Victoria Interrante, Lane Phillips, and Frank Steinicke. Redirecting walking and driving for natural navigation in immersive virtual environments. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4):538–545, 2012.

[BJH+01] Allen Bierbaum, Christopher Just, Patrick Hartling, Kevin Meinert, Albert Baker, and Carolina Cruz-Neira. VR Juggler: A virtual platform for virtual reality application development. In *Proceedings of IEEE VR*, pages 89–96, 2001.

[BM07] Doug A. Bowman and Ryan P. McMahan. Virtual reality: How much immersion is enough? *Computer*, 40(7):36–43, July 2007.

[BW10] Hrvoje Benko and Andrew D. Wilson. Multi-point interactions with immersive omnidirectional visualizations in a dome. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 19–28, 2010.

[BWÖ04] Aysu Baskaya, Christopher Wilson, and Yusuf Ziya Özcan. Wayfinding in an unfamiliar environment: Different spatial settings of two polyclinics. *Environment and Behavior*, 36(6):839–867, 2004.

[CGBL98] Sarah S. Chance, Florence Gaunet, Andrew C. Beall, and Jack M. Loomis. Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments*, 7(2):168–178, 1998.

[CNSD+92] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. The CAVE: Audio visual experience automatic virtual environment. *Communication of the ACM*, 35(6):64–72, 1992.

[Cor16] HTC Corporation. Htc vive. `www.htcvive.com`, June 2016.

[Csi75] Mihaly Csikszentmihalyi. *Beyond boredom and anxiety*. Jossey-Bass Publishers, 1975.

[Das16] Dassault Systèmes. 3DVIA Virtools. `http://www.3dvia.com/products/3dvia-virtools/`, June 2016.

[DCC97] Rudolph P. Darken, William R. Cockayne, and David Carmein. The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 213–221. ACM, 1997.

[DE00] Ufuk Dogu and Feyzan Erkip. Spatial factors affecting wayfinding and orientation: A case study in a shopping mall. *Environment and Behavior*, 26(6):731–755, 2000.

[DKND11] Sebastian Deterding, Rilla Khaled, Lennart Nacke, and Dan Dixon. Gamification: Toward a definition. In *CHI 2011 Gamification Workshop Proceedings*, 2011.

[DM95] N. I. Durlach and A. S. Mayor. *Virtual reality: Scientific and technological challenges*. National Academy Press, Washington, DC, 1995.

[DP02] Rudolph P. Darken and Barry Peterson. *Handbook of virtual environments*, chapter Spatial orientation, wayfinding, and representation, pages 493–518. Lawrence Erlbaum Associates, Inc, 2002.

[DS93] Rudy P. Darken and John L. Sibert. A toolset for navigation in virtual environments. In *Proceedings of the 6th annual ACM symposium on User interface software and technology*, pages 157–165. ACM, 1993.

[EMP09] Stefan Eilemann, Maxim Makhinya, and Renato Pajarola. Equalizer: A scalable parallel rendering framework. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):436–452, 2009.

[Fis14] Matthew T. Fish. *Casual video game play as an augmentation intervention for anxiety: A controlled study*. PhD thesis, East Carolina University, 2014.

[Fra16] Fraunhofer IGD. instant**reality**. `www.instantreality.org`, June 2016.

[FRE03] Kiran J. Fernandes, Vinesh Raja, and Julian Eyre. Cybersphere: The fully immersive spherical projection system. *Communication of the ACM*, 46(9):141–146, 2003.

[GKF+13] Philip Grewe, Agnes Kohsik, David Flentge, Eugen Dyck, Christian Bien, York Winter, Mario Botsch, Hans J. Markowitsch, and Martina Piefke. Learning real-life cognitive abilities in a novel 360°-virtual reality supermarket: A neuropsychological study of healthy participants and patients with epilepsy. *Journal of NeuroEngineering and Rehabilitation*, 10(1):42, 2013.

[GLK+14] P. Grewe, D. Lahr, A. Kohsik, E. Dyck, H. J. Markowitsch, C. G. Bien, M. Botsch, and M. Piefke. Real-life memory and spatial navigation in patients with focal epilepsy: Ecological validity of a virtual supermarket task. *Epilepsy & Bahavior*, 31:57–66, 2014.

[GM98] Stuart C. Grant and Lochlan E. Magee. Contributions of proprioception to navigation in virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3):489–497, 1998.

[GRL+13] P. Grewe, O. Risius, D. Lahr, H. Ohmann, E. Dyck, C. G. Bien, M. Botsch, and M. Piefke. Evaluation eines Trainingsprogramms in der virtuellen Realität für visuell-räumliche Fähigkeiten bei Patienten mit symptomatisch fokaler Epilepsie. *Klinische Neuropsychologie*, 44:175, 2013.

[Hal08] Jonathan Halton. Virtual rehabilitation with video games: A new frontier for occupational therapy. *Occupational Therapy Now*, 10(1):12–14, 2008.

[HB95] C. Hendrix and W. Barfield. Presence in virtual environments as a function of visual and auditory cues. In *Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)*, VRAIS '95, pages 74–82. IEEE Computer Society, 1995.

[HHN+02] Greg Humphreys, Mike Houston, Ren Ng, Randall Frank, Sean Ahern, Peter Kirchner, and James T. Klosowski. Chromium: A stream-processing framework for interactive rendering on clusters. *ACM Transactions on Graphics (SIGGRAPH)*, 21(3):693–702, 2002.

[HJCBD09] Emily A. Holmes, Ella L. James, Thomas Coode-Bate, and Catherine Deeprose. Can playing the computer game Tetris reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PLoS One*, 4(1):e4153, 2009.

[HJHL08] Gregor Heinrich, Christoph Jung, Volker Hahn, and Michael Leitner. A platform for audiovisual telepresence using model- and data-based wave-field synthesis. In *Proceedings of 125th Audio Engineering Society Convention*, 2008. Paper 7608.

[HLCJ12] Michael Havranek, Nicolas Langer, Marcus Cheetham, and Lutz Jäncke. Perspective and agency during video gaming influences spatial presence experience and brain activation patterns. *Behavioural and Brain Functions*, 8(1):34–47, 2012.

[HLL01] Christoph Helmstaedter, Michael Lendt, and Silke Lux. *Verbaler Lern- und Merkfähigkeitstest*. Beltz-Test, 2001.

[Hol02] John M. Hollerbach. *Handbook of virtual environments technology*, chapter Locomotion interfaces, pages 239–254. Lawrence Erlbaum Associates, Inc, 2002.

[HWK+08] Fumiko Hoeft, Christa L. Watson, Shelli R. Kesler, Keith E. Bettinger, and Allan L. Reiss. Gender differences in the mesocorticolimbic system during computer game-play. *Journal of Psychiatric Research*, 42(4):253–258, 2008.

[HWK+12] Matthias Hofera, Werner Wirtha, Rinaldo Kuehnea, Holger Schrammb, and Ana Sacauc. Structural equation modeling of spatial presence: The influence of cognitive processes and traits. *Media Psychology*, 15:373–395, 2012.

[Inc16] Google Inc. Liquid galaxy. `code.google.com/p/liquid-galaxy`, June 2016.

[IRA07] Victoria Interrante, Brian Ries, and Lee Anderson. Seven league boots: A new metaphor for augmented locomotion through moderately large scale immersive virtual environments. In *IEEE Symposium on 3D User Interfaces*, pages 167–170. IEEE, 2007.

[JAB+12] Heidi Janssen, Louise Ada, Julie Bernhardt, Patrick McElduff, Michael Pollack, and Michael Nilsson. Physical, cognitive and social activity levels of stroke patients undergoing rehabilitation within a mixed rehabilitation unit. *Clinical Rehabilitation*, 28(1):91–101, 2012.

[JL05] Jeffrey Jacobson and Michael Lewis. Game engine virtual reality with CaveUT. *Computer*, 38(4):79–82, 2005.

[JSB10] Alex Juarez, Billy Schonenberg, and Christoph Bartneck. Implementing a low-cost CAVE system using the CryEngine2. *Entertainment Computing*, 1(3-4):157–164, 2010.

[KL04] Joshua M. Knapp and Jack M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577, 2004.

[KLB+98] Roberta L. Klatzky, Jack M. Loomis, Andrew C. Beall, Sarah S. Chance, and Reginald G. Golledge. Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, 9(4):293–298, 1998.

[LaV00] Joseph J. LaViola, Jr. A discussion of cybersickness in virtual environments. *SIGCHI Bull.*, 32(1):47–56, 2000.

[LBF+15] Matthew Lombard, Frank Biocca, Jonathan Freeman, Wijnand Ijsselsteijn, and Rachel J. Schaevitz, editors. *Immersed in Media: Telepresence Theory, Measurement and Technology.* Springer, 2015.

[LDP+02] James Jeng-Weei Lin, Henry Been-Lirn Duh, Donald E. Parker, Habib Abi-Rached, and Thomas A. Furness. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Virtual Reality, 2002. Proceedings. IEEE*, pages 164–171. IEEE, 2002.

[LLC16] Oculus VR LLC. Oculus rift. `www.oculus.com`, June 2016.

[LRZ+12] Ana Paula Cunha Loureiro, Cristiane Goncalves Ribas, Talita Gianello Gnoato Zotz, Rebeca Chen, and Flávia Ribas. Feasibility of virtual therapy in rehabilitation of parkinsons disease patients: Pilot study. *Fisioterapia em Movimento*, 25(3):659 – 666, 2012.

[Lyn60] Kevin Lynch. *The image of the city.* The MIT Press, 1960.

[MC03] Fabrizia Mantovani and Gianluca Castelnuovo. *Being There: Concepts, effects and measurement of user presence in synthetic environments*, chapter Sense of Presence in Virtual Training: Enhancing Skills Acquisition and Transfer of Knowledge through Learning Experience in Virtual Environments, pages 167–179. IOS Press, 2003.

[McG12] Jane McGonigal. *Reality is broken.* Vintage, 2012.

[Men99] Sanjay T. Menon. Psychological empowerment: Definition, measurement, and validation. *Canadian Journal of Behavioural Science*, 31(3):161–164, 1999.

[MFW08] Eliana Medina, Ruth Fruland, and Suzanne Weghorst. Virtusphere: Walking in a human size VR "hamster ball". In *Proceedings of the Human Factors and Ergonomics Society*, pages 2102–2106, 2008.

[Mid16] MiddleVR. MiddleVR for Unity. `http://www.middlevr.com/middlevr-for-unity/`, February 2016.

[MPL98] Joan McComas, Jayne Pivik, and Marc Laflamme. *Virtual Environments in Clinical Psychology and Neurocience*, chapter Current uses if Virtual Reality for Children with Disabilities, pages 161–169. IOS Press, 1998.

[MPS11] Ann McNamara, Frederic Parke, and Mat Sanford. Evaluating performance in tiled displays: Navigation and wayfinding. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, pages 483–490. ACM, 2011.

[MS09] Gesine Marquardt and Peter Schmieg. Dementia-friendly architecture: Environments that facilitate wayfinding in nursing homes. *American Journal of Alzheimer's Disease and Other Dementia*, 24(2):333–340, June 2009.

[MW98] Faith A. McCreary and Robert C. Williges. Effects of age and field-of-view on spatial learning in an immersive virtual environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 42, pages 1491–1495. SAGE Publications, 1998.

[NHM11] Braden Neal, Paul Hunkin, and Antony McGregor. Distributed OpenGL rendering in network bandwidth constrained environments. In *Eurographics Symposium on Parallel Graphics and Visualization*, pages 21–29, 2011.

[NHS04] Norbert Nitzsche, Uwe D. Hanebeck, and Günther Schmidt. Motion compression for telepresent walking in large target environments. *Presence: Teleoperators and Virtual Environments*, 13(1):44–60, 2004.

[NSTOW08] Hendrik Niemann, Walter Sturm, Angelika I. T. Thöne-Otto, and Klaus Willmes. *California verbal learning test—Deutsche Adaptation*. Pearson, 2008.

[Oft02] Sonja H. Ofte. Right-left discrimination: Effects of handedness and educational background. *Scandinavian Journal of Psychology*, 43:213–219, 2002.

[PFW09] Tabitha C. Peck, Henry Fuchs, and Mary C. Whitton. Evaluation of reorientation techniques and distractors for walking in large virtual environments. *Visualization and Computer Graphics, IEEE Transactions on*, 15(3):383–394, 2009.

[PH15] Jiri Polcar and Petr Horejsi. Knowledge acquisition and cyber sickness: A comparison of vr devices in virtual tours. *MM Science Journal*, 19(6), 2015.

[PI67] Jean Piaget and Barbel Inhelder. *The child's conception of space*. W. W. Norton & Company, 1967.

[PK10] Martina Piefke and Sina Kühnel. Empirie- und Beobachtungspraktikum: Physiologische Psychologie, Departement of Psychology, Bielefeld University. Winter Term 2009/2010.

[Pso95] Joseph Psotka. Immersive training systems: Virtual reality and education and training. *Instructional science*, 23(5-6):405–431, 1995.

[RABJ98] F. D. Rose, E. A. Attree, B. M. Brooks, and D. A. Johnson. *Virtual Environments in Clinical Psychology and Neuroscience*, chapter Virtual Environments in Brain Damage Rehabilitation: A Rationale from Basic Neuroscience, pages 233–242. IOS Press, 1998.

[Raj13] Joshua Samuel Rajkumar. Enriched environment: Need in medical setup. *BMJ Clinical Research*, March 2013.

[RBR05] F. David Rose, Barbara M. Brooks, and Albert A. Rizzo. Virtual reality in brain damage rehabilitation: Review. *CyberPsychology and Behavior*, 8(3):241–262, 2005.

[RDS+10] Patrick Renner, Timo Dankert, Dorothe Schneider, Nikita Mattar, and Thies Pfeiffer. Navigating and selecting in the virtual supermarket: Review and update of classic interaction techniques. In *Virtuelle und Erweiterte Realität – 7. Workshop der GI-Fachgruppe VR/AR*, pages 71–82, 2010.

[RE98] Martin Raubal and Max J. Egenhofer. Comparing the complexity of wayfinding tasks in built environments. *Environment and Planning B*, 25:895–914, 1998.

[Rey41] André Rey. Lexamen psychologique dans les cas dencéphalopathie traumatique. *Archives de Psychologie*, 28(3):215–285, 1941.

[RFL07] Felix Rabe, Christian Fröhlich, and Marc Erich Latoschik. Low-cost image generation for immersive multi-screen environments. In *Virtuelle und Erweiterte Realität – 4. Workshop der GI-Fachgruppe VR/AR.*, pages 65–76, 2007.

[Riv03] Guiseppe Riva. Applications of virtual environments in medicine. *Methods of information in Medicine*, 42(5):524–534, 2003.

[Riv05] Guiseppe Riva. Virtual reality in psychotherapy: Review. *CyberPsychology and Behavior*, 8(3):220–230, 2005.

[RJ13] Ann Marie Roetke and Sara Jaffee. Results of a randomized controlled trial: The effects of superbetter on depression, July 2013. The University of Pennsylvania in collaboration with SuperBetter Labs, LLC.

[RJR+15] Ann Marie Roepke, Sara R. Jaffee, Olivia M. Riffle, Jane McGonigal, Rose Broome, and Bez Maxwell. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. *Games for Health Journal*, 4(3):235–246, 2015.

[RKW01] Sharif Razzaque, Zachariah Kohn, and Mary C. Whitton. Redirected Walking. In *Eurographics 2001 - Short Presentations*, pages 105–106. Eurographics Association, 2001.

[RL09] Roy A. Ruddle and Simon Lessels. The benefits of using a walking interface to navigate virtual environments. *ACM Trans. Comput.-Hum. Interact.*, 16(1):5:1–5:18, 2009.

[RSPA+06] Bernhard E. Riecke, Jörg Schulte-Pelkum, Marios N. Avraamides, Markus Von Der Heyde, and Heinrich H. Bülthoff. Cognitive factors can influence self-motion perception (vection) in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 3(3):194–216, 2006.

[RWB98] Alberto A. Rizzo, Mark Wiederhold, and J. Galen Buckwalter. *Virtual Environments in Clinical Psychology and Neurocience*, chapter Basic Issues in the Use of Virtual Environments for Mental Health Applications, pages 21–42. IOS Press, 1998.

[SB11] Anna Spagnolli and Cheryl Campanella Bracken. Do you feel as if you are there? Measuring presence in cybertherapy. In *Proceedings of the International Society for Presence Research Annual Conference (ISPR)*, 2011.

[Sch08] Stefan Schlechtweg. miniCAVE – a fully immersive display system using consumer hardware. In *Proceedings of the 14th Eurographics Symposium on Virtual Environments*, pages 73–80, 2008.

[She08] Jesse Shell. *The Art of Gamedesign*. Morgan Kaufmann Publishers, 2008.

[SKHR01] A. Fleming Seay, David M. Krum, Larry Hodges, and William Ribarsky. Simulator sickness and presence in a high fov virtual environment. In *Virtual Reality, 2001. Proceedings. IEEE*, pages 299–300. IEEE, 2001.

[Sla99] Mel Slater. Measuring presence: A response to the witmer and singer presence questionaire. *Presence Teleoperators and Virtual Environments*, 8(5):560–565, 1999.

[Sla03] Mel Slater. A note on presence terminology. *Presence Connect*, 3(3), 2003.

[SS00] Mel Slater and Anthony Steed. A virtual presence counter. *Presence: Teleoperators and Virtual Environments*, 9(5):413 – 434, 2000.

[STM+10] Gustavo Saposnik, Robert Teasell, Muhammad Mamdani, Judith Hall, William McIlroy, Donna Cheung, Kevin E. Thorpe, Leonardo G. Cohen, and Mark Bayley. Effectiveness of virtual reality using Wii gaming technology in stroke rehabilitation: A pilot randomized clinical trial and proof of principle. *Stroke Journal of the American Heart Association*, 41(7):1477–1484, 2010.

[SUS94] Mel Slater, Martin Usoh, and Anthony Steed. Depth of presence in virtual environments. *Presence*, 3(2):130–144, 1994.

[SUS95] Mel Slater, Martin Usoh, and Anthony Steed. Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 2(3):201–219, 1995.

[Tay69] Laughlin B. Taylor. Localization of cerebral lesions by psychological testing. *Clinical Neurosurgery*, 16:269–287, 1969.

[UCAS00] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments*, 9(5):497–503, 2000.

[Uni11] East Carolina University. Study: Casual video games demonstrate ability to reduce depression and anxiety. `www.ecu.edu/cs-admin/news/newsstory.cfm?id=1906`, February 2011. press release.

[UWM+13] Lena Ulm, Dorota Wohlrapp, Marcus Meinzer, Robert Steinicke, Alexej Schatz, Petra Denzlerand Juliane Klehmet, Christian Dohle, Michael Niedeggen, Andreas Meisel, and York Winter. A circle-monitor for computerised assessment of visual neglect in peripersonal space. *PLoS One*, 8(12):e82892, 2013.

[VK78] Steven. G. Vandenberg and Allan. R. Kruse. Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47(2):599–604, 1978.

[VM98] Francesco Vincelli and Enrico Molinari. *Virtual Environments in Clinical Psychology and Neuroscience*, chapter Virtual Reality and Imaginative Techniques in Clinical Psychology, pages 67–72. IOS Press, 1998.

[VWGO04] Peter Vorderer, Werner Wirth, Feliz R. Gouveia, and Others. MEC spatial presence questionnaire (MEC-SPQ): Short documentation and instructions for application. *Report to the European Community, Project Presence: MEC (IST-2001-37661)*, 2004.

[Wan96] John P. Wann. Virtual reality environments for rehabilitation of perceptual-motor disorders following stroke. In P. Scharkey, editor, *Proceedings of the European Conference on Disability, Virtual Reality and Associated Technology*, pages 233–238, 1996.

[WHB+07] Werner Wirth, Tilo Hartmann, Saskia Böcking, Peter Vorderer, Christoph Klimmt, Holger Schramm, Timo Saari, Jari Laarni, Niklas Ravaja, Feliz R. Gouveia, Frank Biocca, Ana Sacau, Lutz Jäncke, Thomas Baumgartner, and Petra Jäncke. A process model of the formation of spatial presence experiences. *Media Psychology*, 9(3):493–525, 2007.

[WK98] Bob G. Witmer and Paul B. Kline. Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments*, 7(2):144–167, 1998.

[WNM+06] Betsy Williams, Gayathri Narasimham, Tim P. McNamara, Thomas H. Carr, John J. Rieser, and Bobby Bodenheimer. Updating orientation in large virtual environments using scaled translational gain. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 21–28. ACM, 2006.

[WNR+07] Betsy Williams, Gayathri Narasimham, Bjoern Rump, Timothy P. McNamara, Thomas H. Carr, John Rieser, and Bobby Bodenheimer. Exploring large virtual environments with an hmd when physical space is limited. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, pages 41–48. ACM, 2007.

[Wol02] Barry E. Wolfe. The role of lived experience in self- and relational observation: A commentary on Horowitz. *Journal of Psychotherapy Integration*, 12(2):147–153, 2002.

[Wor16] WorldViz. Vizard. `www.worldviz.com`, February 2016.

[WS98] Bob G. Witmer and Michael J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 1998.

[WSB+07] Susan L. Whitney, Patrick J. Sparto, Sabarish V. Babu, Larry F. Hodges, Joseph M. Furman, and Mark S. Redfern. Perceived anxiety and simulator sickness in a virtual grocery store in persons with and without vestibular dysfunction. In *Annual Review of CyberTherapy and Telemedicine*, volume 5, pages 173–178. Interactive Media Institute, 2007.

[WSV+99] Edward J. Wegman, Jürgen Symanzik, J. Patrick Vandersluis, Qiang Luo, Fernando Camelli, Antoinette Dzubay, Xiaodong Fu, Nkem-Amin Khumbah, Rida E. A. Moustafa, Robert L. Wall, and Ying Zhu. The MiniCAVE—a voice-controlled IPT environment. In *3rd International Immersive Projection Technology Workshop*, pages 10–11, 1999.

[WW11] David Weibel and Bartholomäus Wissmath. Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology*, 2011, 2011. Article ID 282345.

[YLPM05] Sung-Eui Yoon, Peter Lindstrom, Valerio Pascucci, and Dinesh Manocha. Cache-oblivious mesh layouts. *ACM Transactions on Graphics (SIGGRAPH)*, 24(3):886–893, 2005.

[You03] Christine Youngblut. Experience of presence in virtual environments. Technical Report IDA Document D-2960, Institute For Defense Analyses, 2003.