HAL
open science

Visuo-proprioceptive conflicts of the hand for 3D user interaction in Augmented Reality

Andreas Pusch

To cite this version:

Andreas Pusch. Visuo-proprioceptive conflicts of the hand for 3D user interaction in Augmented Reality. Human-Computer Interaction [cs.HC]. Institut National Polytechnique de Grenoble - INPG, 2008. English. NNT : . tel-00347430

HAL Id: tel-00347430

https://theses.hal.science/tel-00347430v1

Submitted on 15 Dec 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INSTITUT POLYTECHNIQUE DE GRENOBLE

N° attribué par la bibliothèque

|__|__|__|__|__|__|__|__|__|__|

THESE

pour obtenir le grade de

DOCTEUR DE L'Institut polytechnique de Grenoble

Spécialité Mathématiques et Informatique

préparée à **l'INRIA Grenoble – Rhône-Alpes et au LIG, équipe i3D**

dans le cadre de **l'Ecole Doctorale MSTII**

presentée et soutenue publiquement

par

Andreas PUSCH

le 16 octobre 2008

VISUO-PROPRIOCEPTIVE CONFLICTS OF THE HAND FOR 3D USER INTERACTION IN AUGMENTED REALITY

DIRECTRICE DE THESE

Sabine COQUILLART

JURY

M. LEON Jean-Claude, Président
M. van LIERE Robert, Rapporteur
M. VERCHER Jean-Louis, Rapporteur
M. ARNALDI Bruno, Examinateur
Mme. COQUILLART Sabine, Directrice de Thèse

ABSTRACT

This thesis explores potentials of applying spatial visuo-proprioceptive conflicts of the real hand to 3D user interaction in Augmented Reality. A generic framework is proposed which can generate, manage and reduce sensory conflicts at hand level while providing a continuous interaction cycle. Technically, the system is based on a video see-through head-mounted display that allows for embedding the real hand into a virtual scene and to visually manipulate its position in 3D.

Two novel methods are introduced on top of this basis: an intuitive virtual object touching paradigm and a hand-displacement-based active pseudo-haptics technique. Both approaches are studied with respect to their benefits, limitations, effects on the behaviour of the user and consequences for the design of Virtual Environments. It is demonstrated that new forms of human-computer interaction are possible exploiting the described visuomotor conflicts of the hand. Promising future perspectives are presented.

ABRÉGÉ

Cette thèse concerne l'étude d'un conflit visuo-proprioceptif de la main appliqué à l'interaction 3D. Un cadre de travail générique est proposé afin de générer, contrôler et réduire le conflit sensoriel en cours d'interaction. Le système utilise un visiocasque semi-transparent vidéo permettant l'intégration de l'image de la main réelle dans la scène virtuelle ainsi que la manipulation visuelle de sa position 3D.

Deux nouvelles méthodes sont introduites: un paradigme d'interaction intuitif pour toucher des objets virtuels et une technique pseudo-haptique active basée sur le déplacement de la main. Ces méthodes sont étudiées en considérant leurs bénéfices, limitations, effets sur le comportement de l'utilisateur et les conséquences sur la conception d'applications en environnements virtuels. Ces travaux montre que de nouvelles formes d'interaction 3D sont possible en exploitant les conflits visuo-proprioceptifs de la main. Des perspectives prometteuses sont présentées.

TABLE OF CONTENTS

1 Introduction....9

1.1 General motivation....9

1.2 Thematic emphases....10

2 Theoretical foundations and technical backgrounds....15

2.1 Visuo-proprioceptive conflicts and multisensory processing....16

2.2 Within hand's reach interaction....19

2.3 Augmented and Virtual Reality technologies....23

3 Visuo-proprioceptive conflict generation and management....27

3.1 Introduction....27

3.2 System requirements....28

3.3 Existing system infrastructure....30

3.3.1 Hardware platform....30

3.3.2 Software platform....31

3.4 VPC framework....35

3.4.1 Hand texture carrier object repositioning....35

3.4.2 Other system enhancements....38

3.5 Proof of concept....39

3.6 Discussion....40

3.7 Conclusion and future work....41

4 Intuitive control for touching virtual surfaces....43

4.1 Introduction....43

4.2 Related work....44

4.3 Interaction paradigm fundamentals....45

4.4 Hypotheses....47

4.5 Experiment....48

4.5.1 Subjects....48

4.5.2 Factorial design....49

4.5.3 Procedure....50
4.5.3.1 Pointing task....50
4.5.3.2 Questionnaire....52
4.5.4 Data acquisition and analysis....53
4.6 Results....54
4.6.1 Coarse hand oscillation around the target....54
4.6.2 Target entering depth....55
4.6.3 Hand movement duration....56
4.6.4 Hand trajectory length....57
4.6.5 Hand representation evaluation....57
4.7 Discussion of the experiment....60
4.8 Conclusion and future work....62

5 Hand-displacement-based active pseudo-haptics....65
5.1 Introduction....65
5.2 Related work....66
5.3 Technical simulation fundamentals....69
5.4 Force field illusion and response model....70
5.4.1 Illusion....70
5.4.2 Response model....72
5.4.2.1 Perceptual constraints....72
5.4.2.2 Device constraints....73
5.4.2.3 Force field properties....74
5.4.2.4 Adaptation to hand movements....76
5.5 Hypotheses....78
5.6 Experiment....79
5.6.1 Subjects....79
5.6.2 Factorial design....79
5.6.3 Procedure....80
5.6.3.1 Force field strength comparison task....81
5.6.3.2 Illusion evaluation questionnaire....82
5.6.4 Data acquisition and analysis....83

5.7 Results....85
5.7.1 Pectoralis major activity....86
5.7.2 Overall condition ranking....86
5.7.3 Force combination differences....87
5.7.4 Force combination zones....88
5.7.5 Force combination senses....89
5.7.6 Subjective illusion evaluation....89
5.8 Discussion of the experiment....91
5.9 Conclusion and future work....93

6 Visual-to-proprioceptive hand feedback convergence....97
6.1 Introduction....97
6.2 Related work....99
6.3 System integration fundamentals....100
6.4 Hand feedback convergence model....101
6.5 Hypotheses....105
6.6 Experiment....106
6.6.1 Subjects....106
6.6.2 Factorial design....107
6.6.3 Procedure....109
6.6.3.1 Interaction task....110
6.6.3.2 Questionnaire....111
6.6.4 Data acquisition and analysis....112
6.7 Results....113
6.7.1 Qualitative questionnaire results....113
6.7.2 Global method comparison....115
6.7.3 Task completion time regression analysis....116
6.7.4 Target-characteristics-based method assessment....117
6.8 Discussion of the experiment....119
6.9 Conclusion and future work....122

7 General conclusion....125

8 Future perspectives.....129
8.1 Online hand trajectory manipulations.....129
8.2 Beyond positional visuo-proprioceptive conflicts of the hand.....131

List of figures.....135

List of tables.....139

List of abbreviations.....141

References.....143

Appendices.....149
A.1 Questionnaire of the experiment of Chapter 4.....149
A.2 Questionnaire of the experiment of Chapter 5.....152
A.3 Questionnaire of the experiment of Chapter 6.....153

1 Introduction

In this introductory chapter, an orienting overview of the thesis is given by presenting its general motivation (see Section 1.1), thematic emphases and the overall structure (for the last two, see Section 1.2).

1.1 General motivation

The present work is motivated by the idea that exploiting sensory conflicts between vision and proprioception of the user's real hand in space (see Fig. 1-1) can lead to novel approaches of human-computer interaction (HCI), specially focusing on 3D user interfaces (3D UI) in Augmented Reality (AR).

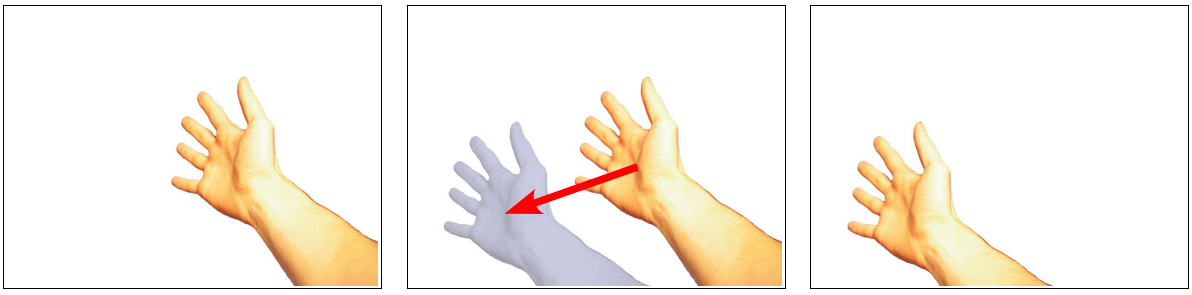


Figure 1-1: Basic hand displacement principle

This motivation was mainly stimulated by the following three points. First, it has been demonstrated that visual manipulations of the user's actions can add value to 3D UI in different contexts, for instance, in the area of haptic illusions (e.g. pseudo-haptics, see also Chapter 5). Second, video see-through (VST) head-mounted displays (HMD) provide the technical basis for spatially inducing and controlling visuo-proprioceptive conflicts (VPC), even of real limbs. Novel interaction techniques which make use of a static or dynamic visual hand repositioning can thus easily be explored. Third, our knowledge on the processing, combination and integration of multisensory inputs as well as on how perception and

behaviour are or can be affected rapidly grows. Considering these aspects in the design and the development of 3D UI may allow for novel forms of applications and hopefully opens a fruitful multidisciplinary view on often still isolated research domains.

1.2 Thematic emphases

The main questions addressed throughout this work are defined below. For a review of the relevant theoretical and technical backgrounds, refer to Chapter 2. All specific fundamentals (see Sections 4.2, 5.2 and 6.2) and hypotheses (see Sections 4.4, 5.5 and 6.5) will be discussed within the scope of each particular chapter.

The thesis rests on four principal axes:

1. *Visuo-proprioceptive conflict generation and management (see Chapter 3).*

   In order to create and flexibly control VPC of the user's hand in space, a software framework is developed on top of the existing laboratory infrastructure. The latter consists of a distributed scene graph real time rendering and interaction system driving a VST-HMD AR setup. Conceptual requirements for the VPC framework are derived from the intended visuomotor manipulation goals and then translated into a common functional basis. Additional system enhancements will be described, too.

2. *Intuitive control for touching virtual surfaces (see Chapter 4).*

   Several known concepts for a classical pointing-like interaction are integrated and merged with a spatially manipulable feedback of the real hand. The main purpose is to develop a generic and more intuitive virtual object or surface touching paradigm without haptic devices. Intuitive means that the user should benefit both subjectively and objectively from a control similar to handling his “every day life interaction tool”. It is meant to facilitate even spatially constrained selection and manipulation tasks in Virtual Environments (VE). The proposed paradigm aims at an interaction within hand's reach (see Fig. 1-2).

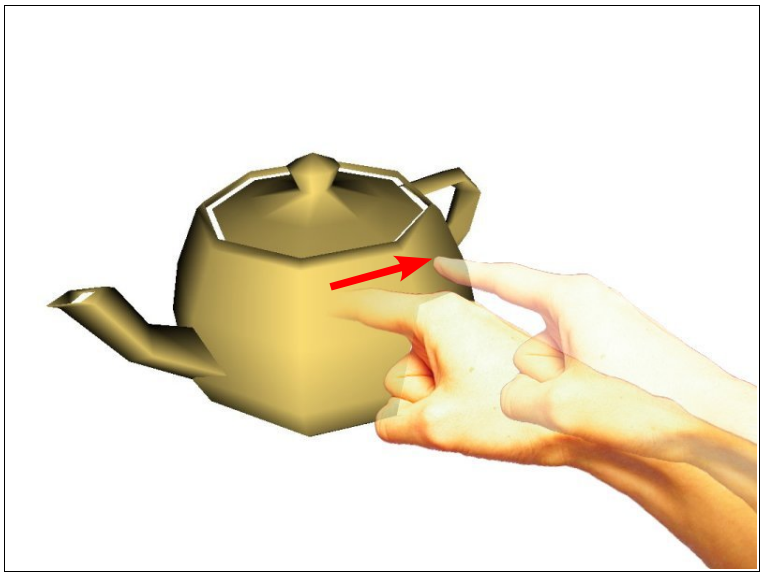


Figure 1-2: Compositing of a user touching a virtual object, with the visual hand constrained on the object's surface and the real hand entering it.

3. *Hand-displacement-based active pseudo-haptics (see Chapter 5).*

   Active haptic systems are known for a long time and can supply VE with convincing haptic feedback. But such devices do often suffer from a number of limitations (e.g. constrained interaction space, actual employment difficulties, expensive maintenance) frequently addressed by recent research.

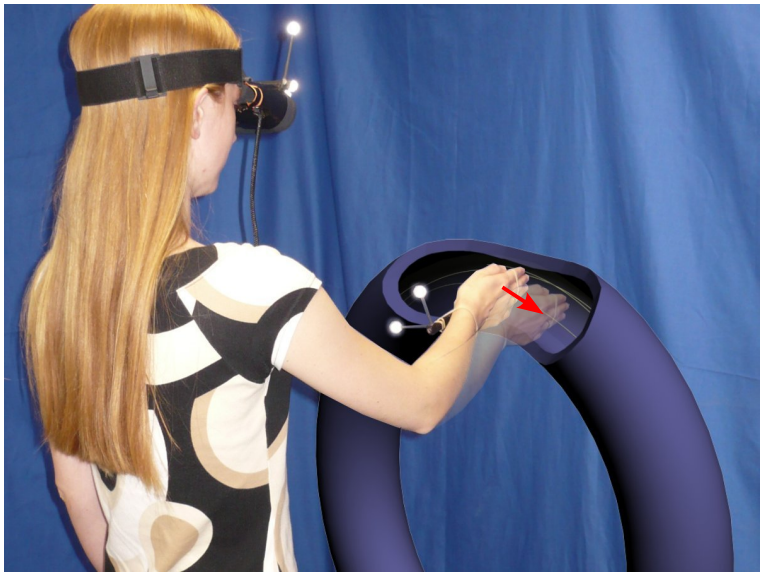


Figure 1-3: Compositing of a user reaching into a virtual force field, including the visualised hand displacement to the right.

When looking at boundaries and capabilities of the human sensory system, it seems to be worthwhile investigating alternative approaches to active haptics. A controlled, event-based visual hand displacement in conjunction with a certain motor activity triggered at arm / hand level could induce a haptic-like or pseudo-haptic impression. Neither active nor passive haptic devices are required for the proposed force field application (see Fig. 1-3).

4. *Visual-to-proprioceptive hand feedback convergence (see Chapter 6).*

Beyond the novel interaction techniques covered by the previously referenced chapters, one can envisage a lot more cases in which spatial VPC at hand level could be useful or even necessary to evoke. For multifarious perceptual, technical and methodological reasons, each visuomotor discrepancy should again be reduced, whenever possible. The most important prerequisite for this treatment is to guarantee an unperturbed and continuous interaction. Several aspects are taken into account to perform a fast, but unnoticeable hand feedback convergence (HFC) until the formerly disconnected hand representations are spatially aligned (see Fig. 1-4).

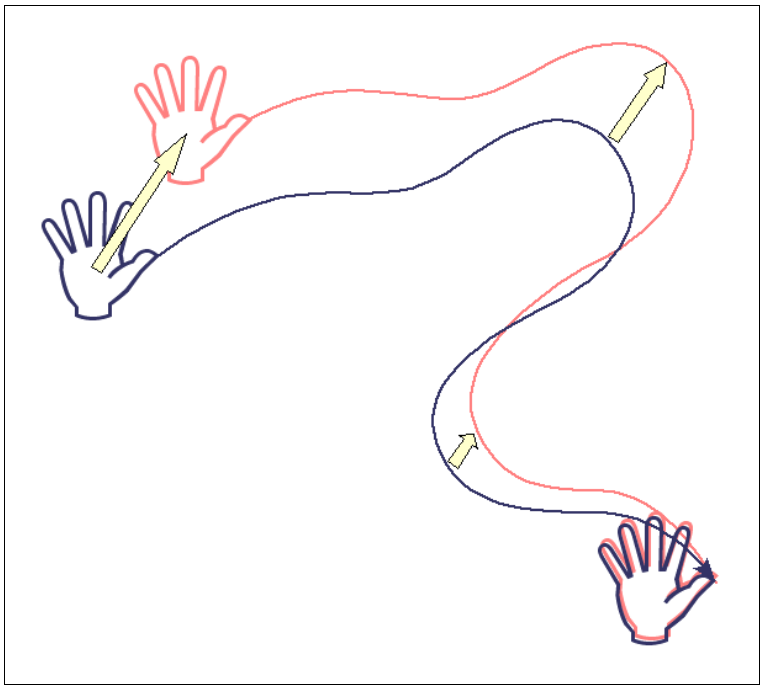


Figure 1-4: Gradual convergence effect until spatial alignment
(dark blue: real hand, salmon: visual hand).

After the general conclusion (see Chapter 7), this thesis will be closed with a discussion of diverse future work directions (see Chapter 8). In this regard, emphasis is not only put on 3D UI, but also on a wider range of disciplines in human and medical sciences.

A specific investigation of the underlying brain processes and their consequences on the user's perception and behaviour are not in the scope of this work. However, a few links are established in order to stimulate a broader, multidisciplinary view on the topic.

2 THEORETICAL FOUNDATIONS AND TECHNICAL BACKGROUNDS

The purpose of this chapter is to provide common foundations upon which all concrete topics of this thesis will build or which they will enhance. For specific related work, please refer to the respective chapters.

Important for the development of visuo-proprioceptive-conflict-based interaction techniques is an understanding of the main principles and mechanisms involved in the provoked sensory conflicts (see Section 2.1). Not only because stretching manipulations of the sensory supply beyond certain limits may unnecessarily stress the user's perception and thus, at some point, the user himself. Also, knowing about manipulation potentials would eventually allow for more efficient and richer interaction methods.

All approaches presented in this work rely on at least roughly aligned visually perceived and kinaesthetically occupied spaces. 3D UI techniques which have proved beneficial for such kind of co-located interaction within hand's reach and several insightful neuroimaging results will be discussed in Section 2.2.

From a technical viewpoint, Augmented and Virtual Reality (AR and VR) technologies have to be assessed with respect to their applicability and practicability (see Section 2.3). To recall the overall system-side objective: visuo-proprioceptive conflicts (VPC) need to be managed flexibly in space while preserving the visual appearance of the user's real hand.

2.1 Visuo-proprioceptive conflicts and multisensory processing

Designed for a multimodal world, the human sensory system acquires information about the environment through various senses (e.g. vision, touch and hearing). As a complementary source to this external data, the body's internal (articular) motion and position state (i.e. proprioception) is considered, too. All these sensory signals feed a complex processing chain in the central nervous system (CNS) steadily enabling perception as well as the execution of voluntary and involuntary motor (re-)actions at limb and body posture level.

It is known for a long time that vision has typically a strong influence on both perception and action. It serves, amongst others, building mental representations of the environment, navigating through and interacting with them. Guiding goal-directed movements is one of the most frequent tasks for which vision is essential. After a target has been located in the egocentric frame of reference, motor plans (i.e. muscle activation patterns) are generated to transport the limb towards this target. Most of the corrections during voluntary or involuntary movements (e.g. pointing actions or balance control, resp.) are supported or even dependent on vision. For this reason, if vision is absent, initial errors in the articulation estimation will be propagated to the movement execution [Sc05]. [Wi02]

There is often an extensive interaction between vision and other sensory modalities, in particular proprioception. The origin of proprioceptive signals are muscle and joint receptors of articular structures and the circular canals and otolithic organs of the vestibular system. These mechanoreceptors transduce mechanical deformations of special tissue into frequency modulated neural signals. Once transmitted to the CNS, this sensory information is used to compute the whole body posture and the spatial limb states based on forward kinematics of joint angles. At this point, proprioception directly contributes to physical actions. Because the knowledge about the system's current state allows for the creation of actual motor commands. Adequate muscle activity is triggered in order to achieve the desired goals. Limb motion and positioning, postural stability or more complex movements such as running are controlled in this manner. [Le00]

To produce a coherent and stable percept, the brain merges the sensory feedback about environmental properties using different strategies [Er04]. Two central mechanisms have been identified: sensory combination and sensory integration. In the former, the CNS tries to

resolve perceptual ambiguities, a method also referred to as disambiguation, by maximising "information delivered from the different sensory modalities". The utilisation of depth cues (e.g. shadows, object occlusion or optical flow) is an example for this strategy. Sensory integration aims at the reduction of "variance in the sensory estimate to increase its reliability". That is, the brain attempts to obtain a more robust sensation by applying weights to the senses which contribute to the particular percept [Er02]. A Maximum Likelihood Estimate (MLE) model was proposed to predict the relation between the visual and the haptic modality, also depicted by the simplified psychometric function in Figure 2-1. Noise or uncertainties in one of the modalities (e.g. expressed by the distribution of nervous excitation patterns) would reduce its respective integration weight.

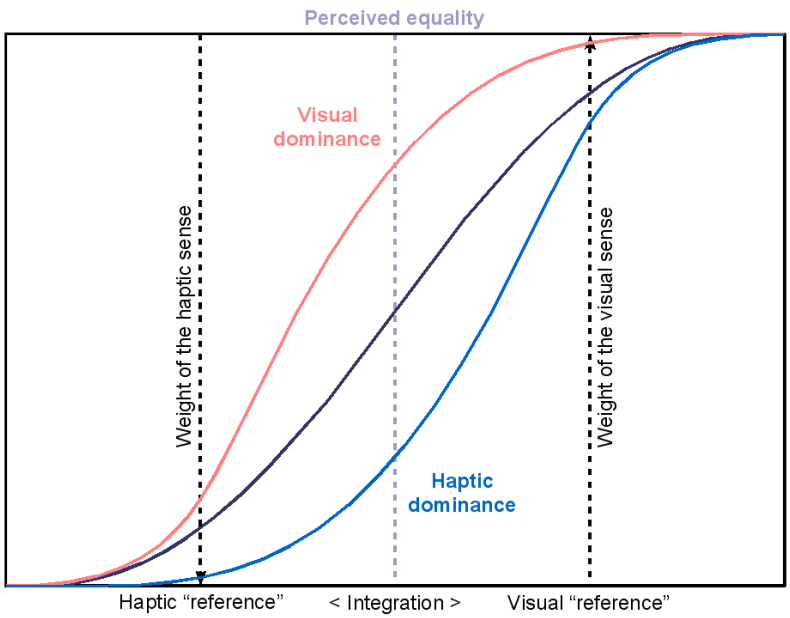


Figure 2-1: Psychometric functions for different cases of visuo-haptic sensory integration.

An artificial decoupling of a limb's visual and proprioceptive spatial appearance (i.e. VPC), whether in motion or position, would lead to such a weighted integration of the conflicting sensory inputs. Depending on the weight assigned to each modality, the one receiving the stronger weight is likely to dominate the other. In the literature, the notions visual dominance or visual capture are used to indicate that the final percept is mainly influenced by vision. One of the earliest reported examples of this phenomenon, although not based on a spatial sensory

divergence, is Charpentier's size-weight-illusion (1891) [Mu99]. Objects of equal mass were presented at different visual sizes. When compared in their weights, the smaller object was found to be heavier. The main reason for this is that a larger visual appearance implies a stronger load, because of the mass-volume proportionality: $m = \rho \cdot V$. But if the actual haptic sensation does not match the anticipated greater weight, then a cognitive interpretation of the integrated senses would rather suggest the object to be lighter. This means that vision has successfully altered the judgement in this visuomotor task. But it also of great

Various static and dynamic perceptual manipulations for the purpose of 3D UI seem to be possible by exploiting these fundamental mechanisms. Due to flexibility of brain processes, a carefully biased perception can detach motor control from known environmental constraints. Sensorimotor representations adapt to artificially induced VPC, resulting in visuomotor skill acquisition or perceptual recalibration [We08]. The first adaptation type is assumed to take place in larger-scale, multidimensional VPC and involves higher level cognitive processes as learning. It does usually not cause after-effects known to occur in perceptual recalibration (e.g. pointing errors). This second type of adaptation is more likely to be triggered on smaller and lower-dimensional VPC (e.g. lateral visual hand shifts). Dual adaptation, an alternation of adaptation and re-adaptation to normal or even other conflicting conditions, has the potential to reduce undesirable (after-)effects. Research in 3D UI could demonstrate a reconfiguration of visuomotor coordination for a gradual VPC at hand level [Bu05]. Subjects wore an opaque HMD and had to perform a sequential pointing task in a game-like scenario. Head and hand were tracked. While moving the real hand towards the target panels, an increasing drift of the visual hand avatar in the vertical plane (i.e. left, right, up or down) was introduced at a supposedly imperceptible velocity of 0.458 degrees/s. Not being informed about this manipulation, several subjects showed pointing deviations of up to 60 degrees without noticing them. The mean drift extent was found to be approximately 40 cm. When, after the first trial, subjects were adverted to the visual treatment, the mean detection threshold dropped to 20 cm. It has also been remarked that, in the absence of any intended hand movement, visual drifts are detected much earlier. Thus, attention may play a key role for the final sensory integration weighting, as the identification with the limb representation does (see next section).

Vision is often the predominant modality in VPC. But there are special cases in which proprioception gains in importance. In [Sn06], for instance, direction-dependent reaching errors were investigated under the influence of the mirror-illusion. At the hand movement starting position, subjects saw a visual substitute of their right hand (i.e. a mirrored image of either the left hand or a wooden block). The right hand itself was placed at varying locations behind the mirror. Pointing towards associated targets was done without visual feedback of the reaching limb. What Snijders et al. (2006) found is a "direction-dependent weighting, with vision relatively more dominant in the azimuthal direction (i.e. left- and rightwards, author's note), and proprioception relatively stronger in the radial direction (i.e. in depth, author's note)". Also, a significant interaction between the target location and the hand feedback at the movement starting point was observed, "suggesting stronger visual-proprioceptive integration from viewing the hand than the block of wood".

Summary

To conclude, sensory integration and perception, as they take place in static or dynamic spatial VPC, appear to be influenced at least by signal reliability, adaptation and attention. The first emanates from a number of environmental and modality-based factors. According to the variance in the estimate, signal reliability directly affects the weights assigned to each of the sensory input channels. That is, whatever raises variance may reduce the final integration weight. A conditional gradual dominance shift between the modalities can hence be expected. The second factor, adaptation, would again readjust this shift. Attention was rarely studied so far within this particular research context. However, evidence for a top-down impact of the attentional focus on sensory integration does already exist. Whether there are more higher level functions interacting with the addressed perceptual mechanisms and to which extent these different layers can influence each other are just a few of the important open questions.

2.2 Within hand's reach interaction

Previous research in 3D UI has isolated some prerequisites for an efficient interaction within hand's reach, among them co-location. This notion refers to the exact spatial alignment of both the visual and the proprioceptive appearance of an interacting limb. Co-location has been

studied in several ways. For instance, in [Mi97], subjects had to complete a near space docking task in which either a direct (i.e. no visual offset) or a distant control at two different levels (i.e. static or linearly increasing offsets) was imposed. Analysis of the task completion time revealed a significant performance benefit in the case of co-location while the two offset conditions did not differ from another. Further, the direct control was subjectively preferred suggesting that it provides a more convincing or intuitive feeling for the manipulation.

The relevance of the interaction distance has also been investigated for a 3D location task [Pa02] on the Responsive Workbench™ [Kr95]. It was tested whether close control (i.e. no or a 20 cm manipulation distance) would result in a better performance than acting at a distance of 40 or 55 cm. There was no significant difference found, neither between 0 and 20 cm nor between 40 and 55 cm. A significant effect of the manipulation distance was only observed for the close control compared to its farther variants. In near space, user actions seem to be tolerant at least against static spatial offsets. Displaying the hand kinematics at a scale of 1.5 have led to significantly worse task completion times. Interaction performance deficits as a function of VPC were reported by Burns et al. (2005), too.

In the remainder of this thesis, the term co-located space will be used in the sense of the close distance presented above. It is therefore understood as the space which is usually situated around the lower segments of the upper extremities (i.e. around forearm and hand, see Fig. 2-2).

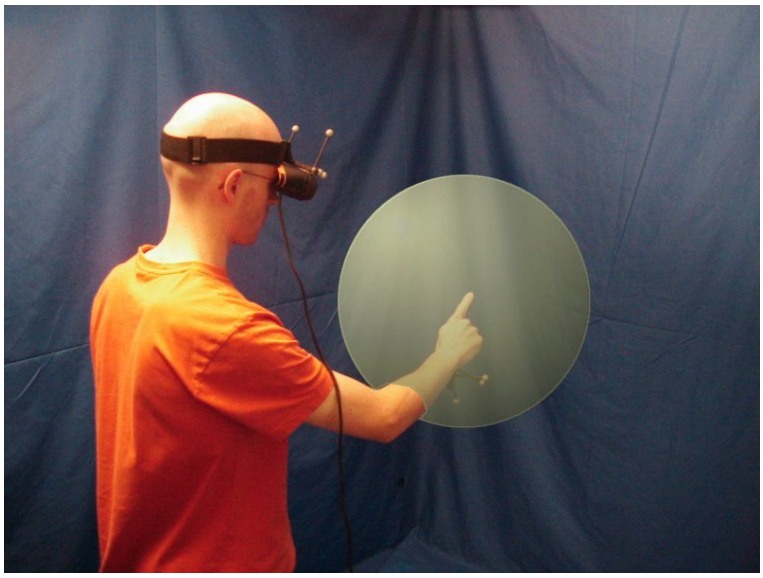


Figure 2-2: Co-located space as the spheric tolerance surrounding "perfect co-location" (i.e. tip of the real index finger).

Beside profiting from a reduced transformation load on the sensorimotor system due to co-location and thus a more accustomed access to virtual objects, another principle has been demonstrated to deliver a more natural experience for an interaction within hand's reach. While it is inherent in nonocclusive VE (i.e. VR systems allowing for a direct view onto the own body), the fact of seeing the user's hand has also been adopted to occlusive VE (i.e. systems without such direct view). Representations range from static figurative pointers over dynamic articulated hand objects to captured video. The extent to which different visual fidelity levels of the so-called self-avatar of the hand can affect performance in a spatial cognitive task was asked in [Lo03]. Specially textured blocks had to be arranged in order to match given patterns. First, in a real world scenario, the reference performance was measured. After, participants had to repeat the same task either in a purely virtual setting (i.e. blocks and hands as virtual objects) or in a hybrid scenario with a generic self-avatar of the hands (i.e. with unicolour rubber gloves worn and the hands registered as video) or in a visually faithful hybrid environment (i.e. embedded video of the bare hands). In the hybrid conditions, visually reconstructed real cubes were mixed with the respective video hand feedback. It was found that handling real objects significantly improves the task performance whereas visual fidelity has only a limited impact. That is, between the two hybrid cases, there was only a slight advantage for the visually faithful hand representation. But the realistic hand feedback was again subjectively preferred.

In recent neuroimaging experiments, specific brain correlates have been identified to be involved in watching hand grasping actions [Pe01]. Subjects were passively observing gestures made by a real hand in real space, two 3D hand models of a low and high level of realism in VR and a real hand on a 2D TV screen. Analysis of positron emission tomography (PET) data showed “different functional correlates for perceiving actions performed by a real hand in a real environment, in comparison with 3D Virtual Reality and 2D TV hand motor sequences”. However, common activities were observed in areas mainly devoted to motor planning, the perceptual and cognitive representation of action and the recognition of “biologically plausible motion”, for instance (see furthermore [Ve06]). These results are of interest for this thesis, even though passive observations may not necessarily engage exactly the same neural networks as action. Several potential system attributes can be derived (e.g. spatial visual feedback and control, apparent motion coherence and quality) likely to reinforce reliability of manipulated or discrepant sensory information about the actions performed.

Researchers have identified a “feeling of ownership of a limb” (for a review, see [Bl03]) and attempted to locate it in the premotor cortex [Eh04]. Such a sensation is considered to be fundamental for bodily self-consciousness and attribution, providing perceptual reliance: a strong perceptual link to the corresponding limb. Perceived actions are understood as self-generated and thus easier trusted, if they match predicted sensory feedback patterns (i.e. the internal model of action [Wo95a]) or appear otherwise faithful. Also, functional magnetic resonance imaging (fMRI) was used to detect variations in brain activity while subjects were presented with a systematically altered rubber hand illusion [Bo98]. To evoke this illusion, itself an example for visual dominance, the real hand has first to be hidden, but replaced by a co-aligned realistic rubber hand. Second, both hands have to receive simultaneous brush strokes (i.e. the real one tactilely and the rubber hand visually). After a short time, subjects typically develop the experience that the rubber hand belongs to them. By spatially and temporally modifying the illusion parameters (i.e. 2 x 2 factorial design, spatial component: rubber hand visually aligned or 180 degrees turned towards the subject, temporal component: synchronous or asynchronous brushing), neuronal responses in the premotor cortex are expected to change accordingly. The effect on the activity was found to be significantly stronger in the aligned synchronous condition compared to all others. Further, proprioceptive recalibration of the upper limb, probably a “key mechanism for the elicitation of the illusion”, took place reflected by significantly lower activity after the illusion onset. Multisensory integration in a body-centred frame of reference is proposed to be the source of self-attribution.

Summary

After all, manipulations of the sensory inputs accompanying an interaction within hand's reach would clearly gain effectiveness from a direct, co-located limb control, because natural sensorimotor processes remain widely untouched. User preference and neuroimaging results as well as widely accepted theoretical models which describe the phenomenon of recognising own actions suggest that both a more lifelike visual hand representation and motion feedback may support acquiring, but also maintaining some advantages of the natural behaviour (e.g. trusting the interacting limb). How far the different methodological and theoretical aspects discussed can be afforded from a technological perspective will be topic of the next section.

2.3 Augmented and Virtual Reality technologies

Interactive immersive VE are systems designed to incorporate the user into an artificial world by providing him with multimodal computer-generated feedback. VR is usually referred to as largely synthetic. Real components do only occur, if they are integral parts of the scenery (e.g. the user himself or some special devices) [Bu03]. At the other end of the system continuum (see Fig. 2-3) stands AR or Mixed Reality (MR) [Oh99]. In these systems, real and virtual elements are combined to varying degrees, ranging from computer-generated real world overlays to an integration of real objects into an otherwise virtual world, also called Augmented Virtuality (AV) [Mi94].

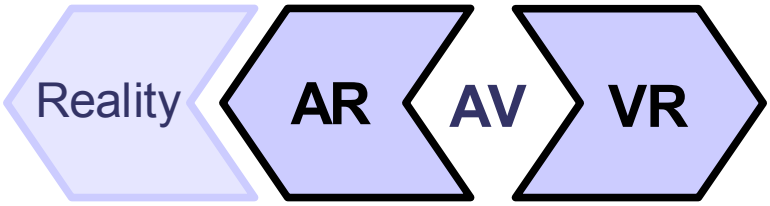


Figure 2-3: AR-VR system continuum (see Milgram et al., 1994).

Common in interactive immersive VE is the need to register or track the user's state, including, for instance, his head and hand positions, gestures, speech, facial expressions or even biosignals. This data can then be used to update the simulation state. Responses are computed in real time and displayed through their respective feedback channels (e.g. for a viewing-dependent rendering, controlled object displacements, or an event-based tactile or force feedback). In so doing, the user's reaction to the new situation finally closes a continuous interaction loop.

In the course of this chapter, a number of general prerequisites were identified which are either essential or desirable for an efficient hand-displacement-based interaction. Essential is the ability to create and flexibly modify VPC in space. In a way, this would be possible in classical projection-based systems like the CAVE™ [Cr92] or the Responsive Workbench™ used by Paljic et al. (2002). But there are at least two serious limitations. First, if the real hand is not covered from the user's view, the sensory discrepancy is visible making it hard for the brain to fuse the conflicting inputs. Putting a hand cover or employing some kind of mirror system as it was used by Snijders et al. (2006) may solve this problem at the expense of

convenience and a drastically reduced interaction space. Second, even without visual hand offsets, a co-located interaction would cause perturbing occlusion violations (see Fig. 2-4), when the user tries to reach with his real hand behind virtual objects. Since the hand is physically always in front of the display, it appears in front of closer presented objects, too.

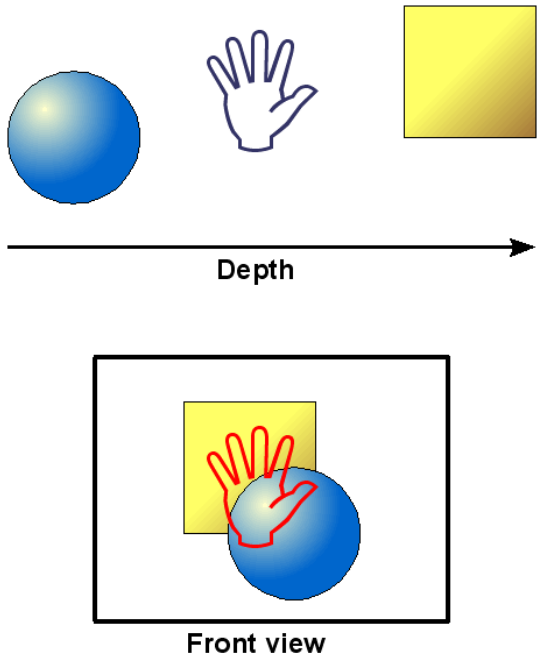


Figure 2-4: Occlusion violation problem,
top: given depth ordering, bottom: incorrect visual
impression with the real hand always in front.

One opportunity to overcome the above-mentioned difficulties would be to use an HMD. The graphical output is directly delivered to the eyes by small displays mounted on a helmet. Without see-through options (i.e. without mutual-occlusion-enabled optical [Ki00] or video see-through [Ed93]), the visual sense of the user is often completely occupied preventing the view of the real environment. Burns et al. (2005) have chosen such an opaque device for their study in which they induced gradual hand shifts. The experimental scene consisted of only virtual objects, the hand included. An initial co-location could thus immediately be achieved by simply showing the virtual hand at the real hand's position. The occlusion violation problem is automatically solved.

To benefit from a perceptual reliance as strong as possible, that is, to assure a wide visuo-proprioceptive manipulation range and impact, a last factor should be considered. Virtual hand avatars mostly appear in a uniform size, shading and a static shape or gesture. Matching all these properties to the actual appearance of the real hand would require complex and expensive techniques (e.g. based on articular motion tracking or 3D reconstruction, resp.). Alternatively, one could use existing video see-through (VST) technology and capture the user's hand as video before embedding it into a virtual scene. This is where terms like AR, MR or, even more appropriate, AV emerge. The only open question is how to control the intended VPC and therefore the spatial visual hand position. A promising starting point which additionally accounts for most of the other prerequisites has been proposed in [Or07]. Figure 2-5 shows the principal approach. For further details, see Chapter 3.

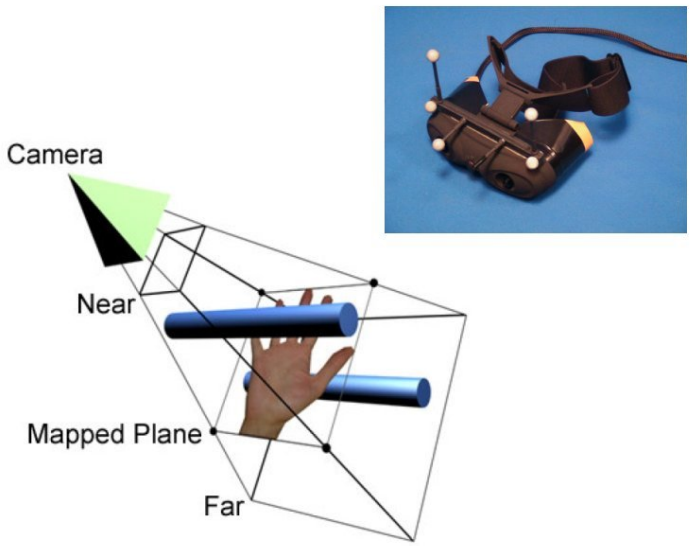


Figure 2-5: Real hand integration approach (see Ortega, 2007) using a VST HMD with built-in cameras (top right).

Summary

VST AR provides the most suitable technological basis for simulating VPC of the hand. It allows for eliminating several hindering drawbacks coming along with other classical VE.

3 VISUO-PROPRIOCEPTIVE CONFLICT GENERATION AND MANAGEMENT

Interaction methods which exploit visuo-proprioceptive conflicts (VPC) at hand level under the conditions described in Chapter 2 can be expected to share a certain hard- and software basis. The current chapter addresses the design and implementation of this common ground.

After the definition of all essential requirements accruing from the VPC manipulation goals (see Section 3.2), the given system infrastructure will be analysed in order to identify potential needs for extensions (see Section 3.3). The resulting VPC framework enabling a flexible control over the spatial position of the user's real hand is developed in Section 3.4. For a proof of concept of this framework, refer to the Sections 3.5 and 3.6. The chapter will finally be closed by pointing out some future methodological and technical improvements (see Section 3.7).

3.1 Introduction

As already mentioned, the primary purpose of this work is to explore novel interaction approaches relying on deliberately conflicted visual and proprioceptive sensory information about the real hand in space. Not only technical achievements allow for this new 3D UI research direction. Also, a better understanding of the involved brain processes and how they can effectively be altered, plays a very important role. The principal factor in this regard is sensory integration influenced by visual dominance (see Section 2.1).

Practically, inducing VPC means to present the user's hand visually at a different location than it really is, either by applying a fixed offset or a continuous displacement. Suchlike may at best be performed using video see-through (VST) AR technology which has the overall advantage of a full control over all displayed elements, including real objects.

The main requirements for a VPC framework are, beside actual hand shifting assignments, first, to offer a generic decoupling interface and second, to remain an encapsulated modular system add-on. In the course of the next sections, these points will be elaborated.

3.2 System requirements

System characteristics which are mandatory for the generation of variable VPC at hand level while delivering a convincing real or quasi-real hand feedback (i.e. a familiar limb substitute immediately reflecting, for instance, gesture changes) are listed hereafter:

1. *Three-dimensional (quasi-)real hand feedback.*

   The user should be presented with co-located, preferably stereoscopic images of his real hand. Because this is presumed to retain the natural sensorimotor processing, amongst others, due to space organisation, so-called biological hand motion feedback and a more natural look (see Section 2.2). Further, the interaction comfort can be ameliorated. To assure a perception as less influenced as possible by other effects than the intended VPC, the user must never see his real hand directly.

2. *VE capable of mixing virtual and spatially manipulable real objects.*

   To combine real and virtual objects (e.g. real hand interacting with a computer-generated 3D scene), there are at least two ways possible. Either virtual objects are placed on top of the real world with occlusion masks applied preventing the rendering of hidden parts. Or real objects are captured and embedded into the virtual scene being shown at a correct spatial ordering with respect to the user's viewpoint. This would automatically avoid occlusion violations. In the first case, artificial real objects shifts are impossible. Consequently, VST AR is the only alternative for the envisaged kind of manipulations (see Section 2.3).

3. *Robust real time background segmentation.*

   As the user's hand shall be seamlessly arranged with virtual objects and eventually undergo a concealed visual repositioning, it needs to be properly extracted from the live image data. A cluttered real environment would complicate the segmentation which may result in visible artifacts (i.e. noisy images) or a decline in the system performance. Quality and speed of the background subtraction algorithm can be improved by simplifying the segmentation problem at two levels. The back of the scenery could easily be coated with unicolour material different enough from skin properties (i.e. green or blue cover). In the standard colour-based subtraction algorithm

itself, an additional assumption about the luminosity distribution could be considered (i.e. background generally darker than foreground).

Once VPC of the hand can be created corresponding to the requirements presented above, the main tasks of the runtime VPC management are as follows:

1. *Visual hand repositioning (i.e. actual VPC production).*

   Independent of the embedding strategy, the underlying structure is supposed to allow for a real time spatial adjustment according to tracking information. That is, even unconstrained co-located actions like hand movements sweeping through a virtual scene have to be reflected by a correct depth placing. In any case, the visual hand is expected to deviate from the real one only translationally (i.e. at three degrees of freedom). The inherent 2D character of the video images makes higher degrees of freedom redundant. A set of elementary VPC control functions would be important, too (e.g. to set fixed 3D offsets, shift vectors and velocities).

2. *Generic real time access to the VPC attributes.*

   The sensory conflict parameters need to be accessible (i.e. reading and writing) in real time to enable arbitrary static and dynamic discrepancies. Moreover, the access should be provided through a generic application interface. Inspired by the classical model view transformation sequence in traditional 3D computer graphics, VPC could also be induced in a local and a global frame. Local means that the applied VPC refer to the user's view whereas global indicates the world reference.

Supplementary components would comprise a VPC supervision and limitation interface as well as a feedback and storage subsystem for manipulation and tracking data. Since these extended functionalities belong to a more comprehensive VPC framework rather than to the essential VPC generation and management requirements, they will be discussed in detail in Section 3.4.

3.3 Existing system infrastructure

In this section, the laboratory's system infrastructure is described as an example of a typical AR / VR framework. The description will be divided into a hardware (see Section 3.3.1) and a software part (see Section 3.3.2). Isolated conceptual and implementation issues are addressed during the VPC framework synthesis (see Section 3.4).

3.3.1 Hardware platform

The available AR setup consists of four subsystems or devices (see Fig. 3-1):

1. VST HMD: opaque head-worn SVGA stereo display with two built-in VGA cameras (see Fig. 2-5, top right).
2. Video acquisition and post processing: image capturing, correction and background segmentation (see Section 3.3.2).
3. Tracking system: infrared optical tracking of mulitple six degrees of freedom bodies (i.e. here: A.R.T.[1]).
4. Compositing system: real time simulation, interaction and rendering (client-server).

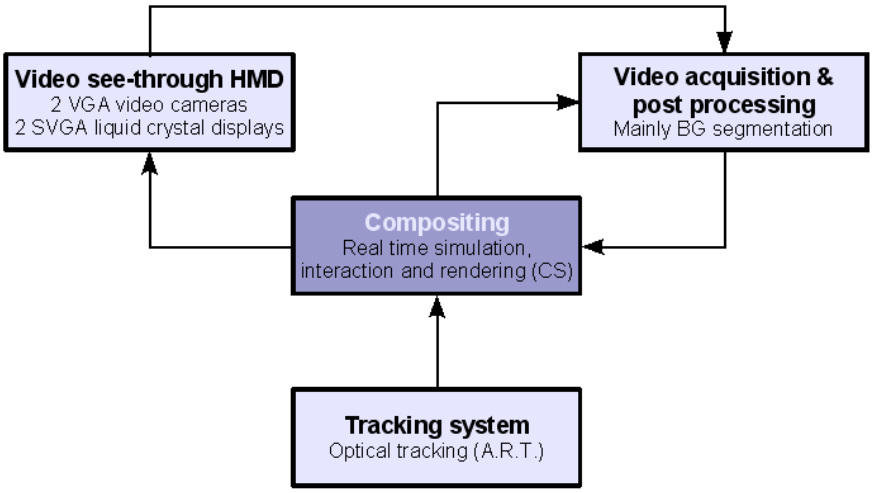


Figure 3-1: Existing system-subsystems scheme.

1 A.R.T.: Advanced realtime tracking GmbH, Germany

The HMD's internal liquid crystal displays (LCD) are connected to the dual-head graphics board of the compositing system. Video data is acquired over USB using proprietary devices and libraries on a separated PC. A VR peripheral network (VRPN) server, running in an A.R.T. machine, provides tracking data to the compositing system. This central unit integrates all information and manages the applications. Each PC is linked to a local gigabit network.

Due to the VST HMD, the most important system requirements are fulfilled. That is, the user's hand is protected from being directly viewed and can theoretically be represented by either its co-located or spatially manipulated visual counterpart. Video of the real environment and thus of the real hand can be captured within the field of view of the built-in cameras. The larger this observation range is the larger could become the VPC. Image processing, including background subtraction, can be performed. Also, a robust tracking of head and hand positions and rotations is possible. The core system collects all data and might render multisensory feedback (e.g. visual and acoustic output). A variety of input devices could finally close an interactive application cycle.

Summary

In conclusion, though device fidelity is limited (e.g. VST HMD: field of view and resolution of cameras and displays, moderate ergonomics), there are no serious problems in the hardware system. However, a larger field of view and a larger stereo overlap of the video cameras would push the effective VPC boundaries [St01].

3.3.2 Software platform

On top of the hardware platform resides a complex software system [Or07]. Each of its components can, more or less, be related to the main blocks shown in Figure 3-1. An overview of the complete software platform is given with Figure 3-2. Isolated parts serve as stand-alone applications.

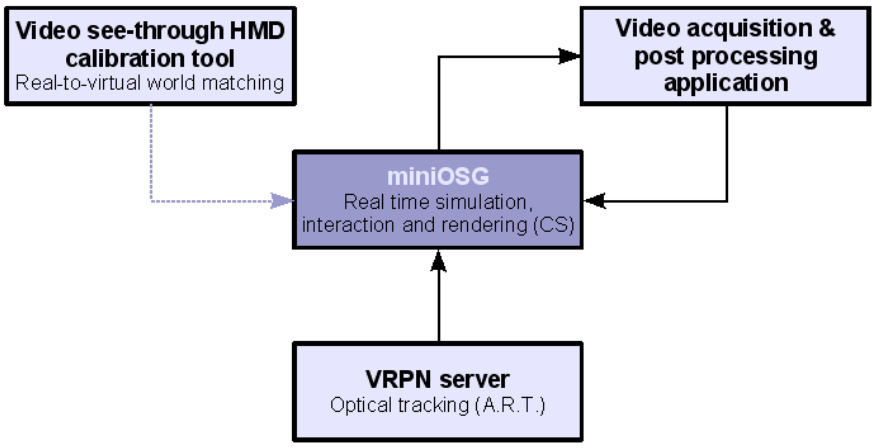


Figure 3-2: Overview of the existing software platform.

The VST HMD calibration tool offers a simple way to solve the real-to-virtual world matching. First, the cameras' intrinsic parameters are determined by utilising the OpenCV [OCV] function `cvCalibrateCamera`. Corners of a known chessboard are detected in multiple snapshots and passed to this method. Each camera has also to be located within the HMD's frame of reference, that is, with respect to the head tracking body attached to the HMD. In order to compute these extrinsic parameters (i.e. their six degrees of freedom spatial properties), the function `cvPOSIT` is called while the coordinates of a tracked known 3D object are used for the inverse pose estimation. As a result, the transformation matrices obtained for each camera can later, during simulation runtime, be applied for an exact superimposition of the spaces viewed by the cameras real and virtual spaces.

Next, the video acquisition and post processing application reads image data from the camera devices. Before transmitting the video feedback to the compositing system, several processing steps take place. A look-up table is generated for the full RGB[2] colour set (i.e. 256^3 elements). For each field of this table, a boolean permission value is stored deciding whether this particular colour has to be considered as background or not. The table is used for the processing of the acquired images. A hue test (i.e. not an RGB test) is performed to create pixel-wise binary background maps which are delivered together with the original images to the central compositing system. There, the images are combined to separate the foreground from the background (see Fig. 3-3). The actual test hue, with a certain tolerance, is determined

2 RGB: Red Green Blue colour scheme

at core system side and transferred to the video acquisition PC. To reduce network traffic, colour images are sent in their initial Bayer format.

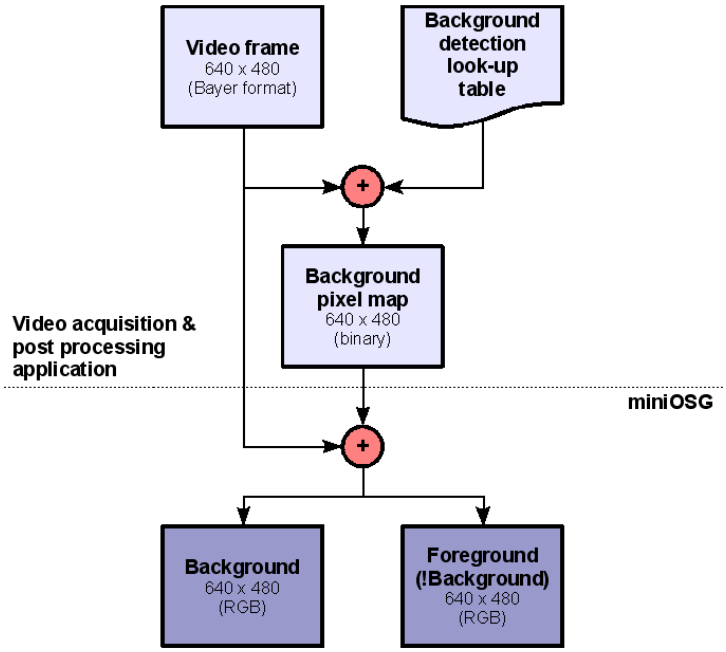


Figure 3-3: Foreground-background generation and separation.

As mentioned in Section 3.3.1, tracking information is accessible through the A.R.T. subsystem. The VRPN server provides the corresponding position and, in case of six degrees of freedom devices, rotation measures for all tracking bodies found. A client module of the laboratory's own AR / VR platform (i.e. miniOSG, see below) continuously requests this data.

miniOSG acts as the nerve centre of the entire presentation-oriented system. It can run both helmet or large screen VR systems (e.g. opaque HMDs or the Responsive Workbench™, resp.) and VST HMD AR setups.

The main features of miniOSG are:

- Client-server-based real time rendering, one control client and several display servers.
- Management of a distributed scene graph environment extending OpenSG [OSG].
- Configuration interface for defining the runtime distribution of the application over the network, rendering and display properties, tracking devices, certain haptic devices, static descriptions for virtual objects, their organisation within the scene and an object-to-tracking-body assignment for a rudimentary interactivity.
- VST functionality by receiving the video data, separating fore- and background output (see Fig. 3-3) and embedding the results as shown in Figure 3-4.
- Communication layer for VST data transfer.
- Generic VPRN client to acquire tracking data.
- Auxiliary tools (e.g. background colour selection, simple event management).

Of particular relevance for this work is the VST feature. Its current implementation in brief: the captured and segmented hand is applied as an RGB-alpha texture onto one invisible carrier object per stereo channel. Each of these objects is computed as the perpendicular section of the corresponding rendering camera frustums at head-to-hand distance. The VST HMD calibration takes practical effect at this point, because the origins of the rendering cameras can now be superimposed on the estimated origins of the real cameras.

A number of essential aspects are either missing in the software system or likely to be insufficient for the intended VPC generation and management. The first drawback can be found in the video post processing application. Background is only determined by the colour hue. Saturation and luminance of the implemented HSL[3] colour scheme are excluded what unnecessarily hampers the segmentation. Foreground can often assumed to be brighter than the background. Technical tests also showed that the sensors of the HMD's built-in cameras produce a strong chromatic noise, especially in darker regions. This leads to an incomplete background subtraction (i.e. pixel ghosting). miniOSG itself does not permit spatial

3 HSL: Hue Saturation Luminance colour scheme

manipulations of the embedded video hand representation. Although, the invisible texture carrier objects do implicitly provide at least a basic shifting potential. Due to the apparent presentation character of the platform, interactive scene modifications beyond linking virtual objects to tracking bodies cannot be realised. A concept for a bi-directional simulation-to-platform data flow does not exist.

Summary

In conclusion, improvements to the background detection should be made in order to avoid visual artifacts adversely affecting an alteration of perception. Moreover, the complete VPC generation and management has to be integrated. An interactive dynamic scene control and some data exchange functions need to be added, too. If possible, the rather low AR frame rates should be accelerated (i.e. currently <12 frames per second) and the noticeable video lag should be reduced (i.e. currently >250 ms).

3.4 VPC framework

After the requirements definition and an analysis of the given system infrastructure, the resulting VPC generation and management framework (i.e. henceforth: VPC framework) is developed in Section 3.4.1. Generality and extendability have to be ensured. Further system deficiencies will be addressed in Section 3.4.2, if they are related to the overall functional goals of the VPC framework.

3.4.1 Hand texture carrier object repositioning

The lightweight solution for embedding the user's hand into a virtual scene can easily be exploited for static and dynamic visual discrepancies. Since the texture carrier objects consist of OpenGL [OGL] quads, redefined and put to the scene graph at each tracking client thread loop, their spatial attributes are available for additional computations. As described above, these quads are the result of a virtual camera frustum section performed in the viewing frame. This operation yields the final carrier objects' vertices later to be transformed into the virtual

world frame (see the VST HMD calibration). A visual repositioning of the real hand can thus be achieved by shifting these objects or rather their geometry (see Fig. 3-4) within the different frames of reference.

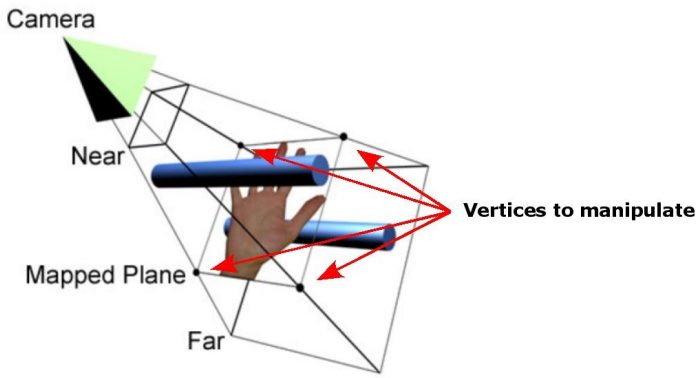


Figure 3-4: Mixing approach indicating the vertices to manipulate of one hand texture carrier object (see Ortega, 2007).

Moving the carrier objects in depth automatically rescales the hand as it would happen in reality. When inducing offsets in the vertical plane, disparity distortions may occur, because of the invariant relative video camera viewpoints (see Fig. 3-5).

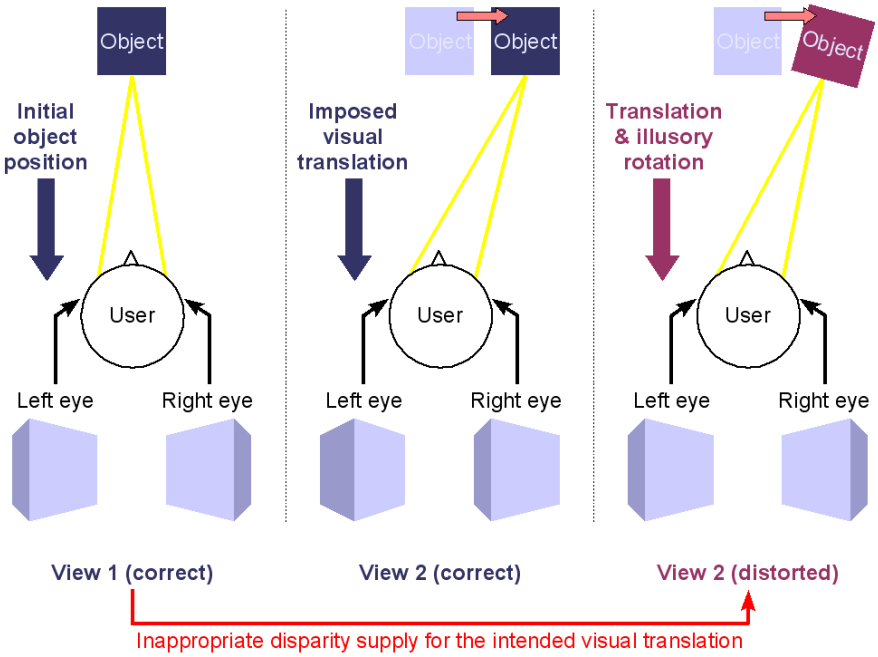


Figure 3-5: Disparity distortion due to invariant relative camera viewpoints.

However, considering the benefits of this planar, but stereoscopic real image embedding approach and the fact that actions in a co-located space (see Section 2.2) are not expected to exceed certain distortion limits, the technique can be used for generating VPC.

To manage static offsets as well as dynamic displacements in either the local viewing or the global world frame, a data flow concept and a set of basic control functions have to be provided. At the lowest level of the AR mixing, respective input fields have to be opened. Enabling this access represents the only intervention to the AR / VR platform kernel.

Regarding the VPC framework architecture (see Fig. 3-6), a concrete application has the possibility to trigger a fixed or a continuous decoupling of the visual hand feedback while specifying the manipulation target frame. In the static case, an offset vector $\vec{o}$ or a direction vector $\vec{d}$ plus an offset distance l in cm have to be passed. These parameters are pushed to the actual hand shift matrix m_{HS} . For dynamic displacements, a direction vector $\vec{d}$ and a displacement velocity v_{displ} in cm/s are required. The displacement control adapts to the duration of the tracking client thread loop and computes m_{HS} assuming a linear offset development. After the actual shift has been calculated, an optional target-frame-dependent feasibility test (FT) takes place in order to prevent inconvenient VPC (e.g. outside the video capturing or display range, potential perception stress). The test is based on the estimated visual-to-real hand deviation angle and returns an error message, if given constraints have been violated. Alternatively, the hand representation can fall back to any static 3D hand model. Offset thresholds are declared in a plain text configuration file. It is recommended to experimentally determine the true device properties for an optimised VPC bandwidth (see Section 5.4.2.2). m_{HS} is finally applied to the viewing- or world-related coordinates of the hand texture carrier objects.

At each tier before handing over the VPC data to the AR / VR platform, previous inputs, intermediate outputs and the current head and hand tracking information can be read by the concrete application or be stored by the VPC framework for diverse analysis reasons. Storage formats are defined in a plain text file, too.

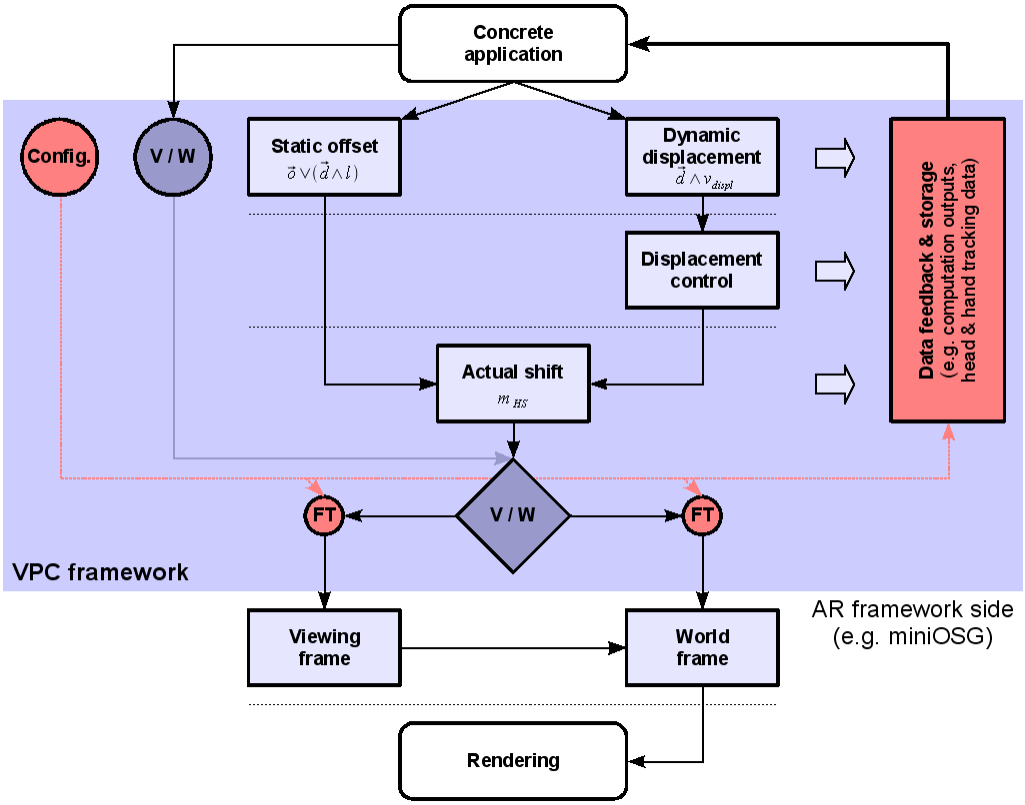


Figure 3-6: VPC framework model, including AR framework interaction.

Generality and extendability are important characteristics of the VPC framework. An access to the AR / VR software system is only necessary for the ultimate assignment of the visual hand shift. Beside the open framework entries, also m_{HS} is generic in terms of applicability. Even virtual hand avatars can be managed using this matrix. Albeit the upper two architecture layers can easily be extended, more specific modules should preferably be situated at the application side. From a development point of view, the VPC framework is designed as a Singleton and consists of a C++ application programming interface (API).

3.4.2 Other system enhancements

In the course of this section, additional system enhancements relevant for an efficient VPC-based interaction are presented. One critical side effect which would make sensory alterations (i.e. visual hand shifts) undesirably obvious to the user is the pixel ghosting mentioned in Section 3.3.2. What happens is that the existing background segmentation fails in darker

image regions, because of the chromatic noise produced by the low fidelity cameras of the employed VST HMD. To improve the subtraction quality, a luminance factor is added to the former simple hue test implicitly accounting for the brighter foreground hypothesis and solving the identified ghosting problem. When capturing the background, the average hue and luminance measures of the selected area are transferred to the video acquisition and post processing application. The hue test will only be done after a successful luminance high-pass. Pixels blocked by the high-pass filter are declared as background.

System performance is generally low. Various lags caused by, for instance, the video image transfer or some scene graph access operations, often cumulate to considerable display delays. Suchlike is known to increase the probability of subjective discomfort or cyber sickness. It can also directly affect sensorimotor behaviour due to spatiotemporal feedback incoherencies. However, a number of source code, algorithmic and scene graph access optimisations resulted in a 33% performance gain finally delivering 16 frames per second for the display loop and video lags of less than 150 ms. The video and virtual world contents were synchronised by buffering tracking for the duration of the average video latency (i.e. about 50 ms).

miniOSG has been extended to allow for interactive scene modifications (i.e. virtual object creation, removing and transformation), a basic limb-object collision detection, an event control, a local correction for head and hand tracking bodies (e.g. for precise pointing tasks) and supplementary audio. Most of these features are essential for the approaches presented in this work or their specific experimental investigation. But, at the same time, the main AR / VR platform's overall functional abilities are enhanced. A closer link to the VPC framework is established through the collision detection add-on within which, for the computation of apparent visual intersections, the position of the user's real hand can be multiplied by the actual hand shift matrix m_{HS} .

3.5 Proof of concept

The VPC framework has been integrated and informally verified on the system described in Section 3.3. Tests included gradual hand shifts and fixed offsets in all three dimensions. As long as the video cameras' capturing spaces were not exceeded and the user could still see his hand, the sensorimotor system appeared to adapt successfully to the imposed manipulations.

In the next two chapters, it will be shown that the developed and implemented concept can serve the required VPC generation and management needs. The system was further used in a preliminary psychological study. Other potential fields of applications will be discussed in Chapter 8.

3.6 Discussion

The goal of this chapter was to propose a generic and extendible VPC framework capable of inducing static and dynamic spatial VPC at hand level. Both co-located real hand feedback and 3D hand avatars are supported. The first is meant to intensify visual dominance and hence perceptual reliance on the limb used for interaction in VE. This merging of real and virtual elements qualifies the notion AR or, more precisely, AV.

The existing AR / VR system infrastructure was analysed with respect to the envisaged visuomotor conflicts. It was found that in the hardware platform, there are only a few constraining device properties (e.g. visual ergonomics of the VST HMD such as the field of view, resolution and colour consistency of the cameras and displays). The software platform exhibited a number of limitation to overcome in order to generate and manage VPC in the desired way (e.g. noisy background segmentation, no interface for visual hand shifts at all). Additionally, more general enhancements were necessary (e.g. revised data flow concept, interactivity functionalities, system performance improvements).

Based on this analysis and the VPC generation and management requirements developed beforehand, the VPC framework was elaborated. Modifications to the kernel of the underlying AR / VR platform are minimal (e.g. access to the methods used for the spatial visual integration of real objects into the virtual scene). The main features of the VPC framework are to provide an API for concrete applications, a set of basic control functions and a dedicated data flow concept. Drawbacks in the existing system infrastructure were addressed in a more comprehensive context contributing to global platform improvements.

3.7 Conclusion and future work

A flexible VPC framework has been designed and implemented. First, the general system characteristics were defined and second, the given hardware and software conditions analysed in order to isolate potential points to ameliorate. The synthesis of all these factors has led to the final framework which will be used and tested in the course of the following chapters.

However, limitations exist which could, in some cases, constrain the applicability of the approach or weaken the intended link to perception. The hand is integrated in the form of 2D textures. Although captured in stereo, the properties of the employed VST HMD and the embedding technique do allow for just a small image overlap (i.e. small effective stereoscopic region). Viewing the hand monoscopically could cause depth positioning difficulties, if there are no other depth cues available. State et al. (2001) showed how to overcome at least parts of this issue. Object intersections do normally produce linear cutting edges which let the hand hardly appear as belonging to the three-dimensional virtual space. The same holds for a visual artifact occurring, if the hand texture carrier objects are displaced away from the image border the real arm crosses. An alternative solution to these problems would be to reconstruct the actual hand shape in 3D and apply projective textures to the resulting object. But there are disadvantages in such often very complex techniques, too (see Section 2.3). Shadows cast by simplified invisible proxy objects could be another option.

Further restrictive elements are the manipulation scope and the tracking. In fact, for each hand, a single sensitive point can be taken into account for both inducing VPC and an interaction with the virtual world. That is, separated finger displacements are not possible. Advanced tracking and image segmentation or 3D reconstruction methods would be required to get closer to a full hand support, though this might not always be a major goal. The interaction purpose should determine the system design here as well.

The VPC framework could be extended to offer multiple control instances. For now, only one global VPC generation and management interface exists. It would thus be up to the application to take care of an effect mixing, if needed.

As mentioned in the last section, another set of problems can directly be related to the hardware used so far. But once performance and ergonomics improve, these limitations can be reduced. And AR / VR technology evolves rapidly.

4 INTUITIVE CONTROL FOR TOUCHING VIRTUAL SURFACES

A novel method for near space interaction is proposed in this chapter. It aims at merging previously independent concepts for an interaction within hand's reach and enrich them by adding visuo-proprioceptive conflicts (VPC) of the user's hand.

Specific related work will be discussed (see Section 4.2), followed by the theoretical and technical fundamentals of the novel approach (see Section 4.3). After, the central hypotheses are formulated (see Section 4.4). An experiment has been conducted in order to evaluate the synthesised paradigm as well as to investigate behavioural consequences of using it (see Sections 4.5 to 4.7). The conclusion and envisaged future work are topic of Section 4.8.

4.1 Introduction

Control and manipulation of virtual worlds require both dedicated tools for highly specific tasks (e.g. in CAD[4]) and immediately understood techniques which can be used by everybody without any prior learning. For touching a virtual surface within hand's reach (e.g. when controlling a 3D application interface or working with close objects), such a technique could provide an intuitive feeling of control, as if the user would control the own hand in reality. Hence, a promising way to bridge the gap between performing near real and virtual world actions in the absence of haptic devices could be to build a more robust perceptual link to the tool used for interaction.

Vision of the real hand or its representation plays a decisive key role, specially during goal-directed pointing or reaching movements. Our brain maintains a continuous observation and correction process to compensate for reaching errors and / or target modifications [Pa96]. The resulting hand trajectories towards the target and the final adjustment on the target itself reflect an optimisation in terms of end effector stability and pointing or touching accuracy. An intensified limb attribution might support this process. (see also Sections 2.1 and 2.2)

4 CAD: Computer Aided Design

To exploit these and other behavioural aspects for 3D UI, the novel video-see-through-based interaction paradigm presented here aggregates the advantages of a) acting in co-location, b) using visual movement constraints, c) avoiding occlusion violations and d) providing a convincing visual feedback of the hand. Important subgoals are spatial employment flexibility and system integration ease.

4.2 Related work

Several studies have shown that a near space interaction in co-location is convenient from a performance and a user preference point of view (see Section 2.2). Further, seeing the own real hand may improve acceptance, control comfort and preserve the natural sensorimotor organisation. Avoiding occlusion violations reduces integration seams between the real and the virtual world (see Section 2.3).

Concerning a more intuitive surface touching impression, researchers have mainly focused on haptic and visual hand movement constraints. Active haptic devices attached to the real limb can limit the interaction space to virtual objects by rendering appropriate intrusion prevention forces. In [Or06], for instance, a decoupled computation of continuous visual collisions and a corresponding constraint-based force feedback applied to the user's hand were proposed. The system they utilised was the Stringed Haptic Workbench [Ta03] (see Fig. 4-1).

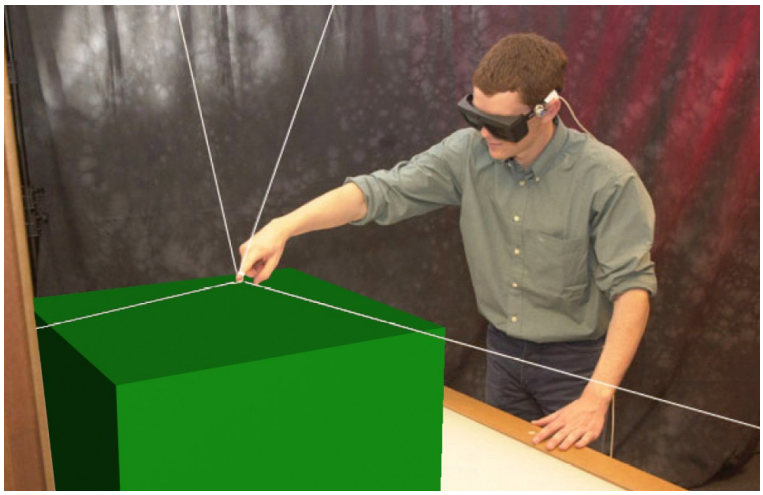


Figure 4-1: Stringed Haptic Workbench,
compositing of a user touching a virtual cube (see Tarrin et al., 2003).

Techniques for visually preventing the hand representation from entering virtual objects usually operate on 3D hand avatars (e.g. static 3D hand models or pointers in occlusive VE, see Burns et al., 2005). Because only those can freely be controlled within the virtual space so far. It has been found that subjects were a lot less sensitive to visual hand movement constraints compared to virtual object interpenetrations. The natural expectation of spatially constrained actions can already be satisfied by blocking the avatar of the user's hand at virtual surfaces. Burns et al. (2005) applied the so-called rubber band method (RB) [Za01] on object entering. This method minimises the distance between the two representations by keeping the visual hand, in case of virtual surface constraints, as close as possible to the real hand. Since a purely virtual scene was presented, including the hand, the potential multisensory integration and attribution barriers driven by displaying artificial limb substitutes remain an unresolved issue.

4.3 Interaction paradigm fundamentals

The overall objective is to combine the indicated benefits of an interaction within hand's reach and enhance them with a visually faithful representation of the hand (i.e. embedded video feedback). To this end, the system infrastructure and the VPC framework discussed in Chapter 3 will be adopted. Haptic devices are excluded from the design because of the intended lightweight implementation (i.e. regarding spatial availability, system complexity, calibration needs and maintenance effort).

VPC occur in the case of visual hand movement constraints to virtual object surfaces. Hand tracking data is used for computing the actual object penetration. On entering, RB is applied to obtain the compensation vector $\vec{c}$ for the visual repositioning of the user's hand. Once the visual hand releases from the virtual surface (i.e. temporarily invariant norm of $\vec{c}$), the incorporated incremental motion method (IM) [Za01] leads to a maximised motion coherence. That is, $\vec{c}$ is kept constant assuring action reliability. The principal process is depicted in Figure 4-2.

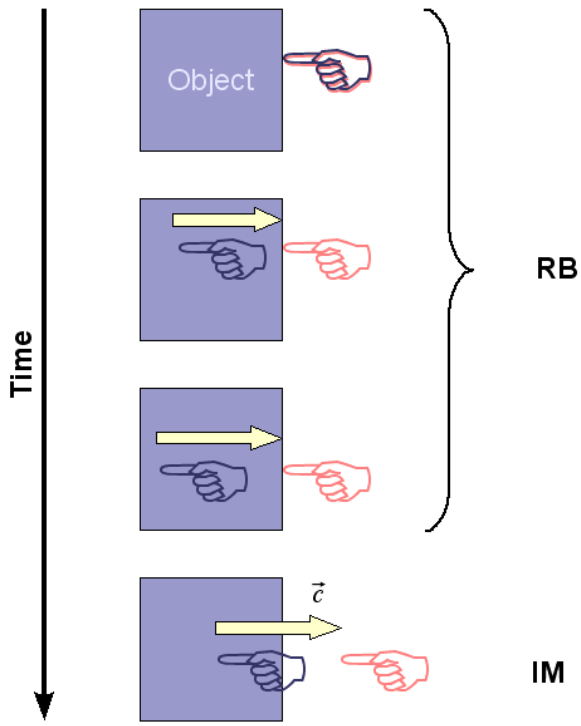


Figure 4-2: RB-IM process on surface entering and release (dark blue: real hand, salmon: visual hand).

In case of intersections with given scene constraints, $\vec{c}$ is declared as a static offset and fed into the VPC framework. The hand texture carrier objects get therefore shifted in space about the same amount (see Section 3.4.1). As a result, the visual hand seems to rest on or slide over the virtual object entered with the real hand.

Additional contact cues beside stereoscopy are either rendered sound (e.g. based on the apparent object material) or simple object highlights. Alternative approaches are pointed out in Section 4.8.

4.4 Hypotheses

It has been stated in different studies that subjects often preferred hand representations at higher visual fidelity levels (see Section 2.2). This is an important fact from a user-centric interaction-ergonomics-oriented viewpoint.

Hypothesis 1 addresses the subjective component of an intuitive control in a virtual object touching task. In agreement with previous observations, it is expected that the appearance of the novel interaction paradigm will be preferred over other classical hand representations or avatars (i.e. detailed 3D hand model, simplified 3D hand model, ordinary 3D pointer arrow, see Fig. 4-3). The advantage should persist for different aspects of a virtual surface touching scenario (i.e. visualisation quality, final pointing accuracy, hand movement naturalness and overall comfort). Further, acceptance may decrease with lower realism levels.

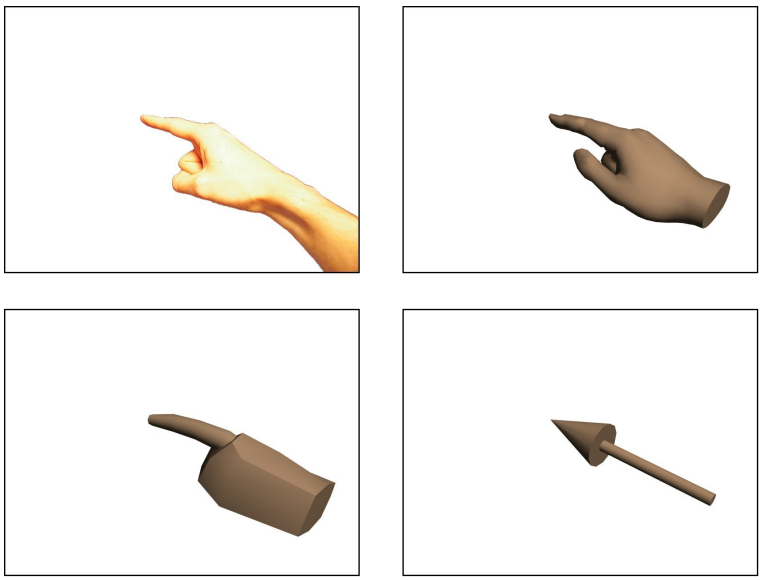


Figure 4-3: Hand representations used during the experiment
(i.e. upper left: video feedback, upper right: detailed 3D model,
lower left: simplified 3D model, lower right: 3D pointer model).

Hypothese 2 focuses on the influence of the above-mentioned hand feedback types on the behaviour in a goal-directed pointing task. It is expected that the novel interaction paradigm can, due to perceptual reliance (see Section 2.2), positively affect motor performance. More precisely, first, an improved final pointing stability is assumed, if the realism level of the hand representation increases. Stability is considered as a rough accuracy measure. Second, it is

thought that the intuitive feeling of control can be reflected by the virtual object entering with the real hand while the visual hand rests on the object's surface. That is, higher feedback fidelity could result in smaller VPC (i.e. reduced target entering depth) until the subjective detection of the touching event.

The reaching duration and the hand trajectory length, between the moments the hand enters the field of view and actually reaches the target, should be studied, too. These are indicators for action optimisation and economics. Since the field of view of the HMD is rather small (see Section 5.4.2.2), the visible hand transport phase is inherently not very long. Effects may thus be marginal, but yet insightful.

The experiment designed to investigate these questions is described in the next section.

4.5 Experiment

To test the hypotheses established in Section 4.4, an experiment was prepared which consisted of two parts: a repetitive 3D pointing task (see Section 4.5.3.1) and a sequential questionnaire (see Section 4.5.3.2). The latter was given directly after the pointing task was completed.

4.5.1 Subjects

Sixteen adult volunteers (i.e. 20 – 40 years old, 7 female, 9 male) participated in the study. None of them reported serious vision problems (i.e. either normal or corrected to normal vision). A few subjects had some prior non-expert knowledge on AR / VR and / or human perception. All were naive about the study purpose and had never used the experimental setup.

4.5.2 Factorial design

The experiment followed a 4 x 2 factorial design of which factor one specifies the number of hand representations (see Fig. 4-3) and factor two the number of pointing target locations on a cube surface (see Fig. 4-4).

Beside video feedback of the hand, another three virtual hand representations of different realism levels were chosen. In descending order with respect to the degree of realism:

1. Real hand video (i.e. the most natural available).
2. Detailed 3D hand model (i.e. close to a realistic shape).
3. Simplified 3D hand model (i.e. strongly simplified shape).
4. Ordinary 3D pointer arrow (i.e. abstract representation).

All 3D hand models, including the arrow, had a common visual appearance in terms of size and a uniform skin-like shading. Additionally, they were displayed at six degrees of freedom according to the real hand tracking. The given video-to-virtual world lag of about 50 ms was simulated for the pure virtual scenes (i.e. without video feedback, see also Section 3.4.2.).

To prevent a fast hand movement adaptation, it was decided to use more than one target location. Two sufficiently separated targets (i.e. top near and bottom far, 3 cm edge length, see Fig. 4-4, right) were expected to meet this requirement. Advanced behavioural analyses may also be permitted when varying the target characteristics in this manner.

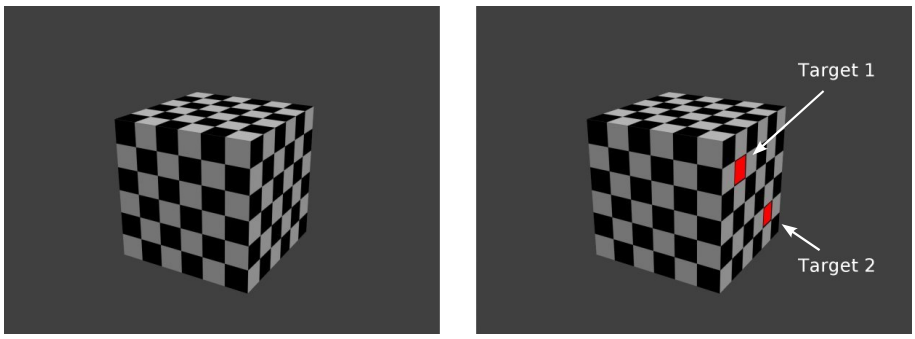


Figure 4-4: Virtual scene (left), with target locations (right).

In sum, there were eight conditions to be randomly distributed and equally weighted over the duration of the experiment to prevent any effect carry over. Hence, an 8 x 8 random latin square was generated to assure balanced trial sets. Each participant had to accomplish 16 repetition per condition making up 128 trials in total.

4.5.3 Procedure

Before volunteers could take part in the experiment, they had to answer several introductory questions. This way, it should be verified that all subjects would fit the desired profile. That is, they had to be naive about the purpose of the experiment, never used the experimental setup, at the most posses only little experience in AR / VR theory and practice and at best never participated in former AR / VR experiments. General information concerning age span, gender, dominant hand, vision problems etc. was asked at the experimental session.

For a detailed explanation of both the pointing task and the administered questionnaire, refer to the next two sections.

4.5.3.1 Pointing task

To be able to present the cube and thus the pointing targets at a similar relative height, the subject's shoulder was considered as the reference. The video see-through (VST) HMD was adjusted to the eye distance by shifting display oculars accordingly. A clear view of both display images had to be confirmed before continuing. Further, the rest position between trials (see Fig. 4-5, left) and the grasp gesture for pointing and holding the hand tracking device (see Fig. 4-5, right) were explained.

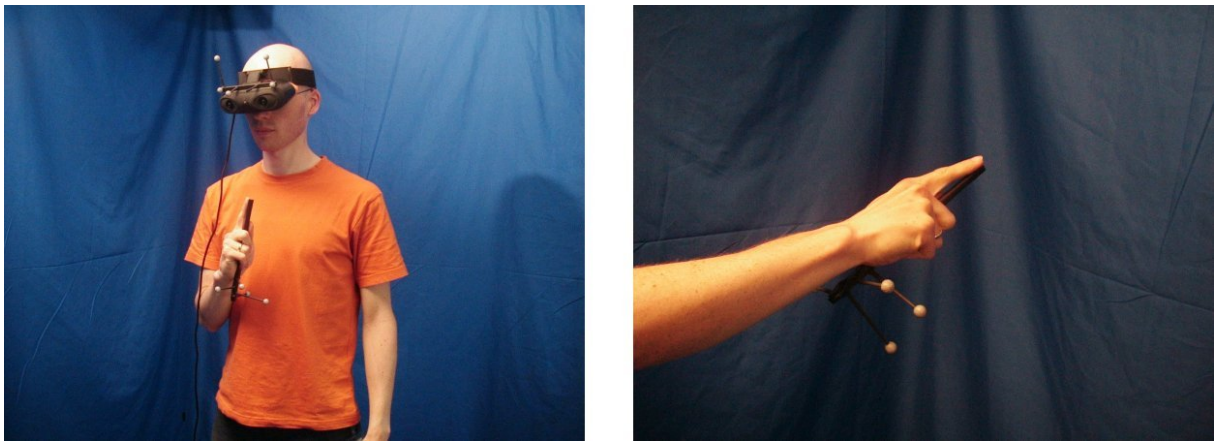


Figure 4-5: Rest position (left), grasp and pointing gesture (right).

Regarding the actual task, subjects were requested to line up on a defined position looking towards the blue-covered walls of the room. They were instructed to use the right index finger in order to touch the centre of the red square that would appear on the right side of the cube (see Fig. 4-6).

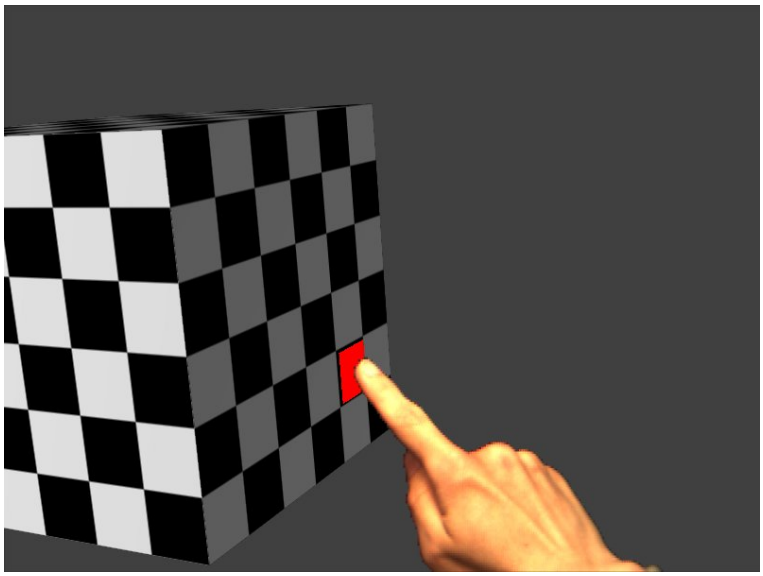


Figure 4-6: Application screenshot of a user touching the target (i.e. VST HMD view).

A simultaneous acoustic trigger notification was played back (i.e. first beep). Subjects were told to have four seconds to do the pointing and to perform a precise rather than a rapid touching movement. Once they thought they had touched the target's centre, subjects had to

return to the rest position. No other contact cues were presented. An acoustic trial end notification (i.e. second beep) would be heard at the same time the target disappears. This procedure recurred 128 times per subject, with a relaxation break of 5 minutes at the half.

4.5.3.2 Questionnaire

A sequential questionnaire (see Appendix 1) was given to the participants immediately after the pointing task was done. It contained consecutive open questions as well as a subjective hand representation evaluation section. It was not permitted to return to previously completed pages. Subjects were free to give written comments to any question. Discussions were not allowed, except for cases of comprehension problems.

Part one of the questionnaire focused on the differences perceived between trials over the whole experiment. If differing conditions were remembered, their total number, the globally preferred, helping or even interfering ones had to be indicated and the sensation be explained. The next set of questions asked more precisely for the recognised hand representations, again including their number and the subjective preference. A sketchy drawing of each recalled hand type was requested.

The second part of the questionnaire mainly consisted of a multi-level hand representation evaluation. Subjects were shown images of the four hand representations used during the experiment. They had to assess the visualisation quality, the final pointing accuracy on the target, the naturalness of the hand movement or transport towards the red squares and the overall comfort while performing the tasks. Each aspect had to be evaluated, ranging from 1 for best to 5 for worst, as a function of the visual hand representation. A special justification was required for the overall comfort assessment. General remarks could be given in the end of the questionnaire.

4.5.4 Data acquisition and analysis

The basis for the behavioural analysis was head and hand tracking information recorded at approximately 60 Hz. A first processing of this raw data yielded specific action events (i.e. entering the field of view, approaching the cube surface, stabilising the finger on the target and releasing the visual hand from the cube surface). In addition, supplementary measures like the target entering depth (i.e. the maximum VPC), the cumulative hand movement path length and a sample-wise 3D hand velocity were computed. However, the four behavioural main variables used for statistics were:

1. Coarse hand oscillation around the target (i.e. repeated visual target contacts before returning to the rest position, see Section 4.6.1).
2. Target entering depth (i.e. maximum target penetration before returning to the rest position, see Section 4.6.2).
3. Hand movement duration (i.e. time between entering the field of view with the visual hand and stabilising it on the final target, see Section 4.6.3).
4. Hand trajectory length (i.e. path length between entering the field of view with the visual hand and stabilising it on the final target, see Section 4.6.4).

From the questionnaire, the hand representation assessment section was considered for the analysis (see Section 4.6.5). Other responses and comments served, at least for now, only as a supporting source for a clearer interpretation of the results.

Behavioural and subjective evaluation data was analysed using descriptive statistics (i.e. mean and standard deviation, SD), followed by a repeated measures Analysis of Variance (ANOVA) and, if adequate, correlation and / or post-hoc tests (e.g. Pearson's product moment correlation and / or pairwise comparison or Fisher's Least Significant Difference, LSD, resp.).

4.6 Results

The results of the hand trajectory analysis will be presented first. Thereafter, in Section 4.6.5, attention is drawn on the qualitative hand feedback assessment.

For not yet understood continual reaching errors in 4 subjects, they had to be excluded from the analysis. Aside from that, a few trials of the remaining experimental data were removed, mostly for technical reasons (e.g. corrupted tracking data, malfunctions of the system or the simulation).

4.6.1 Coarse hand oscillation around the target

At the moment, a subject had touched a target for the first time a counter was launched. All subsequent touching repetitions led to counter increments as long as the current trial's data acquisition ran. It is therefore a measure for the hand movement guidance quality.

There was no effect of the hand representation on the coarse pointing accuracy found ($F(3, 33) = 0.8$; $p > 0.5$). The target location effect was also not significant ($F(1, 11) = 3.47$; $p > 0.09$), although touching stability appeared to be more than 7 times higher on the far target (i.e. $SD_{far} = 0.009$ vs. $SD_{near} = 0.065$). Figure 4-7 shows these results. The number for the hand representation refers to the list of Section 4.5.2.

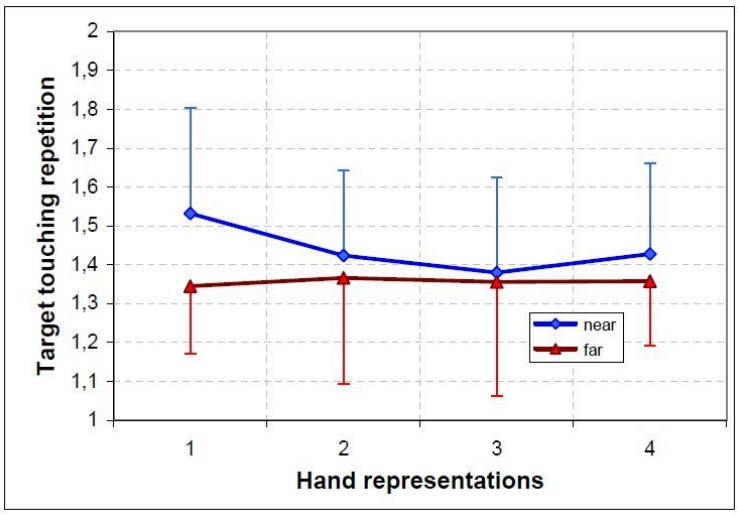


Figure 4-7: Effect of the hand representation on the coarse hand oscillation (means, SDs).

4.6.2 Target entering depth

This variable reflects the maximum target penetration before subjects decided to move their hand back to the rest position. Thus, the target entering depth tells about the lateral hand position estimation error or touching overshooting. The visual hand was always constrained to the cube's surface.

Analysis yielded a significant effect for both the hand feedback ($F(3, 33) = 2.89$; $p < 0.05$) and the target location ($F(1, 11) = 2.89$; $p < 0.003$). Regarding the hand representation, a post-hoc LSD test (i.e. pairwise comparison) revealed significant differences between the real hand video and the ordinary 3D pointer arrow ($p < 0.021$) as well as between the simplified 3D hand model and the ordinary 3D pointer arrow ($p < 0.015$). The detailed 3D hand model was situated at an intermediate level without any statistically relevant performance variation. Correlation effects were not found, neither for the target factor in general ($r^2 = 0.005$; $t = 0.72$; $p > 0.4$) nor for any specific target ($r^2_{near} = 0.008$; $t = 0.59$; $p > 0.5$ and $r^2_{far} = 0.004$; $t = 0.44$; $p > 0.6$).

Independent of the hand representation, participants had a better control over their limb when pointing at the far target (see Fig. 4-8).

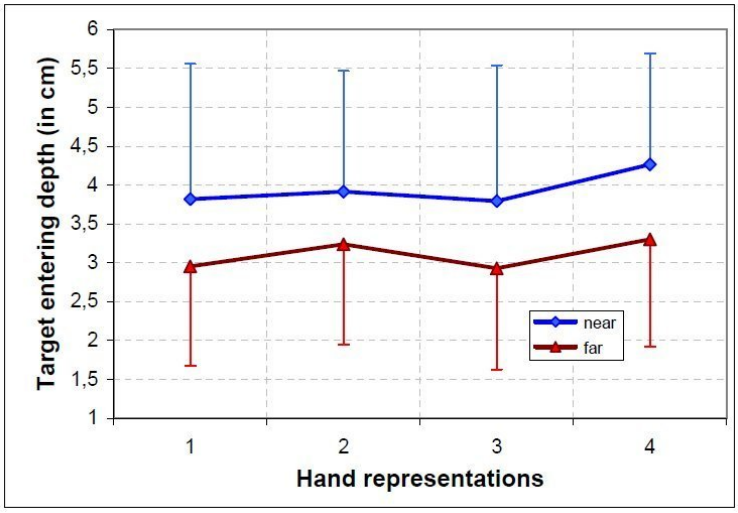


Figure 4-8: Effect of the hand representation on the target entering depth (means, SDs).

4.6.3 Hand movement duration

Expressing the action in an economic way, the hand movement duration describes the time a subject has spent between the following two key events:

1. Hand entering the field of view (i.e. hand tracking position intersected with at least one of the two virtual viewing frustums).
2. Stabilising the visual hand on the target square (i.e. last target contact of the displayed hand, incl. VPC, before trial end).

An ANOVA indicated that the hand movement duration was not influenced by the hand representation ($F(3, 33) = 1.25$; $p > 0.3$). But there was a target location effect observed ($F(1, 11) = 10.63$; $p < 0.008$). That is, pointing towards the close target was significantly faster performed (see Fig. 4-9).

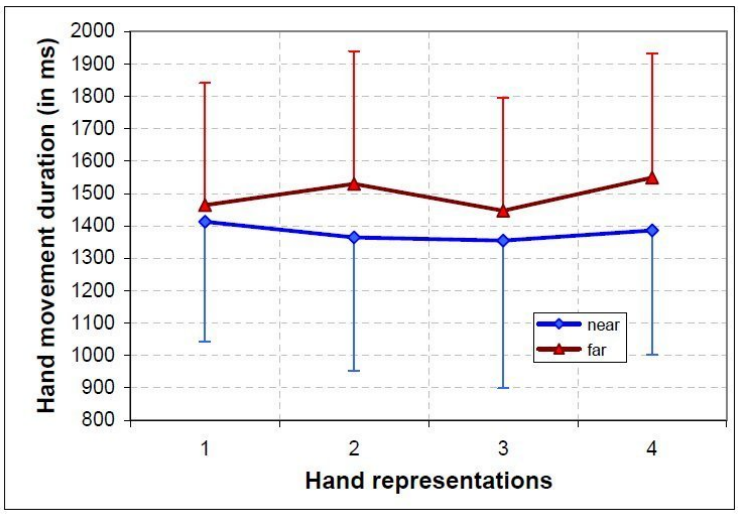


Figure 4-9: Effect of the hand representation on the hand movement duration (means, SDs).

4.6.4 Hand trajectory length

The events delimiting the hand trajectory length and hence the second economics measure were the same like for the hand movement duration. No effect on the hand trajectory length was found, neither caused by the hand representation ($F(3, 33) = 0.42$; $p > 0.7$) nor by the target location ($F(1, 11) = 3.5$; $p > 0.09$). The latter factor shows only a light tendency (see Fig. 4-10).

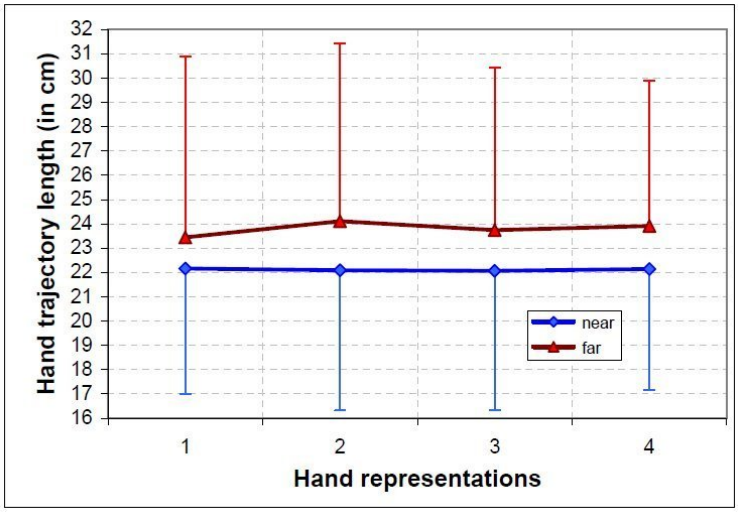


Figure 4-10: Effect of the hand representation on the hand trajectory length (means, SDs).

4.6.5 Hand representation evaluation

The analysis of the subjective hand representation evaluation scores was concentrated on (for details see Section 4.5.3.2):

1. Visualisation quality.
2. Final pointing accuracy.
3. Hand movement naturalness.
4. Overall comfort.

Subjects had to rank each of these aspects as a function of the hand representation from 1 to 5 (i.e. best to worst, resp.). One participant did not complete the whole evaluation because of strong uncertainties in some cases. His data was excluded from the global analysis and all subquestions concerned.

In total, the hand feedback affected the subjective responses significantly ($F(3, 30) = 42.01$; $p < 0.0001$). The Pearson's product moment correlation test showed also a highly significant positive correlation ($r^2 = 0.45$; $t = 12.19$; $p < 0.0001$). This indicates that ranks decreased with the realism level of the hand representation (see Table 4-1). A question effect was not found ($F(3, 30) = 0.62$, $p > 0.6$).

Table 4-1: Overall assessment means for the hand representations used.

Real hand video	Detailed 3D hand model	Simplified 3D hand model	Ordinary 3D pointer arrow
1.52	1.67	2.85	3.73

A post-hoc Newman-Keuls test yielded significantly better total results for the real hand video and the detailed 3D hand model compared to the other two hand representations (i.e. for both: $p < 0.0002$). Moreover, the ordinary 3D pointer arrow was rated significantly worse than the simplified 3D hand model ($p < 0.0003$). Even if seeing the real hand was generally preferred, no difference was found between the real hand video and the detailed 3D hand model ($p > 0.6$),

After having this global acceptance image obtained, subquestions were analysed separately. Results are as follows (i.e. main effect and correlation, see also Fig. 4-11):

1. Visualisation quality: $F(3, 33) = 24.36$; $p < 0.0001$ and $r^2 = 0.56$; $t = 7.69$; $p < 0.0001$.
2. Final pointing accuracy: $F(3, 33) = 5.54$; $p < 0.004$ and $r^2 = 0.23$; $t = 3.72$; $p < 0.0006$.
3. Hand movement naturalness: $F(3, 30) = 30.56$; $p < 0.0001$ and $r^2 = 0.54$; $t = 7.07$; $p < 0001$.
4. Overall comfort: $F(3, 30) = 33.19$; $p < 0.0001$ and $r^2 = 0.5$; $t = 6.61$; $p < 0.0001$.

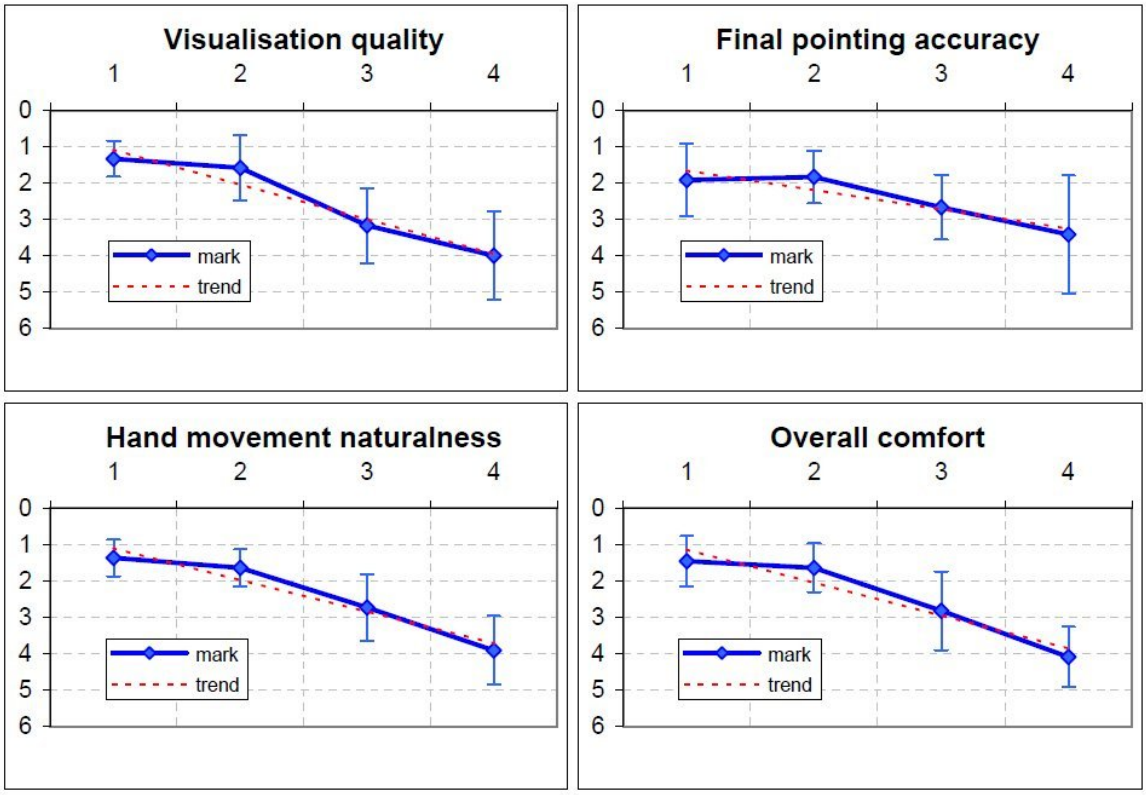


Figure 4-11: Ranking of the hand representation as a function of the task aspects (means, SDs), with “marks” ranging from 1: best to 5: worst.

Pairwise comparisons (i.e. Fisher's LSD) within each subquestion revealed similar factor constellations as they were found for the global view (i.e. in most cases: $p < 0.001$ or smaller). The only exception was the final pointing accuracy. Here, the simplified 3D hand model was, from a statistical point of view, not evaluated differently from all others.

4.7 Discussion of the experiment

The main purpose of the experiment was to test both the user preference and the behavioural consequences hypotheses stated in Section 4.4.

In *Hypothesis 1*, a subjective visual advantage of the novel interaction paradigm over other classical hand representations was expected to occur and to persist for several steps of a goal-directed pointing movement (i.e. touching a virtual surface). A lower realism level should lead to a decline in acceptance. Qualitative results show strong preference and correlation effects which widely confirm observations made in previous studies. The more realistic the hand appears in a generic near space interaction task, the better users feel when acting in a VE. Although no statistically significant difference in the evaluation was found between the real hand video and the detailed 3D hand model conditions, participants mostly preferred to see the own hand: “seems to be very intuitive”, “preferred the real hand”, “was easy to move towards the red square”, “felt to hit the target more quickly with my hand”, “one can better estimate the hand position and it is more comfortable”, “my hand was the most natural”, “because it was my hand”, “better surface understanding and space perception”, “comforting to know that I can see my own hand”. Contrary opinions were sometimes expressed as well, for instance: “preferred the arrow for accuracy reasons”, “the arrow for its precision”, “the virtual hand, since it 'fits' with the virtual cube”, “the 3D hand looked clean (...) did not like the pixelisation of the video hand”, “the virtual hand, because it looked '3D'”. Beside technical fidelity, it was mostly the precision which was criticised. The accuracy subquestion was actually the only one which showed slightly less distinct hand representation and correlation effects. In summary, the detailed 3D hand model was often able to compete with the proposed interaction paradigm. Reasons for that could have been of technical nature due to capturing and mixing limitations. Further, the task type did not require complex hand or finger movements (e.g. grasping), so that the benefits of a real time visual feedback were not fully exploited. In fact, a static gesture was sufficient. The initial hypothesis has hence to be modified. In the studied case, for a typical pointing-like virtual object touching situation within hand's reach, a high realism level is desirable, but real hand video is not essential. However, seeing the real hand, for now as video, seems nonetheless to deliver the most intuitive subjective interaction impression.

Hypothesis 2 addressed behavioural consequences of using different hand representations, including the novel interaction paradigm, in a virtual object touching task. The visual hand feedback was spatially constrained to the object's surface. It seems that the visual limb fidelity has only little influence on motor behaviour under the given conditions. The 3D pointing stability within a 3 x 3 cm target was not affected. Lateral overshooting was found to be the strongest for the most abstract hand feedback (i.e. ordinary 3D pointer arrow). This was expected. But, interestingly, several subjects thought to be “more precise” when handling the arrow because of its sharp end. One reason for the increased actual position estimation error could have been the non-hand-like shape which was integrated the least “natural” (i.e. limited attribution) into the visuomotor control process. The ordinary 3D pointer arrow may thus be considered as the least intuitive. A clear performance benefit of the real hand video compared to the other 3D hand models could not be seen. There was also no hand representation correlation effect found suggesting that, in the described pointing scenario, a highly realistic hand feedback does not improve hand movement precision and stability. A reformulation of the initial hypothesis is hence necessary. The above-mentioned control deficits hold solely true for purely abstract virtual hand substitute (e.g. an arrow or a ray).

As presumed, the hand representations did neither affect the hand movement duration nor the hand trajectory length of the visible hand motion towards the targets. This happened probably for two reasons. First, the main task consisted of touching a visual target with a certain accuracy. Subjects had therefore to focus their attention on the target and peripheral visual guidance of the hand did not play a major role. Second, the limited fields of view of the VST HMD (i.e. cameras and displays) prevented participants from seeing the hand for the most part of the transport phase. A detailed hand representation might prove beneficial, if the user really focuses on it, for instance, when performing precise virtual object manipulations. That is, to reduce system lags, one could use a gaze-based level of limb realism to avoid complex real hand embedding techniques as long as the attentional focus is somewhere else.

Concerning target effects, the target entering depth as well as the hand movement duration were statistically dependent on the target location. However, the coarse hand oscillation and the hand movement trajectory length showed only tendencies. It appears nonetheless to be valid to speak of two widely separated targets initially introduced to limit adaptive behaviour. The far target revealed better overall stability results. The stereoscopic quality could have been worse for viewing the hand manoeuvring towards the close target (see State et al., 2001).

4.8 Conclusion and future work

In this chapter, a novel near space interaction method was proposed. The basic concept relies on a VST HMD AR system and the VPC framework developed in Chapter 3. The new technique integrated useful characteristics for an interaction within hand's reach, like a) co-location, b) visual movement constraints, c) correct occlusions and d) a convincing visual feedback of the hand. Adding the latter was meant to reinforce an intuitive control for virtual surface contacts by directly reflecting intentionally performed actions.

Two hypotheses were established predicting a better user acceptance and an improved action stability and precision with an increased realism level of the hand. The experimental procedure comprised a goal-directed 3D pointing task and an evaluation questionnaire. Four alternating realism levels of the hand were presented while the pointing targets appeared at two spatially varying locations. Questionnaire results show that a higher visual fidelity of the interacting limb is preferred. Subjects further clearly indicated the intuitive character of the novel interaction paradigm. However, an overall ranking did not reveal a statically significant benefit of the real hand video over the detailed 3D hand model. In the behavioural data analysis it was found that the touching event detection was worst in terms of lateral target overshooting for the most abstract hand representation used in the experiment (i.e. 3D pointer arrow). The perceptual link might have been the weakest in this case leading to a reduced involvement of natural visuomotor control processes.

This suggests that for classical virtual world pointing operations in near space, a high realism level of the limb visualisation can improve the subjective feeling of control and comfort. But there is no evidence that providing real hand feedback has an impact on motor performance in pointing interaction with a quasi-static hand gesture. When designing AR or VR systems, challenging real limb embedding techniques should be counterbalanced with the actual interaction goal. Applicability, spatial employment flexibility and system integration ease of the novel paradigm could be demonstrated (see also Chapter 3).

Future work includes a number of technical system and approach enhancements. Beside what has been envisaged in Section 3.7 (e.g. improved near space stereoscopy, additional depth cues for the texture-based hand feedback), the system and the VPC framework could profit from a gaze-based level of hand realism taking the attentional focus of the user into

account. This would help to reduce the computational system load. Material and hand-kinematics-dependent sound cues providing acoustic tapping, knocking or sliding impressions may support the intuitive feeling of an otherwise only visually constrained spatial interaction. Moreover, one could think of vibrotactile feedback at the fingertip on virtual surface contacts and displacement adjustments while approaching the surface. A common issue of all the techniques relying on VPC is the cumulative sensory discrepancy, for instance, on repeated object contacts. In order to allow for a continuous interaction without being forced to actively reset VPC, to limit potential adaptation problems and the interaction performance loss (see Sections 2.1 and 2.2), the visual hand representation should eventually be aligned again with the real hand. To this end, Chapter 6 proposes a general purpose solution.

Finally, effectiveness and relevance of the technical and methodological enhancements have to be experimentally verified. An exciting open question is also whether more complex manipulation tasks requiring variable hand gestures and / or finger movements could elicit definite advantages of the novel interaction paradigm. Hence, the principal idea of gaining in intuition due to perceptual reliance is still under investigation.

5 HAND-DISPLACEMENT-BASED ACTIVE PSEUDO-HAPTICS

The goal of this chapter is to show that visuo-proprioceptive conflicts (VPC) at hand level can evoke a novel type of force illusion: *H*and-displac*EM*ent-based *P*seudo-haptics (HEMP). The feedback can be qualified as *active* and does not require any haptic support devices.

Specific related work is reviewed in Section 5.2. The general technical basis extending the VPC framework will be introduced in Section 5.3, followed by an elaboration of the force field (FF) illusion and response model (FIRE) as well as its computational core (see Section 5.4). The main hypotheses are stated in Section 5.5 and experimentally investigated from Section 5.6 to 5.8. Conclusion and future work will be addressed thereafter in Section 5.9.

5.1 Introduction

Haptics has become an important modality for recent VE. While research is often focusing on active haptics, an increasing number of works explores alternative ways. Lighter techniques which have the potential to reduce system and device complexity include passive haptics, pseudo-haptics and sensory substitution. Due to properties of the human sensory system, further simplifications seem to be possible. That is, since the brain is able to merge under certain conditions even conflicting sensory modalities into a stable precept (see Section 2.1), the user's experience can efficiently be deceived. In the field of pseudo-haptics, an altered passive force perception was achieved by coupling a manipulated visual feedback with a force sensor (for more details, refer to Section 5.2).

The active pseudo-haptics approach proposed in this chapter continues prior work in this area, but it is novel in several respects. First, the term “active” indicates that the illusory force should be perceived as being exerted. Second, haptic-like sensations may be induced in the absence of any active or passive haptic device. That is, virtual phenomena could deliver some force feedback without the constraints frequently found in haptics hardware (e.g. limited interaction space, large room occupation). Third, relying on sensory integration and visual

capture, adaptive behaviour, a sustainable limb and action attribution (see Sections 2.1 and 2.2), an illusory force sensation might be generateable by triggering a certain motor activity at arm level combined with a plausible visual hand displacement on exposure to a virtual FF. The sensation should be elicited without prior learning. To meet the sensory manipulation requirements, a video see-through (VST) HMD AR setup (see Chapter 3) will be used.

However, this technique is not supposed to replace real haptic rendering. But it could, in adequate environments, provide a supplementary force feedback channel.

5.2 Related work

Haptic illusions which are based on the modulation of given force stimuli by vision have a long history, leading back to Charpentier's size-weight illusion (see Murray et al., 1999). In this pioneering work, it was shown that subjects usually estimated the weight of objects of equal mass depending on their apparent visual size. That is, the larger the object appeared, the lighter it was perceived (see Fig. 5-1). Reasons for this anticipatory effect are pointed out in Section 2.1.

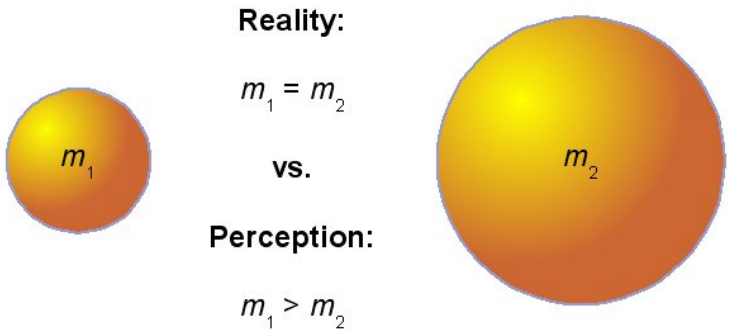


Figure 5-1: Size-weight illusion experiment.

Recent pseudo-haptics approaches follow similar strategies of altering perception through vision. In [Lé00], for instance, a stiffness feeling was simulated by linking a perturbed visual feedback to the “internal isometric device resistance” of a Spaceball™ (see Fig. 5-2). Two experiments were carried out in which subjects had to discriminate compliances of either only virtual springs or between real and virtual ones. The Just Noticeable Difference (JND) in the

real-to-virtual comparison case was found to be consistent with previous results. Lécuyer et al. (2000) concluded that the passive apparatus they used could simulate some kind of haptic information.

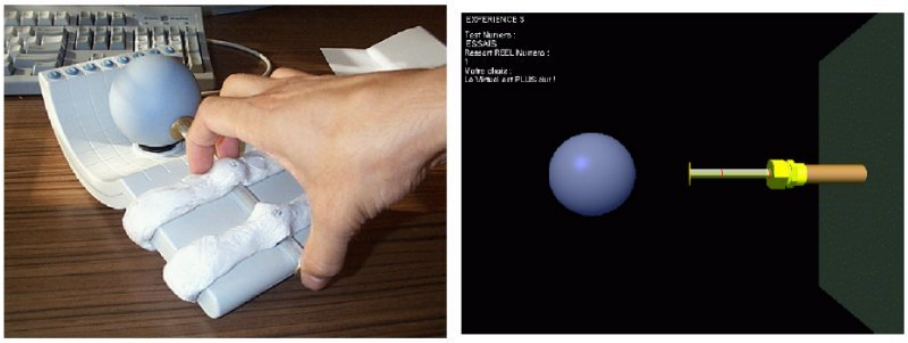


Figure 5-2: Pseudo-haptic spring stiffness experiment (see Lécuyer et al., 2000), device (left) and visual feedback (right).

It has also been demonstrated that a torque impression could successfully be induced by employing isometric as well as slightly elastic force input device combined with a virtual torsion spring [Pa04] (see Fig. 5-3).

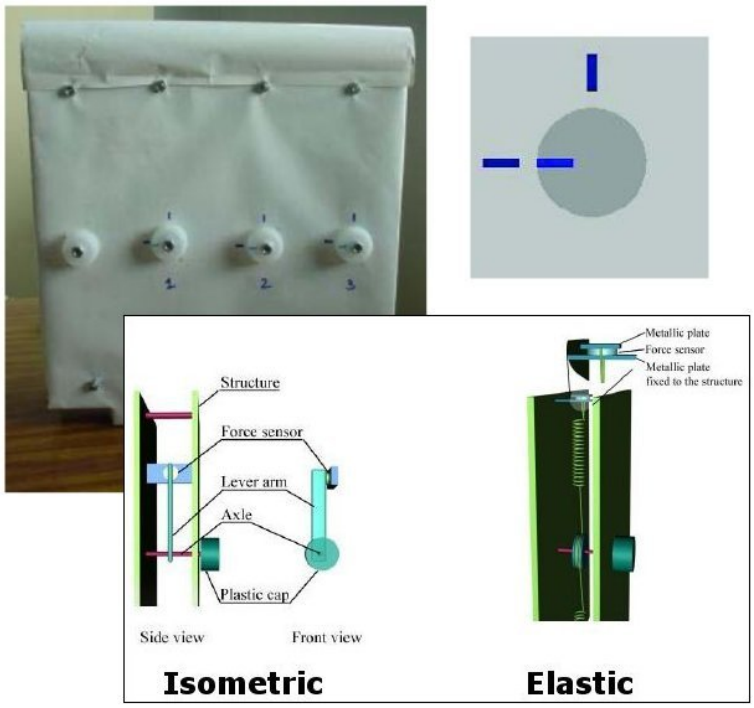


Figure 5-3: Pseudo-haptic torque experiment (see Paljic et al., 2004), device (top left, bottom) and visual feedback (top right).

The authors showed that elastic devices could produce a better resolution (i.e. smaller JND), but a “higher subjective distortion of perception compared to the isometric device”. Paljic et al. (2004) referred to a “perceived mechanical work” cue to explain the source for the subjects' general discrimination ability or even the illusion effect.

A “boundary of illusion” was identified for the simulation of haptic stiffness [Lé01]. The participants were required to compare the stiffness of two virtual springs. These springs were haptically simulated by a active force feedback device (i.e. the PHANToM™) and visually displayed on a 2D computer screen. The control spring behaved in a realistic manner whereas the “pseudo-haptic one” was stiffer at the manual level, but sometimes less stiff in the visual modality. By computing the Point of Subjective Equality (PSE), the boundary of illusion was determined. Contingent on the degree of the sensory conflict, a monotonically increasing “distortion of perception” occurred. It was inferred that “more visual deformation is necessary to compensate large haptic differences”. Furthermore, it was reported that “this boundary varies greatly depending on the subjects and their strategy of sensory integration”.

Pseudo-haptic feedback has also been explored for 2D desktop environments using active mouse cursor displacements [Me02] [Me08] and, moreover, applied to gaming interfaces (e.g. less force-oriented vibrotactile feedback in video game controllers), the simulation of musical instruments (for an overview, see [Ha06]) and to systems for training [Cr04]. The Virtual Technical Trainer (VTT), for instance, permits milling of virtual workpieces (see Fig. 5-4). Different material properties are simulated by a varying velocity of the virtual tool. That is, “a strong resistance of the material is associated with a strong deceleration of the tool on screen.”

Figure 5-4: Virtual Technical Trainer (see Crison et al., 2004),
device (left) and visual feedback (right).

5.3 Technical simulation fundamentals

On top of the basic system named in Section 5.1 reside the visualisation and computation components of the simulation. Visually, a virtual tube object is presented having the upper segment opened (see Fig. 5-5). Through the tube, line particles are flowing meant to illustrate a stream-like phenomenon. Particle size, density, shading, streaming direction and speed are variable. Event cues (e.g. sound on FF entering) can be added. The sensitive region affecting the visual hand can, if needed, adapted to other arbitrary shapes.

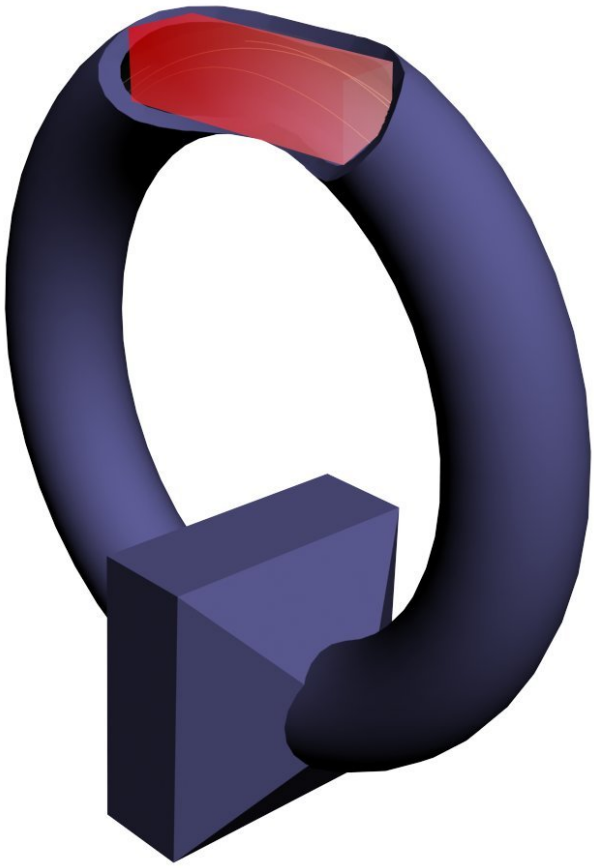


Figure 5-5: Virtual stream tube object,
with the sensitive region highlighted (red).

Since the goal is to visually displace the user's hand on exposure to the FF, the VPC framework has to be supplied with the according hand shift information. To this end, FIRE, the FF illusion and response model, computes at each time step the hand attraction kinematics in the form of the direction vector $\vec{d}$ and the displacement velocity v_{displ} . System response is generated in real time to guarantee a continuous interaction.

Internally, FIRE works as a state machine with the conditional transitions shown in Figure 5-8. The sensory decoupling parameters are dependent on a number of factors, among them perceptual and system constraints, the current simulation state and the user's actual hand movements. Some of the states can be skipped, if, for instance, the performed real hand movements would be excessively fast or far. Otherwise, they are linearly interconnected and will only be parsed again, if the visual hand reenters the FF or changes its movement direction with respect to the FF. To protect or rather regain the simulation range once the hand offset was increased, a link to the hand feedback convergence module (HFC, see Chapter 6) can be established. The reduction of the VPC would also help limiting specific disadvantages of perceptually recalibrating adaptation processes (see Section 2.1).

5.4 Force field illusion and response model

The discussion of the approach is divided into two main parts:

1. Illusion (i.e. theoretical basis of how to induce the force sensation, see Section 5.4.1).
2. Response model (i.e. core algorithm for the FF response behaviour, see Section 5.4.2).

5.4.1 Illusion

To design an effective force illusion setting for VE, the principal question to ask would be: what contributes to comparable force impressions in real life? First of all, a perceptible load needs to be applied to a body part. If, for instance, a book is held in the hand, then its weight can be “felt”. Or, if someone puts his hand into a streaming river, the flow pressure can be “sensed”. In both cases, the sensation can easily be traced back to the effort the motor system has to apply in order to stabilise the involved limb or to transport it towards a target. Paljic et al. (2004) called this cue the “perceived mechanical work”.

The corrective process to maintain a desired postural stability is known as compensatory postural adjustment (CPA, see Wise et al., 2002). That is, a load which perturbs a given state naturally triggers a certain motor reaction. This information combined and integrated with other sensory signals (e.g. skin deformation due to surface pressure, feedback of joint, muscle

and tendon receptors or vision of the causative object, see also Section 2.1) may constitute the impression of the actual force – or even alter it. It can be assumed that, as long as the brain is able to merge given multimodal inputs and the sensorimotor system can adapt to the imposed constraints, an appropriate motor activity could be interpreted as a consequence of an event. This event is perhaps only visually observed, but at no time questioned by the user, neither at cognitive nor at a lower sensory processing level. If the visual feedback of the own actions finally appear plausible (see, amongst others, Blakemore et al., 2003), then an artificial or modified force perception could be elicited.

Triggering a CPA at arm level in conjunction with the visually perceived FF effect (i.e. visual hand drift) should induce such an illusory sensation of force. Haptic devices are not necessary. All VPC treatments have to be well-concealed to preserve visual dominance and thus perceptual reliance. Details on the VPC management and which other factors FIRE has to consider will be discussed in Section 5.4.2.

The following two scenarios for an interaction with the virtual FF demonstrate how the force sensation is expected to occur:

1. *Hand stabilisation within the FF.*

   In this scenario, the user tries to resist a simulated force by keeping his hand at a certain position within the FF. At the moment the visual hand involuntarily starts to move along with the flow, a CPA is triggered. In order to visually stabilise the hand, the user will compensate for the displacement by unconsciously moving his real hand in the opposite direction of the flow (see Fig. 5-6). This motor effort (i.e. muscle work) is integrated with an almost stationary visual representation of the hand. The illusion of a force to resist might occur as long as the generated feedback remains reliable

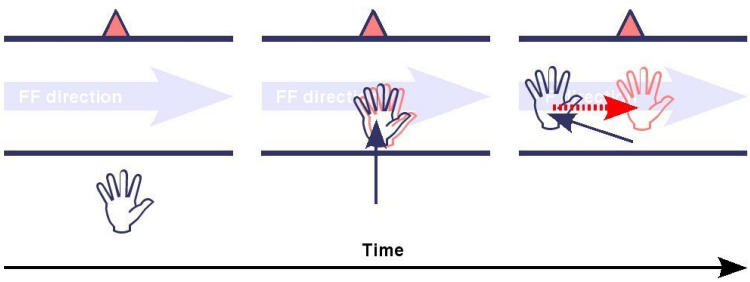


Figure 5-6: Hand stabilisation within the flow (dark blue: real hand, salmon: visual hand, stabilisation area indicated by salmon triangle).

2. *Hand movement along with the FF.*

   In this second scenario which adopts the already described principles, the user moves his real hand voluntarily along with the flow. The effect of a faster visual hand motion should be interpreted as a movement support provided by the FF. That is, the user could get the impression of an easier hand movement requiring less muscle work.

The impact of the illusion will most likely be influenced by the user's reliance on what the system visually feeds back. Hence, a convincing simulation within the given constraints is envisaged. This includes the properties of the hand representation and the FF response dynamics. Once the user has exposed his real hand to the virtual FF, he should believe in what he feels rather than questioning his perception.

5.4.2 Response model

The model's general mode of operation in terms of data flow and VPC handling is addressed in Section 5.3. Here, attention is turned to the core algorithm for the FF response behaviour. Four points are taken into account for the computation of the visual hand displacement:

1. Perceptual constraints (see Section 5.4.2.1).
2. System constraints (see Section 5.4.2.2).
3. Force field properties (see Section 5.4.2.3).
4. Adaptation to hand movements (see Section 5.4.2.4).

5.4.2.1 Perceptual constraints

This section emphasises factors which could, if not respected, cause an undesirable break in the visual capture and therefore in the illusion. For a comprehensive overview of the known perceptual issues recommended to be considered in FIRE (e.g. maximum hand offsets, co-located interaction, real hand embedding), refer to the Sections 2.1 and 2.2.

Many things are still unclear, though. In particular, the perceptual and behavioural effects of dynamic VPC have not yet been studied under the conditions relevant for this work. Several informal experiments were conducted to encircle applicable best practice measures. Tasks consisted of passive hand drift observations with and without attentional distraction, perturbed straight-ahead pointing and various spatially manipulated active hand movements. After all, the favourable displacement dynamics as they will be used in FIRE seem to be:

- Effective displacement velocity range: approximately 3 – 8 cm/s.
- Maximum displacement acceleration: approximately 10 cm/s².

5.4.2.2 Device constraints

This class of problems is generally related to current VST HMDs. Due to the limited field of view of the attached or built-in cameras (see Fig. 2-5, top right), the maximum possible deviation of the user's hand from the video capturing centre is predetermined. So, if the real hand disappears from the cameras, it cannot be registered anymore. Consequently, the hand texture carrier objects cannot be supplied with the live image data. The properties of the device used so far are depicted in Figure 5-7 and assigned to the VPC framework (see Section 3.4). Based on an assumed average interaction distance of about 50 cm to the virtual FF, the maximum offset o_{max} according to the smaller leftwards deviation threshold should not exceed 17 cm for horizontal VPC because of $o_{max} \leq \sin(20°) \cdot 50\,cm$. In fact, o_{max} has been set to 15 cm for the simulation.

As already mentioned in the previous chapters, some other drawbacks come along with the limited display and camera quality (e.g. in resolution, brightness and colour consistency) and the often narrow field of view of the displays themselves. But since the critical bottleneck is the field of view of the cameras, only o_{max} is regarded as a constraint.

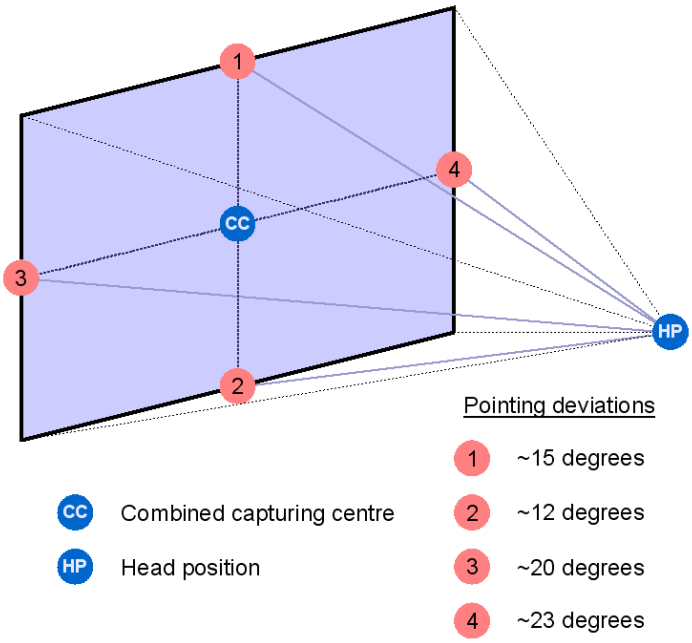


Figure 5-7: VST deviation angles for a just visible pointing.

5.4.2.3 Force field properties

In order to create a convincing system response, a few basic principles for an interaction with a stream-like phenomenon should be adapted. Looking at the river example of Section 5.4.1, what happens in terms of action and reaction, if someone exposes his hand to flowing water? Suppose the flow is strong enough, then the hand gets attracted and pushed to the side. Further suppose that the person wanted to feel or resist the steaming force, then a voluntary CPA would activate different muscles from shoulder to forearm and hand just to overcome the given flow pressure. Due to visuomotor control loops, this muscle work will finally result in a more or less stabilised hand. Only turbulences or swirls within the flow may lead to little position instabilities. If the hand is taken out of the stream, muscle pretension directed against the prior quasi-steady force would cause hand positioning errors similar to those when putting the hand into the stream. Thanks to the readjustment processes mentioned before, the hand will nevertheless quickly be stabilised again.

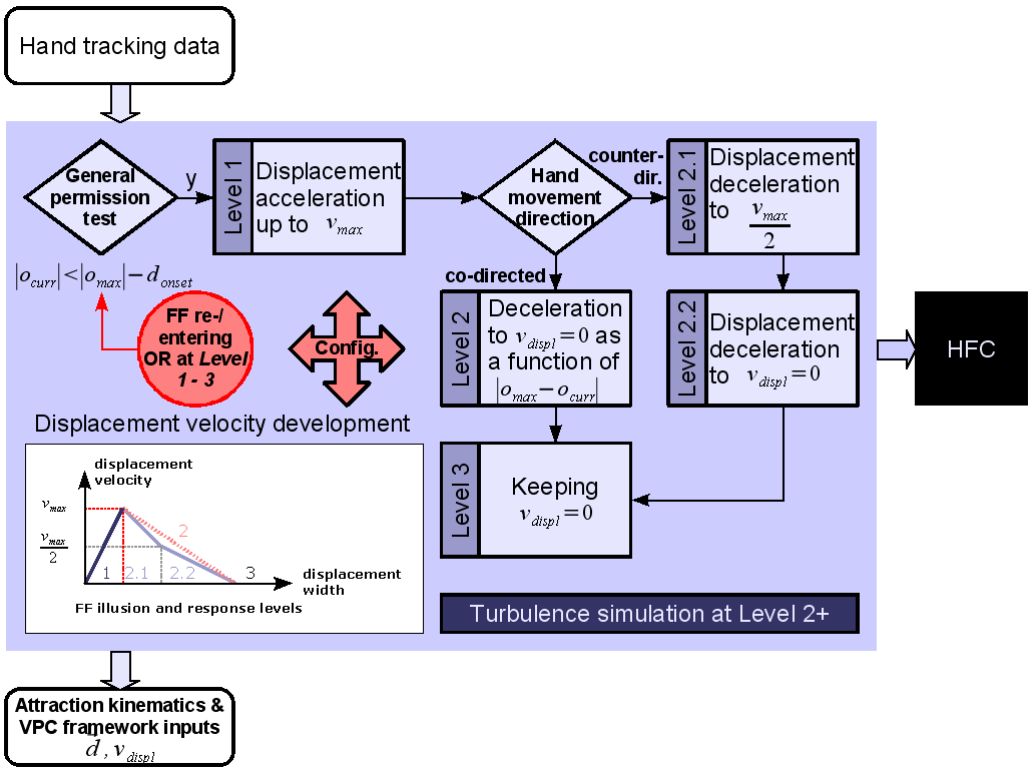


Figure 5-8: FIRE (i.e. FF illusion and response model), with all internal state levels and the displacement velocity development with level correspondences (bottom left).

As illustrated in Figure 5-8, the state *Level 1* represents the illusion onset phase. It is triggered at the moment the real hand is put into the FF. For each FF strength level F, there exists one onset displacement distance d_{onset} within which a related maximum displacement velocity v_{max} will be reached. That is, the visual hand will be accelerated and shifted along with the flow, away from the real hand's position (see Fig. 5-6). During this phase, at least in the hand stabilisation scenario, the user's reaction (i.e. CPA) might be the most intense, if not the most important. The visual onset drift is meant to provoke a motor reaction globally reflecting the effort to resist the induced force. d_{onset} and v_{max} can hence be understood as the key response parameters or primary FF properties. To support visual capture, a turbulence effect (i.e. minor positional oscillation of the visual hand) is activated after the onset phase.

A supplementary set of parameters is required for practical reasons. One can imagine that several problems would emerge, if the drift velocity would be kept at a certain value once the onset phase was passed. The user would need to move his real hand further into the opposite direction of the FF to see his hand visually stabilised at the desired position. This will not be

possible for a longer time. Because at some point, due to the limited field of view of the VST HMD's cameras, it will not be possible to capture the real hand anymore. Also, the illusion would most likely break under extreme conditions, for instance, when the user points far sidewards while he still sees the visual hand in front of him (i.e. stressed or overstrained sensory integration, see Section 2.1). One opportunity to tackle this issue is to keep the user's hand within the field of view of the cameras by fading the displacement velocity out. FIRE provides a hand-movement-dependent adaptation mechanism which does exactly that (see Section 5.4.2.4).

Even if not explicitly studied in the course of this chapter, the response model is prepared to handle two additional interaction cases. First, for a continuous sweeping through the virtual FF, the displacement can be reinitiated as long as $|o_{curr}| < |o_{max}| - d_{onset}$. Second, if either the displacement limit is reached or the visual hand leaves the sensitive FF region, then FIRE establishes a connexion to HFC. The overall purpose of this operation is to reduce the offset between the real hand and its visual counterpart.

5.4.2.4 Adaptation to hand movements

The system response has to adapt to the user's hand movements for both perceptual and practical reasons. However, none of the treatments should distract the user's attention from the interaction with the FF. This also requires that the user cannot directly see any part of his real arm or hand.

The adaptation consists of two methods (i.e. states *Level 2*, *2.1* and *2.2*, see Fig. 5-8):

1. *Hand stabilisation and movements against the FF.*

   The displacement velocity fade-out was originally based on a negative exponential function, but got later transformed into a more appropriate reciprocal derivate. The 1D distance of the real hand from the FF entry point along the FF axis served as function value whereas the parameters depended on F (see Section 5.4.2.3). This yielded a progressive velocity reduction reflecting the user's “work applied to the system”. In some preliminary tests it was found that the user's sensitivity to the fade-out function shape was not very high. Thus, a linear fitting simplification was chosen. The model

has now two major fade-out steps integrated (i.e. states *Level 2.2* and *3*, see Fig. 5-8), with the first and faster deceleration down to $v_{max}/2$ and the second and slower deceleration down to a displacement velocity of 0. The widths of the fade-out steps (i.e. 1D distance state transition thresholds) scale with F :

$$scale_i = \frac{v_{max}(F_i)}{v_{max}(F_{ref})} \qquad \text{(Eq. 5-1)}$$

With v_{max} set to 5 cm/s for the reference force F_{ref} , $scale_i \leq 1.5$ and the initial fade-out or transition condition step widths to undergo the rescaling for each F_i :

- State 1 (i.e. offset acceleration): 2.5 cm.
- State 2 (i.e. offset deceleration): 5 cm.
- State 3 (i.e. static offset): 10 cm.

2. *Hand movements along with the FF.*

   If the user performs an FF-directed hand movement, then the displacement deceleration after the onset phase can be less strong. In fact, there is no real "force to overcome" for the user and the displacement (i.e. more precisely: hand movement support response) could be applied for the whole time the visual hand stays within the FF. The only limits are the maximum allowed displacement distance (i.e. adaptation threshold to state *Level 3*, see Fig. 5-8) and the FF object boundaries. The displacement velocity is reduced as a function of the "yet allowed VPC growth" $|o_{max} - o_{curr}|$.

The state *Level 3* of FIRE keeps the displacement velocity at 0 until another valid trigger event appears (see Fig. 5-8). Validity is given as long as a sufficient shifting space remains for the onset phase before o_{max} is reached.

If the hand movements are too short, that is, when there is not enough "energy applied" by the user to counteract the FF, then the visual hand will be transported until the end of the sensitive region and finally stopped or "blocked" at its boundaries.

Once the visual hand is taken out of the stream, then a forced position error is introduced by quickly reducing o_{curr} to 85% of its former value. This simulates the muscle pretension effect described in Section 5.4.2.3 and contributes to a first decrease of the VPC.

5.5 Hypotheses

The experimental evaluation of the proposed HEMP approach will be designed based on two questions. First, can FIRE generate different F -dependent levels of muscular activity in the main actuator for an arm-movement-driven hand stabilisation task? This is assumed to be an essential prerequisite for an alterable pseudo-haptic percept with constant visual feedback (i.e. invariant visual flow properties, see Section 5.3). Second, does the illusion finally occur and are several FF levels discriminable?

Hypothesis 1 therefore focuses the ability of FIRE to stimulate the desired variable motor response. Suppose the right hand is successively placed into a rightwards streaming virtual FF of different intensities with the goal of stabilising the visual hand at an indicated location. Then an increased average electromyographic (EMG) activity should be measurable in the principal shoulder traverse flexor (i.e. pectoralis major, see Fig. 5-12). This muscle effectuates the required leftwards compensatory movement of an almost outstretched ahead pointing arm. Specially the onset phase (i.e. displacement acceleration) seems to be eligible for causing a higher average activity, since a stronger contraction is needed to perform a faster CPA. Just a longer hand movement distance due to the scaled fade-out step widths is not likely to produce greater changes in the average activity. All should work without haptic devices.

Hypothesis 2 deals with the conscious perception of the illusory force impression. Based on the assumptions that the employed system can preserve perceptual reliance at a sufficient level and that the integration of plausible sensory inputs can contribute to an altered percept, it can be expected that the illusion occurs in most of the subjects. The restriction “most” is made, because it has been repeatedly shown that the impact of pseudo-haptics is largely dependent on the subjective strategies for assessing, processing and integrating conflicting sensory information. However, in case the muscle activity induced by FIRE is adequately variable, then subjects should be able to discriminate given FF levels with higher accuracy the larger the presented intensity differences are. Further, sensing stronger FF levels might benefit

from an improved signal-to-noise ratio (SNR) in the triggered reaction (i.e. higher average activity). The increased SNR should result in a better discrimination performance, if higher F are involved.

5.6 Experiment

The main purpose of the experiment was to test the hypotheses (see Section 5.5) and to study the potentials of HEMP. Subjects had first to perform a forced-choice FF strength comparison task (see Section 5.6.3.1). After, an illusion evaluation questionnaire was given (see Section 5.6.3.2).

5.6.1 Subjects

Thirteen healthy adult volunteers (i.e. 18 – 55 years old, 4 female and 9 male) participated in the experiment. Ten of them were right-handed and 3 ambidextrous with a large right hand usage in their everyday life. None of the participants suffered from serious vision problems. Corrected vision was not considered to be problematic. Ten have never used a comparable setup. Most were even completely new to AR / VR. The rest had either attended an AR / VR class at the university or experienced VE in demos, workshops or beside their professional work. The subjects were not aware of the goal of the experiment.

5.6.2 Factorial design

Subjects had to perform a repetitive pairwise comparison task, with 5 FF strength levels presented (i.e. F_1 to F_5, corresponding to $v_{max}(F_i) = \{3.57, 4.23, 5, 5.92, 7\}\, cm/s$ with preserved relative difference). Rescaling of the FIRE state transition conditions is based on F_3 as the reference force (see Eq. 5-1). Since a forced-choice protocol was used, all cases of equal forces were excluded, resulting in 20 instead of 25 combinations. Several informal tests revealed a high discrimination performance (i.e. high success rates), when very large force differences were compared. The set of combinations was hence reduced by those exceeding a 2-level difference. While considering only one FF direction for the comparison

(i.e. flow to the right), the final design comprised 14 conditions (C_1 to C_{14} , see Table 5-1) organised in single factorial vector.

Table 5-1: Condition square for the comparison task: greater force was either presented in the first trial (condition 8 to 14) or in the second trial (condition 1 to 7).

trial 1 \ trial 2	F_i	1	2	3	4	5
	1		1	2		
	2	8		3	4	
	3	9	10		5	6
	4		11	12		7
	5			13	14	

Distraction trials (i.e. flow to the left, FF properties as for F_3) were introduced to limit adaptation processes. Within a complete set of 14 comparisons (i.e. 28 trials), there were 8 such distraction trials included, but never in between two trials to be compared.

The comparison conditions and distraction trials yielded a total of 22 conditions to be equally weighted, randomly distributed and protected against effect carry-over. A 22 x 22 random latin square was employed for the condition balancing. With 12 repetitions per condition, each participant had to complete 168 comparisons.

5.6.3 Procedure

Subjects were first asked to fill in a general information form (i.e. about age, gender, handedness, prior knowledge on AR / VR etc.). After, the main experiment began with the comparison task (see next section) and ended with the illusion evaluation questionnaire. The whole comparison task was divided into four parts (approx. 18 to 20 minutes each) so that subjects had some time to rest and recover.

5.6.3.1 Force field strength comparison task

Each comparison consisted of two consecutive trials of reaching movements. Initiated by an acoustic signal (i.e. beep) and starting from the rest position (see Fig.5-9, left), subjects had to move their right hand into the target area. The target itself was the opened upper segment of a virtual stream tube object located approximately 45 cm in front and at the shoulder height of the subject (see Fig. 5-10). There were particles flowing through the tube, always at constant velocity and appearance. A little red sphere specified the horizontal region of the flow to approach. The hand tracking body was fixed by rubber bands so that subjects did not need to actively hold it, for instance, by grasping (see Fig. 5-9, right).

Figure 5-9: Rest position (left), hand tracking body fixation (right).

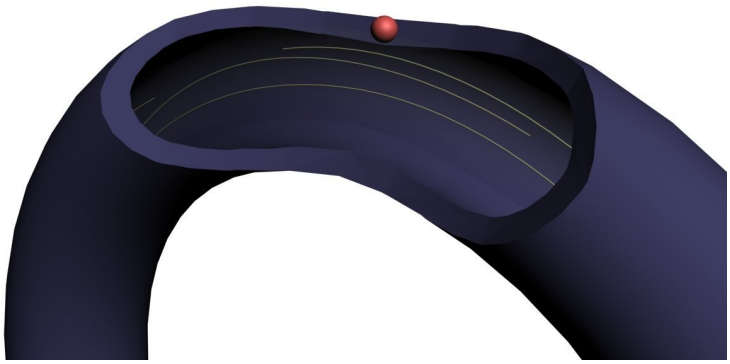


Figure 5-10: Close-up of the stream tube object, with particle flow and the hand stabilisation region indicator (i.e. red sphere).

Once they had entered the FF, participants were instructed to try to keep their hand within the stream at the open part's centre (i.e. red sphere, see Fig. 5-11). The hand stabilisation scenario was chosen for both its action clarity and its control practicability. Moreover, the FF simulation characteristics required relatively slow reactive hand movements. Systems lags, mainly the video latency (see Section 3.4.2), played thus only a minor role. Another beep, 6 seconds after the first, marked the end of a single trial and subjects had to return their hand to the rest position. One trial took about 10 seconds and two such trials made up one comparison unit. After each comparison, subjects had to identify the trial in which they found it harder or had to make a greater effort to hold the hand at the desired position. Oral responses were registered by the experimenter (i.e. "1" for $F_{T1} > F_{T2}$ and "2" for $F_{T2} > F_{T1}$).

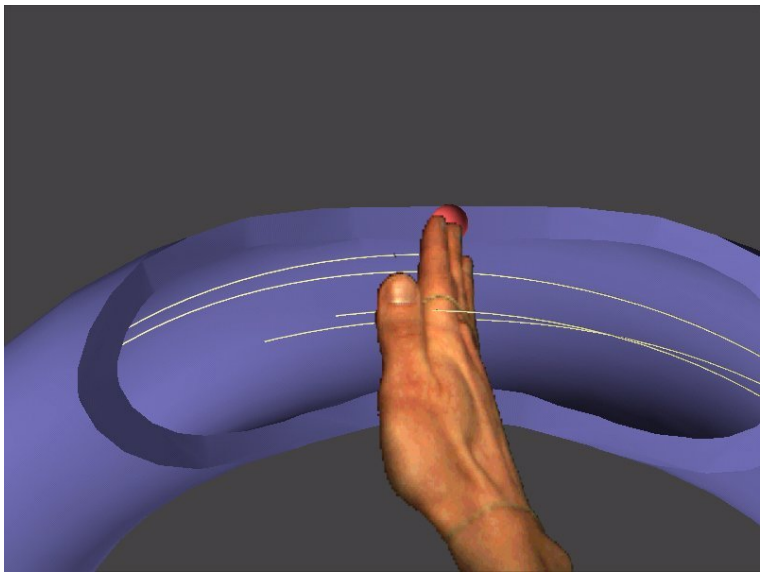


Figure 5-11: Application screenshot of a user reaching into the FF (i.e. VST HMD view).

5.6.3.2 Illusion evaluation questionnaire

The illusion evaluation questionnaire (see Appendix 2) was given after the comparison task was done. It contained 7 questions, mainly designed to get a first impression of the actually induced illusion and to improve understanding of the comparison results.

Subjects were asked to describe their sensation when having the hand exposed to the visual stream and whether this sensation has changed over time. They were further asked to note all differences they perceived during the comparison task as a whole. On a 7-level scale, subjects

had to indicate the extent to which the VR experience correlated with any of their real world experiences (i.e. from 1: inconsistent to 7: consistent, taken from the Presence Questionnaire [Wi98]). This assessment, meant to help situating the evoked sensation between being purely artificial and real, had to be explained. Finally, it was asked for the cues subjects used for their comparison judgements.

The form closed with questions about possibly perturbing factors, if any, and allowed for general remarks.

5.6.4 Data acquisition and analysis

The first quantitative data basis for the analysis was EMG data acquired from the pectoralis major. Skin electrodes have been placed along the muscle fibres at the electric emission spot (i.e. thickest muscle section, see Fig. 5-12). Data was recorded at 1 kHz, rectified and subdivided according to the task phases. The main phase used for the analysis began with the moment the visual hand had entered the FF. It ended as soon as the hand stabilisation criterion was fulfilled (i.e. local real hand movement velocity minimum due to state *Level 3* of FIRE).

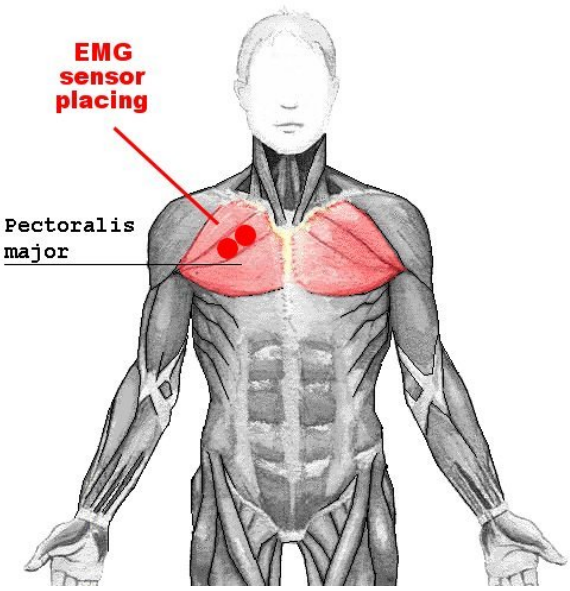


Figure 5-12: Musculus pectoralis major and the EMG sensor positions chosen for the experiment.

The comparison responses (i.e. success rates) served as the second quantitative data source. They were grouped in different ways to isolate a number of specific effects. Groupings for the discrimination analysis were:

1. *Overall condition ranking (see Section 5.7.2).*

   Gives an ordered listing of all presented conditions, providing a global image of the comparison performance.

2. *Force combination differences (see Section 5.7.3).*

   This effect grouping is based on the force level difference between the two trials of a pair (i.e. 1- or 2-level difference).

3. *Force combination zones (see Section 5.7.4).*

   There are 7 possible zones for combining the existing force levels (i.e. Z_1 to Z_7). With an increased zone ID, the force levels to be compared are higher, too. That is, regarding the condition square (see Table 5-1):

   - $Z_1 : C_1 \text{ and } C_2$
   - $Z_2 : C_2 \text{ and } C_9$
   - $Z_3 : C_3 \text{ and } C_{10}$
   - $Z_4 : C_4 \text{ and } C_{11}$
   - $Z_5 : C_5 \text{ and } C_{12}$
   - $Z_6 : C_6 \text{ and } C_{13}$
   - $Z_7 : C_7 \text{ and } C_{14}$

4. *Force combination senses (see Section 5.7.5).*

   This last grouping basically focuses on an interesting subquestion: how are responses affected when the greater force level appeared either in the first trial (i.e. $F_{T1} > F_{T2}$) or in the second trial of a pair (i.e. $F_{T2} > F_{T1}$)?

Subjective results about the participants' conscious perception of the virtual FF were obtained from the illusion evaluation questionnaire.

For each trial, head and hand tracking data and VPC information was recorded at 60 Hz.

The quantitative data was analysed using classical statistics. That is, a descriptive part (i.e. mean and SD) was followed by repeated measures ANOVA, Pearson's product moment correlation tests and, if appropriate, post-hoc analyses (e.g. Bonferroni-corrected LSD).

5.7 Results

The behavioural and comparison response results will be presented first. After, subjective questionnaire statements are processed (see Section 5.7.6).

A few comparisons in which subjects were not able to give any response had to be excluded from the analysis (data loss rate: approx. 2.24%). Beside for technical reasons, responses were mainly not given, for instance, in cases of strong uncertainties or if subjects had lost their concentration on the task.

5.7.1 Pectoralis major activity

The analysis of the arm EMG activity associated with the FF effect compensation behaviour focused on the right pectoral muscle. Its main adductor role at shoulder level is to move the arm laterally from right to left relative to the sagittal plane of the body. Results show that the average activity of this muscle increased significantly with the FF levels when the hand was put into the virtual FF ($F(4,48) = 6.25$; $p < 0.0004$ and $r^2 = 0.83$; $t = 3.78$; $p < 0.033$, see Fig. 5-13).

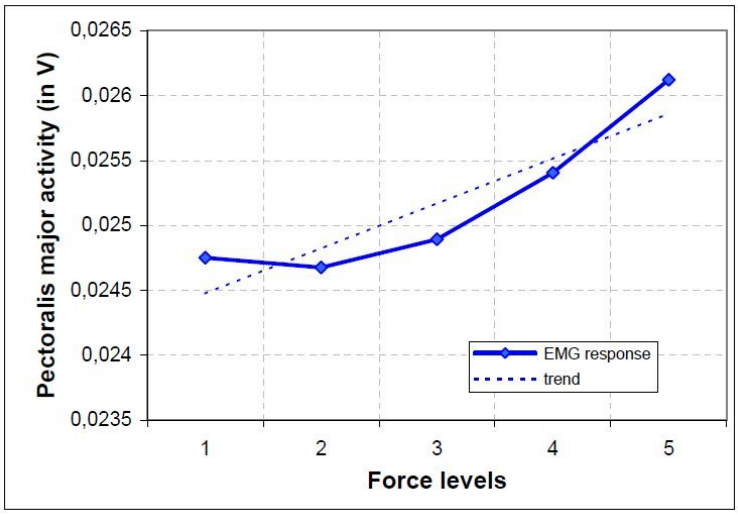


Figure 5-13: Influence of the simulated FF levels
on the mean pectoralis major activity.

It should be remarked that the activity curve seems to have an exponential shape (i.e. $r^2 = 0.83$; $t = 3.82$; $p < 0.032$). This might be due to the FF parameters applied during the experiment. As mentioned in Section 5.6.2, v_{max} varied with an equal relative difference, thus implying an exponential regression.

5.7.2 Overall condition ranking

In Table 5-2, all conditions are shown (i.e. mean in %, "1" = 100%, and SD), ordered by their response success rate. The related main variables (i.e. force combination differences, zones and senses) are also indicated. The particular impact of each on the results will be presented in the next sections.

Table 5-2: Condition ranking and overall results. Each rank shows: the condition with respect to the condition square (cond.), the condition-dependent force level difference (diff.), zone and sense as well as the related mean success rate and SD.

rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14
cond.	**6**	**4**	**2**	**13**	**7**	**9**	**5**	**11**	**3**	**10**	**14**	**12**	**1**	**8**
diff.	2	2	2	2	1	2	1	2	1	1	1	1	1	1
zone	6	4	2	6	7	2	5	4	3	3	7	5	1	1
sense	2	2	2	1	2	1	2	1	2	1	1	1	2	1
mean	**.883**	**.863**	**.837**	**.830**	**.814**	**.760**	**.759**	**.734**	**.723**	**.689**	**.640**	**.639**	**.628**	**.540**
SD	.113	.110	.159	.128	.141	.156	.208	.167	.164	.218	.190	.194	.126	.195

5.7.3 Force combination differences

In the experiment, only 1- and 2-level differences were presented (see Section 5.6.2). The comparison performance found for 2-level differences (i.e. displacement peak velocity ratio: 1.4, mean: 0.818, SD: 0.068) was significantly better ($F(1, 12) = 49.72$; $p < 0.0001$) than for 1-level differences (i.e. displacement peak velocity ratio: 1.18, mean: 0.68, SD: 0.088). Figure 5-14 shows these results.

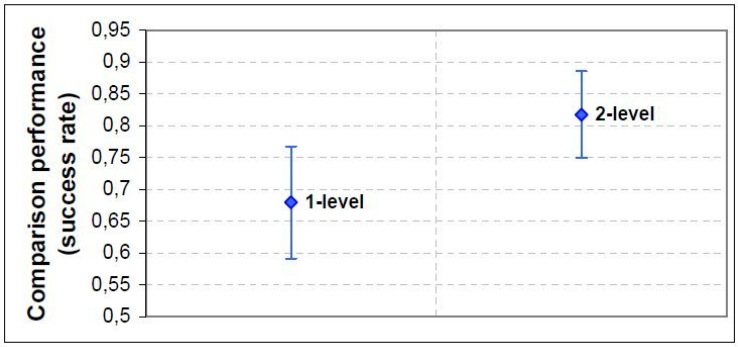


Figure 5-14: Influence of the force combination differences on the comparison performance (means, SDs).

To determine whether the mean for 1-level differences was nonetheless dissimilar from chance, it was compared to a hypothetic 50% success score. A separate variance estimate t-test was employed which actually revealed a significant difference from chance ($t = 7.369$; 2-sided $p < 0.0001$).

5.7.4 Force combination zones

This grouping is based on the 7 force combination zones described in Section 5.6.4. The fact that the force combination differences are inherently coded in the force combination zones, made an effect on the comparison performance expectable ($F(6, 72) = 10.89$; $p < 0.0001$). To analyse whether the subjects' performance correlated with the presented force combination zones, the Pearson's product moment correlation test was applied. A positive correlation was found ($r^2 = 0.04$; $t = 2.82$; $p < 0.006$, see also the trend curve in Fig. 5-15).

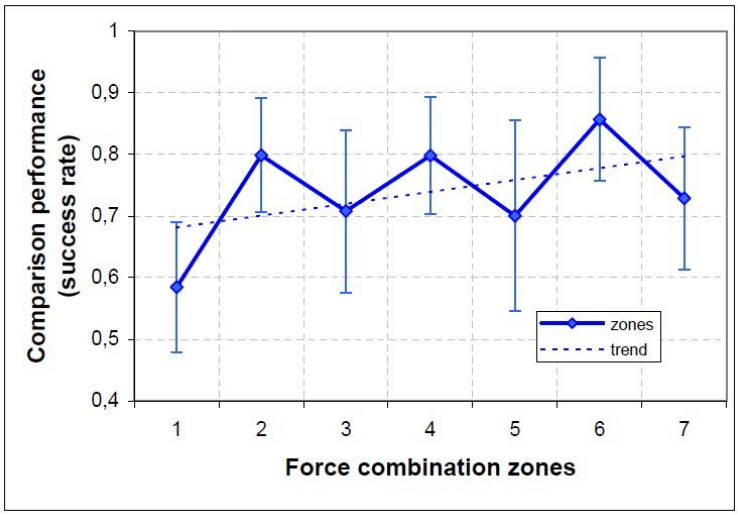


Figure 5-15: Influence of the force combination zones
on the comparison performance (means, SDs).

But when additionally considering the two force combination senses, then this correlation has to be assess separately. That is, if the greater force was presented in the second trial of a pair, a significant correlation was found ($r^2 = 0.09$; $t = 2.91$; $p < 0.005$) whereas in cases with the greater force appearing in first trial such correlation did not exist ($r^2 = 0.02$; $t = 1.37$; $p > 0.1$).

5.7.5 Force combination senses

For the tested force combination senses, a significant effect on the comparison performance was found ($F(1, 12) = 5.58$; $p < 0.036$, see Fig. 5-16). That is, for a greater force simulated in the second trial of a pair (i.e. mean: 0.787, SD: 0.09), the comparison results were better than if the greater force was presented in the first trial (i.e. mean: 0.69, SD: 0.095).

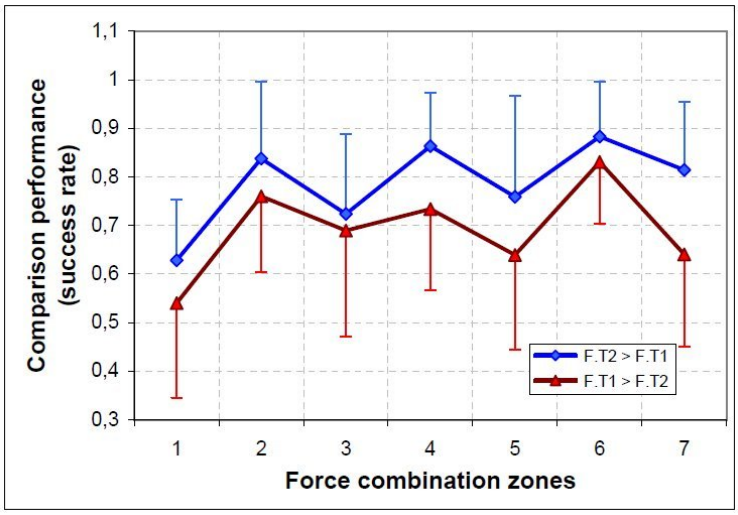


Figure 5-16: Influence of the force combination senses
(red: greater force in the first trial, blue: in the second trial)
on the comparison performance (means, SDs).

While further regarding the subject-wise response quantities for the two force combination senses, then a significant effect is exhibited ($F(1, 24) = 11.41$; $p < 0.003$) showing that subjects rated the second trial as stronger (i.e. mean: 0.55; SD: 0.75) more often than the first trial (i.e. mean: 0.45; SD: 0.75).

5.7.6 Subjective illusion evaluation

The goals of the illusion evaluation questionnaire were to derive the conscious sensations participants had when exposing their hand to the virtual FF, to yield a score for how close these impressions were to real world experiences and to identify possibly disturbing or confusing factors in the simulation.

Concerning their sensations, almost 70% (i.e. 9 out of 13) of the subjects reported, for instance, “a force that you were obliged to stem against with your hand”, “a flow more or less strong as water”, “a force that obliged me to compensate with my muscles”, “a pressure exerted by the flow, tearing the hand away”, “felt that the flow was pushing my arm”, “got the hand pushed to the side, requiring (muscle) tension to resist” and “the flow seemed to push my hand away” or similar statements. Seldom, subjects noted that the sensation diminished over time. Most did either not perceive or not remark any change.

The consistency of the VR experience with something that was already experienced in the real world got rated with 4.54 out of 7 possible points. Subjects explained their sensations, for instance, as “a driving force”, “water stream”, “air stream as perceived when holding the hand outside of the window of a driving car”, “air blowing on my hand”, “holding the hand into some stream”, “putting the hand [...] under an air head-dryer” and “the sensation was very real”. But there were also statements like “only few situations in which we see our hand going away and need to apply a specific force to control it”, “no tactile sensation” or “feeling as if I would have lost my 'tactile' sensation”.

The strategies to come to a comparison judgement seemed to differ among the subjects. Two groups could be separated of which the first relied more on their sensations, that is, “by the feeling of how much I had to force my hand”, “by the feeling of heaviness”, “by muscle contraction and the contraction duration”, “by the first moment of the force impact”, “by the resistance I had to oppose”, “by how much effort I had to put”, “I just 'felt' the strongest flow” or “whether I felt my pectoral muscle fatiguing”. The second group based its judgements mainly on observations of the scene and / or their own actions, as “by the position and the movement of the hand”, “by the displacement of the hand before my reaction”, “by how far I pushed my hand to the left”, “in the strongest I saw my hand shifted more” or “the more separated from the center it was, the more force I had to do to get it (the hand, author's note) back”.

The duration of the experiment and the ergonomics of the see-through HMD were partially remarked as uncomfortable.

5.8 Discussion of the experiment

The hypotheses tested via the experiment concentrated on the appropriateness of FIRE to provoke different levels of muscular activity in the horizontal arm adductor and, secondly, on the resulting perception of the virtual FF in terms of pseudo-haptics. To recall both main questions, see Section 5.5.

In *Hypothesis 1*, it was assumed that FIRE could innervate different mean EMG intensities in the most important muscle for performing the FF strength comparison task. Indeed, data shows that there was an increased average response, thus a stronger contraction provoked in the pectoralis major with higher F simulated. But this activity may have not been caused in the first place by the longer hand movement distances (i.e. scaled fade-out step widths). The dynamics of the onset phase and therefore the primary FF properties (i.e. d_{onset} and v_{max}) were a lot more qualified to trigger a CPA at varying intensities. Faster hand shifts had to be compensated by participants. While the general slope of the EMG activity (see Fig. 5-13) can be explained by the nature of the F-based simulation parameters, the lack of a clear difference between the first two FF levels is not so obvious. Perhaps the muscle responses did not differ much between F_1 and F_2, because a lower bound of the excitation range was found. Consequently, the reduced comparison performance shown in the overall condition ranking could be expected (see Table 5-2). However, the initial hypothesis holds true for most part. Restrictive elements to include would be the minimum FF properties (e.g. those of F_2). For further details about future improvements, refer to Section 5.9.

Hypothesis 2 focused on the participants' conscious perception. In case a convincing and perceptually plausible multisensory supply were provided, then a force illusion should have occurred, at least in the majority of the participants. There were two sources of data, first, subjective illusion experiences and second, the quantitative comparison response data. To answer to the question one, whether there is any sensation of force, the statements made in the illusion evaluation questionnaire appear to be promising pointers. Close to 70% of the subjects felt as if their hand was “pushed” by the flow or as if there was a force “exerted” on their hand. The reported comparison judgement strategies seem to support these subjective impressions. When looking at the virtual-to-real world experience consistency score, a light tendency towards a more realistic than a purely artificial sensation might exist. In any case, it is not impossible though that the statements were, to some extent, primed by the experimental

procedure. In future studies, this risk should be minimised, for instance, by concentrating on behavioural or physiological aspects of the illusion. Regarding the second question about the ability of FIRE to evoke different pseudo-haptic force levels, the user can globally be provide with the required distinguishable sensory information. That is, significantly higher success rates were found for:

1. 2-level differences suggesting that larger differences would perform even better.
2. Greater force levels involved in a comparison and correlated with an increased F .

These results were generally expected and so confirm the initial hypothesis. But the number of subjects to which the illusion was effectively conveyed and the discrimination performance strongly imply the need for system and model enhancements (see next section). The degree of ambiguities in the sensory signals due to an incomplete visual capture might have been one of the critical issues beside the above-mentioned motor activity evocation constraints. After all, it should be remarked that lower scores were mainly found when smaller hand motion differences and / or slower hand movements were provoked. It could hence be worthwhile to have a closer look at the sensitivity of the proprioceptive system under the imposed conditions (see also Scheidt et al., 2005). A comparison to natural flow phenomena (e.g. modified air compressor) or a physically correct haptic rendering may help in this respect.

The effect of the force combination senses on the comparison performance (i.e. $p < 0.036$) and, more subtle, on the force combination zones results (i.e. no correlation for $F_{T1} > F_{T2}$) did most likely emerge because of the observed judgement preference. However, it is difficult to isolate the origin of this response bias. Possible that the duration of a comparison (i.e. approx. 20 sec.) and the HEMP approach by itself have contributed to it. Information decay in short-term memory and a potential action representation interference due to the dynamic VPC could have given a stronger weight to the last and clearer memorised trial. To clarify this point, additional fundamental investigations on the impact of VPC on perception and action are necessary.

5.9 Conclusion and future work

A novel active pseudo-haptics technique has been proposed in this chapter. The system used for the simulation is based on a VST HMD AR hard- and software platform extended by the VPC framework (see Chapter 3). FIRE, the force field illusion and response model, has been developed on top of this system. Internally, it operates as a state machine consisting of an illusion onset state and a number of displacement velocity reduction states. For the actual VPC computation, FIRE takes the primary FF properties into account (i.e. d_{onset} and v_{max}) as well as the current simulation state by considering the user's "work applied to the system" (i.e. FF effect counteraction). Additionally, there are several perceptual and device constraints having an impact on the computation (e.g. applicable displacement dynamics and the limited video capturing space).

The approach relies on two major assumptions. First, the visually presented phenomenon should theoretically be capable of returning some kind of pushing force. To this end, an easily understood virtual FF with parameterisable features is simulated. Second, appropriate haptic information related to this phenomenon has to be provided to the user. Since neither active nor passive haptic devices are employed, a muscular carrier activity is stimulated. This is done by triggering a specific motor reaction once the user has put his hand into the sensitive region of the virtual FF. The hand starts visually to drift along with the flow. As a reaction to this apparent visuomotor perturbation, a CPA at arm level is actuated. Now, by automatically transporting the real hand against the FF, tone is increased in the respective muscles. Voluntary sweeping movements would eventually suffice, too. If the well-defined actions can successfully be coupled with a plausibly manipulated positional appearance of the hand, then the user's CNS may adapt to the new, the artificial multisensory supply – and the illusion occurs.

An experiment has been carried out to test this hypothetical construct both at a system response and a subjective sensation level. Participants had first to perform a forced-choice pairwise comparison task with 5 different FF levels presented. Each F was characterised by a particular combination of d_{onset} and v_{max} which themselves were chosen to preserve an equal relative difference. These parameters were meant to induce different levels of activity in the main shoulder traverse flexor while the FF visualisation never changed. It was expected that this constellation should lead to different pseudo-haptic FF strength impressions. The

analysis of the pectoral muscle EMG data shows a significant influence of the FF levels on the mean contraction intensity during exposure. FIRE is thus able to selectively modulate motor activity. However, for lower F, this effect seems to disappear suggesting that lesser accentuations between smaller adjacent force grades are not useful. The questionnaire results demonstrate that most subjects consciously experienced the pseudo-haptic FF. They often described it as a “stream-like” or “pushing” force. The presented force levels were globally discriminable at an higher accuracy the larger the differences and the stronger the simulated forces are. Nonetheless, a large gap in the success rates was revealed (i.e. from 54 to 88.3%).

The list of potential future work aspects in this novel 3D UI niche is long. It ranges from system improvements (e.g. using more ergonomic devices with better characteristics) to investigations on perceptual and behavioural backgrounds of the employed dynamic VPC, but also the exploration of actual applications and the synthesis of a theoretical model of pseudo-haptics. Some of these points represent challenging engineering problems, others intriguing research questions at a multidisciplinary scope. The FF approach itself can be ameliorated in several ways. For instance, the illusion quality depends to some extent on the visual fidelity of both the user's limb and virtual FF. The hand embedding issues have already been addressed in the previous chapters. A more realistic presentation of the FF (e.g. fluid-like rendering, particle-hand interactions) would again raise the impact of the visual modality. The VPC computations need to be revised and adapted to the experimental findings by including new minimum muscle stimulation parameters and reorganising the remaining F bandwidth. Another promising opportunity for general sensation improvements would be to incorporate a disambiguation support. In HEMP, sensory integration is the principally exploited source for altering perception. However, sometimes hardly resolvable conditions may occur in terms of signal suitability (e.g. oversimplification of the contributing sensory supply, perceptual stress due to overstraining VPC, insufficient or inappropriate muscle activities for the pseudo-haptic force to be simulated). As a result, the percept could become unstable or ambiguous, the merge incomplete and the final sensation perturbing or illogic to the user. One possibility to overcome this problem could be to involve a processing mechanism specially focusing on disambiguation: sensory combination (see Section 2.1). That is, complementary information about relevant environmental properties is delivered to the user's CNS in order to reinforce perceptual reliance. In other words, providing additional cues to make the observed scenario more reliable would help inhibiting undesired perceptual and cognitive “doubts”. Controllable

physical air streams or a few vibro-elements attached to the hand could serve as simple tactile displays. Also, sounds and / or additional visual stimuli could convey the required feedback. Whatever can enrich the context of the force-emitting virtual phenomenon and its effect on the user might be qualified to feed sensory combination – and therefore to compensate for sensory integration deficits. Brain imaging protocols could identify the neural basis of the studied force illusion.

Usability and practicability of the proposed and the enhanced HEMP solutions have to be rigorously tested before applying them to adequate VE. The exploration of scientific data and the development of sensible 3D graphical interfaces are only two productive perspectives. Further interesting fields for applications are educational entertainment, scenarios in which flows of different intensities or even different matters can be experienced without having them "actually present". Beyond virtual FF, other active and passive HEMP approaches are contrivable. Lifting virtual objects (i.e. weight), deforming virtual surfaces (i.e. compliance), simulating friction, stiffness or rudimentary fluid attributes are just a few examples. The sensation will, of course, never reach the quality of equivalent real haptics or an absolutely authentic sensory occupation. But it can be an interesting alternative, though, having its own unique potentials.

Publications

1. A. Pusch, O. Martin, and S. Coquillart, "HEMP – Hand-Displacement-Based Pseudo-Haptics: A Study of a Force Field Application", In Proceedings of IEEE 3DUI 2008, Reno, USA, 2008.

   => "Honorable Mention Award"

2. A. Pusch, O. Martin, and S. Coquillart, "HEMP – Hand-Displacement-Based Pseudo-Haptics: A Study of a Force Field Application and a Behavioural Analysis", to appear in International Journal of Human-Computer Studies, accepted for publication on 24 Sept. 2008.

   => Invited paper

6 VISUAL-TO-PROPRIOCEPTIVE HAND FEEDBACK CONVERGENCE

Visuo-proprioceptive-conflict-based 3D UI approaches suffer by definition from potentially problematic side effects. They generate and / or enlarge multisensory conflicts between the kinaesthetic and the visual hand representations. In this chapter, a possible offset reduction procedure is developed which accounts for this issue and can easily be applied to any interaction technique relying on visuo-proprioceptive conflicts (VPC).

In Section 6.2, previous work related to offset treatments is discussed. After, practical and theoretical consideration contributing to the hand feedback convergence process (HFC) will be addressed (see Sections 6.3 and 6.4, resp.). The experimental investigation of the approach is initiated in Section 6.5 and will be closed in Section 6.8. In the concluding Section 6.9, also a number of future perspectives are pointed out.

6.1 Introduction

It has been demonstrated that providing a deliberately manipulated real hand feedback can add value to near space interaction and deliver a novel type of pseudo-haptic sensation (see Chapters 4 and 5). Both techniques make use of a spatial visuomotor conflict between the real hand and its virtual visual duplicate. Since there is no offset reduction procedure integrated so far, true applicability of the approaches seems to be unnecessarily restricted. For instance, in the case of the intuitive surface touching paradigm, repeated collisions with scene constraints would automatically cause ever deeper object penetrations. That is, the distance between the hand representations grows (see Fig. 6-1). Concerning the force field (FF) application, the closer the actual offset gets to its allowed maximum, the more of the remaining displacement range is lost. At some point, the intended strength of the FF can no longer be simulated leading to diminishing force illusion effects. Finally, in both cases, an unintended adaptation could occur over time.

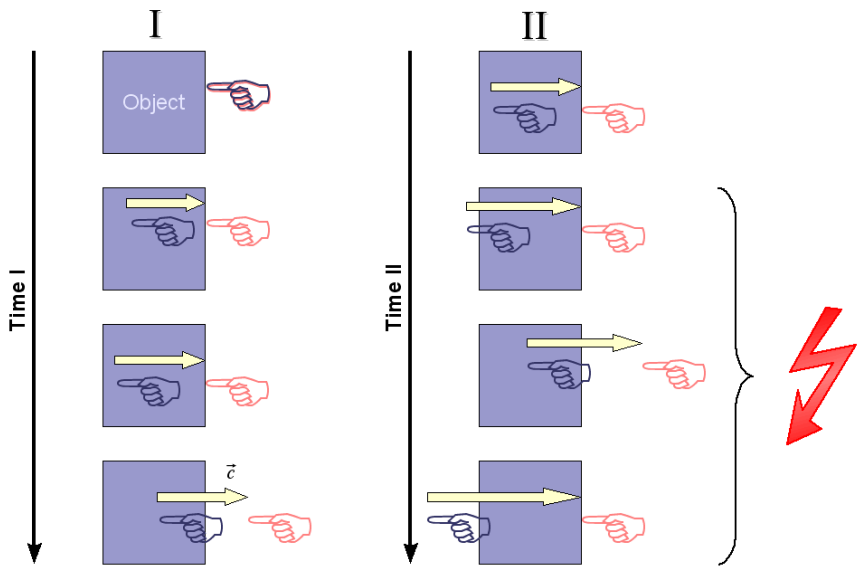


Figure 6-1: Potentially problematic VPC cumulation due to repeated surface contacts or touching events (dark blue: real hand, salmon: visual hand).

Beside these functional problems, perceptual and technical aspects have to be respected as well (e.g. static and dynamic displacement thresholds, video capturing space, advantage of frequently re-adapting to "normal" visuomotor conditions, i.e. to true co-location, and to limit VPC after-effects elicited by perceptual recalibration, see Section 2.1). The VPC framework could prevent too large hand offsets by sustaining maximum values or, regarding video see-through (VST) device constraints, simply replace the video hand feedback by a high quality 3D hand model (see Section 3.4). However, none of these options is satisfying, because they would partially imply radical changes in concrete applications and do not solve the possible adaptation drawbacks.

A generic solution is proposed instead which gradually reduces existing discrepancies based on the HFC model (see Section 6.4). Its overall goal is to minimise VPC as fast as possible without perturbing the user or giving him the feeling of being "controlled by the system". Interventions should ideally not be noticeable and not require interruption of the current task. The HFC computation takes the above-mentioned constraints and several user factors into account (e.g. real hand movement direction and velocity, supposed attentional focus, human field of view). Further, due to its general purpose, the method could become an integral part of the VPC framework.

6.2 Related work

Only a few works have hitherto raised the question of how to manage decoupled sensory information about the user's hand in space. For instance, Burns et al. (2005) emphasised it after comparing “detection thresholds for visual interpenetration (the depth at which they see that two objects have interpenetrated) and sensory discrepancy (the displacement at which they notice mismatched visual and proprioceptive cues)”. The study design and inferred claims may be disputable, since attentive passive observations were compared to actions in a game-like pointing scenario with the “attentional priority” directed to the conflict detection. Apart from this, it was shown that participants were more sensitive to object intersections than to VPC (see also Section 4.2). The authors concluded that “separating the real hand and visual avatar hand to prevent visual interpenetration is beneficial”, but it would “introduce new concerns”. They wanted to reduce the interaction performance decline, avoid “intolerably large” offsets and provide, even in the presence of VPC, “the most perceptually plausible experience to the user”. The maximised hand motion coherence as it can be achieved by the incremental motion method (IM, see Zachmann et al., 2001, and Section 4.3) was considered to be less salient than an avatar sticking to the virtual object like in the rubber band method (RB, see Zachmann et al., 2001, and Section 4.2). The goal of RB is to minimise positional mismatches. IM never reduces any offset.

Neither of these two approaches was found to be optimal and an alternative solution has been proposed: the Credible Avatar Limb Motion technique (CALM, see Burns et al., 2005) which the authors have later referred to as: Management of Avatar Conflict By Employment of Technique Hybrid (MACBETH) [Bu07]. It is situated between the opposing ends of the “discrepancy continuum” delimited by IM and RB (see Fig. 6-2).

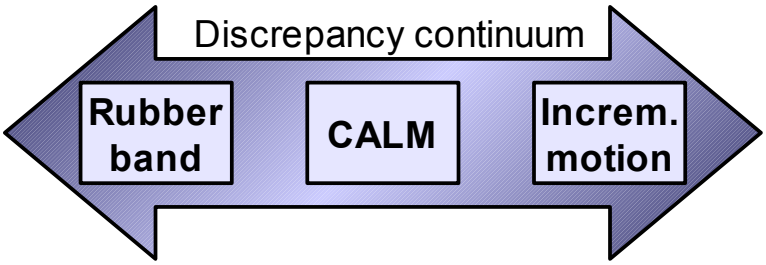


Figure 6-2: CALM / MACBETH situated between RB and IM (see Burns et al., 2005).

Although the necessity of such an intermediate VPC treatment method was indicated, Burns et al. (2005, 2007) did not or could not yet specify details on the computational basis, implementation aspects, the level of applicability or potential limitations. Thus, the design principles remain unclear.

6.3 System integration fundamentals

The HFC procedure is generally independent of the underlying system infrastructure, but may operate as an encapsulated VPC management module. In order to calculate the convergence velocity v_{conv}, access to the following information is mandatory:

- Current head and hand tracking data at 6 degrees of freedom.
- Given offset vector $\vec{o}$.

The need for a VPC reduction implies that a sensory decoupling took already place in one way or another. Suppose the VPC framework has been installed as a control interface to the hand representation, then the listed parameters are immediately available for computations.

To perform the real-to-visual hand feedback convergence, first, an activation flag has to be set. Thenceforth, v_{conv} will be used to update $\vec{o}$ before it is assigned to the hand texture carrier objects or any alternative virtual hand representation. Concrete applications can affect the HFC impact by passing a miscellaneous convergence modifier (see Section 6.4). The configuration of all internal factors, including their ranges and rules with respect to the user state is done in a plain text file. The embedding into the VPC framework requires a redirected data flow and therefore an adaptation of the module interconnexions (see Fig. 6-3).

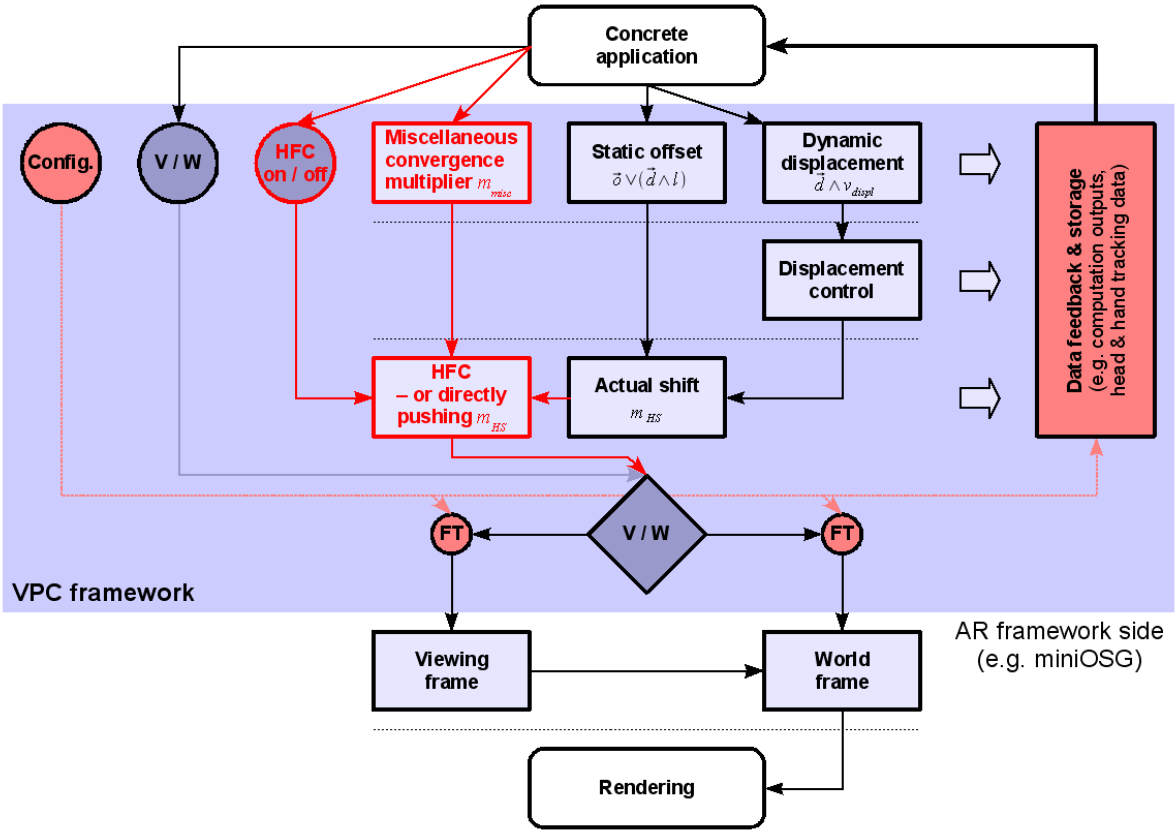


Figure 6-3: HFC embedded into the VPC framework.

6.4 Hand feedback convergence model

The VPC reduction or HFC could theoretically be activated as soon as $\vec{o}$ is not increased anymore. To check for this case, the offset vector will be observed. After an adjustable delay and if the activation flag has been set, HFC becomes effective until $|\vec{o}| \leq 0.05\,cm$. $\vec{o}$ will be cut to zero length below this configurable threshold.

The main equation for the computation of the convergence velocity is:

$$v_{conv} = v_{RH} \cdot m_{H2V} \cdot m_{H2O} \cdot m_{misc} \qquad \text{(Eq. 6-1)}$$

With v_{RH} being the real hand velocity in cm/s, m_{H2V} the visual-hand-to-viewing-centre deviation multiplier, m_{H2O} the hand-movement-direction-to-offset deviation multiplier and m_{misc} the application-controlled miscellaneous convergence modifier. It should be recalled that, in addition to the convergence performance, user comfort constraints play an important

role for the determination of these factors. Hereafter, all relevant points will be described.

The tracked hand position is buffered roughly at each 100 ms. To estimate the real hand velocity, the temporal and positional changes between the current and the buffered data are used:

$$v_{RH} = \frac{HP_{curr} - HP_{buff}}{t_{curr} - t_{buff}} \qquad \text{(Eq. 6-2)}$$

With HP being the hand positions and t the associated timestamps. Hand rotations are not taken into account.

m_{H2V} represents the attentional focus in a way. Since no eye tracking is employed, there can only be a moderate approximation. In previous studies, it has been revealed that online manipulations of the visuomotor organisation are realised a lot earlier, if people direct their attention to the respective event (see Burns et al., 2005, but also Section 5.4.2.1). That is, the further the focus is away from the centre of the viewing space, the faster could be the convergence. First, the head-to-visual-hand vector $\vec{v}_{H2H}$ is determined based on the given tracking data and $\vec{o}$. Next, by considering $\vec{v}_{H2H}$ and the head orientation vector $\vec{v}_{gaze}$, the enclosed visual-hand-to-viewing deviation angle δ_{H2V} can be obtained. m_{H2V} will finally be computed as a function of δ_{H2V} (see Eq. 6-2). Regarding the human field of view which has an oval shape covering about 200 degrees horizontally and 135 degrees vertically with a stereo overlap of about 60 degrees [We91] (see Fig. 6-4), the function parameters could be set as follows:

- At the minimum visible deviation (i.e. at 0 degrees: 25%): $m_{H2V} = 0.25$.
- At the maximum visible deviation (i.e. at 100 degrees: 100%): $m_{H2V} = 1.0$.

These boundaries are best practice measures and yield for a linearly interpolated m_{H2V} :

$$m_{H2V} = \frac{1.0 - 0.25}{100 - 0} \delta_{H2V} + 0.25$$
$$m_{H2V} = 0.0075\, \delta_{H2V} + 0.25 \qquad \text{(Eq. 6-3)}$$

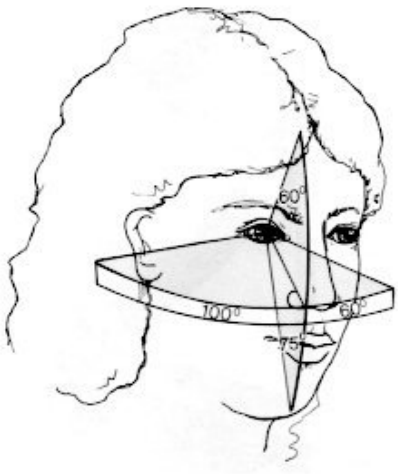


Figure 6-4: Human visual field
(see Werner, 1991).

m_{H2O} is contingent on the divergence of the hand movement direction from $\vec{o}$. The main reason for the existence of this multiplier is an interesting observation made in one of the preliminary VPC dynamics experiments. Subjects had to indicate, after horizontally pointing from one location to another, whether they found their hand movement naturally displayed or somehow manipulated. In the latter case, subjects were requested to specify their sensation by deciding between either a faster or slower perceived hand motion. It seems that they were more aware of movement supports. However, future experiments should pick up this question to narrow the real thresholds and dependencies down. The basic assumption for now is that v_{conv} should be smaller, the larger the angle between the hand movement direction vector $\vec{v}_{HM}$ and $\vec{o}$ is. Rotational VPC are not included. In a first step, $\vec{v}_{HM}$ is determined at the begin of each new 100 ms buffering interval (see above): $\vec{v}_{HM} = HP_{curr} - HP_{buff}$. Then, the hand-movement-direction-to-offset deviation angle δ_{H2O} is computed as just mentioned. The multiplier interpolation function uses these coefficients:

- Co-directed hand movement (i.e. at 0 degrees: 75%): $m_{H2O} = 0.75$.
- Counter-directed hand movement (i.e. at 180 degrees: 25%): $m_{H2O} = 0.25$.

Resulting in:

$$m_{H2O}=\frac{0.25-0.75}{180-0}\delta_{H2O}+0.75$$
$$m_{H2O}=-0.002\bar{7}\,\delta_{H2O}+0.75 \qquad \text{(Eq. 6-4)}$$

Any additional environmental effects on the convergence can be controlled by the concrete application (i.e. by passing m_{misc}). The internal structure of the HFC model is summarised in Figure 6-5. Upper and lower bounds for m_{H2V} and m_{H2O} are defined in a configuration file.

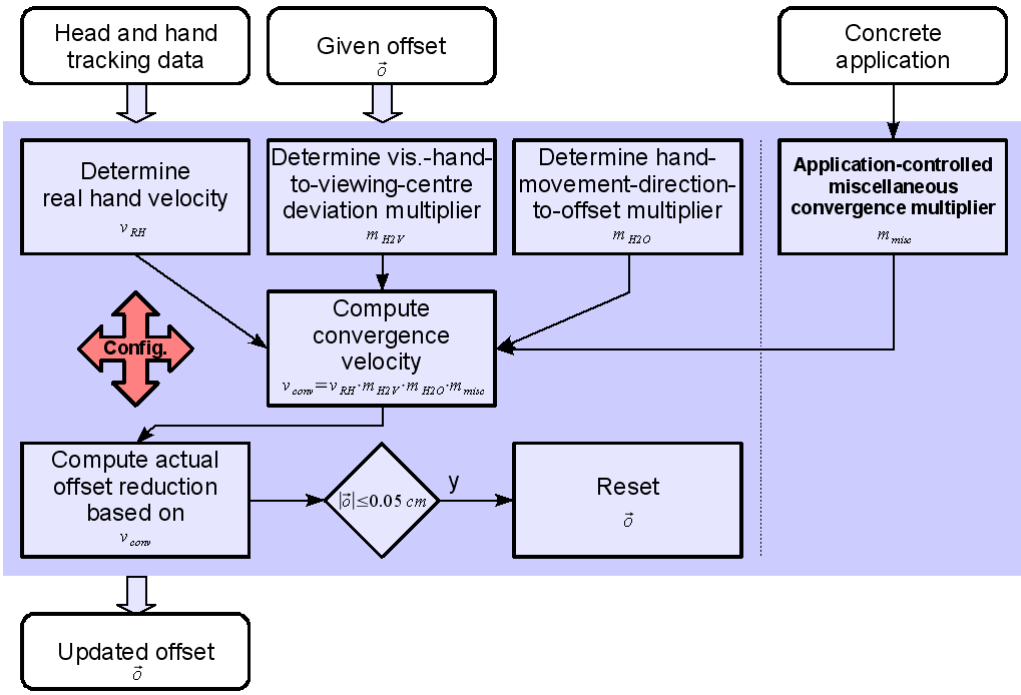


Figure 6-5: HFC model (i.e. hand feedback convergence model).

Based on the proposed multiplier ranges, the maximum convergence velocity for actions within the human visual field can reach $v_{conv}=0.75\,v_{RH}\cdot m_{misc}$. When taking the given VST HMD properties into account (e.g. average maximum eccentricity: 20 degrees, see Section 5.4.2.2), then the achievable maximum velocity would be $v_{conv}=0.3\,v_{RH}\cdot m_{misc}$. What these examples show is that the HFC outcome usually fulfils $v_{conv}<v_{RH}$. A change in this relation may only be enforced through m_{misc} . Thus, implausible visual drifts, an apparent visual

immobility or, even worse, a visual motion opposite to the performed hand movements can, under normal conditions, be avoided. All function parameters represent best practice values for an intendedly unnoticeable HFC. But future research should experimentally verify them in detail. More general perceptual and device constraints are handled by the VPC framework.

6.5 Hypotheses

The two hypotheses arising from the principal goal and the design of HFC concern the user's sensations as well as his motor performance when being confronted with a gradually reduced VPC in a standard reaching task. There is solely RB to which the novel approach can be compared. This fact by itself proves already a major advantage of HFC. Since, if no virtual surface or object constraints are existent and exploitable (i.e. blurred, semitransparent or no shapes at all as in the pseudo-haptic FF, see Chapter 5), then RB would not make sense. For other, more classical cases, a comparison will though reveal illuminating insights on benefits, limitations and improvement potentials.

Hypothesis 1 is mainly motivated by observations made in previous experiments (see Burns et al., 2005) and a number of informal studies conducted within the frame of this work. It appears that visually incoherent hand motions are easier detected than coherent ones which maintain a static VPC at hand level. They may further convey a rather artificial feeling of control. Hence, subjects should notice RB earlier or more frequent than HFC, specially if they focus directly on their actions, and they are expected to experience RB as less comfortable. The influence of HFC on the spatial limb movements should not be perceived.

Hypothesis 2 addresses the impact of the offset reduction methods on motor performance. Regarding surface-induced VPC, RB inherently benefits from co-location as soon as the real hand moves out of the space occupied by a virtual object. Precise pointing movements should thus be faster compared to movements which first, require an adaptation to dynamically changing sensory discrepancies (see also Mine et al., 1997, and Paljic et al., 2002) and second, are perhaps subject to certain offsets at the terminal pointing location. Contrariwise, in HFC, an incomplete offset reduction could be compensated for by the shorter real hand transport. Harmonised target characteristics (see also Section 6.6.2) providing similar indices of difficulty ID [Fi54] for both methods (see Fig. 6-6) should at least negate this advantage.

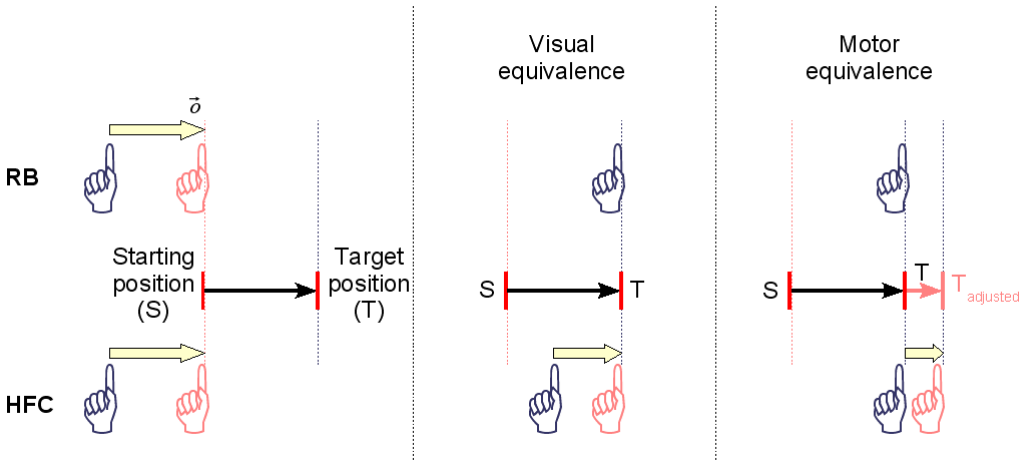


Figure 6-6: Visual and motor equivalence (dark blue: real hand, salmon: visual hand).

The interactive part of the experimental procedure consists of a cube pushing subtask. During this action, the visuomotor conflict later to be reduced is created by modifying the so-called control-to-display ratio (C/D ratio, see also Crison et al., 2004). Participants see the visual substitute of their hand moving slower than the real hand. This condition is principally qualified to cause a pseudo-haptic sensation (e.g. heaviness or sliding friction, see Section 5.9), since a certain muscle activity is combined with a related, but divergent visual feedback. The questionnaire will devote a supplementary question to this phenomenon.

6.6 Experiment

The experimental investigation of the HFC method with respect to the central hypotheses (see Section 6.5) was separated into an interaction task (see Section 6.6.3.1) and a sequential questionnaire for subjective responses (see Section 6.6.3.2). The interaction task was carried out first.

6.6.1 Subjects

Fifteen adult volunteers (i.e. mean age: 28.7, SD: 5, 7 female, 8 male) took part in the study. Fourteen were right-handed and one subject ambidextrous with a strong tendency towards right-handedness. Except for two subjects with a slightly degraded stereo vision capability,

nobody suffered from vision problems (i.e. normal or corrected to normal vision). To each level of prior AR / VR knowledge (i.e. none, some, [very] good), there could 5 individuals be assigned. None of them was informed about the purpose of the experiment.

6.6.2 Factorial design

For the interaction task, a two factors 3 x 4 design was chosen. Factor one describes the employed hand offset reduction procedures (see further down) and factor two the accuracy pointing targets T . The spheric target objects were defined upon their 1D distance D from the initial reaching position (i.e. near: 17 cm, far: 32 cm) and their size or width W (i.e. small: 1 cm, large: 2 cm). This yielded a subsidiary 2 x 2 matrix (see Table 6-1).

Table 6-1: Target conditions.

T_i	near	far
small	T_1	T_2
large	T_3	T_4

The selected values for D and W should help rendering target characteristics as they were required to test *Hypothesis 2*. For each T , ID was calculated using the respective term of the Shannon-Hartley theorem [MK92] which represents an empirically more stable refinement of Fitts' original law:

$$ID = \log_2 \left(\frac{D}{W} + 1 \right) \qquad \text{(Eq. 6-5)}$$

Table 6-2: The targets' IDs.

ID_T	near	far
small	**4.17**	**5.04**
large	**3.25**	**4.09**

Table 6-2 shows an interesting relation between T_1 and T_4. The far spheres lay a little bit outside the field of view of the utilised VST HMD (see Section 5.4.2.2). Will there be an extra load for searching the target under these circumstances? The regression analysis will focus on this question in detail (see Section 6.7.3).

The main offset reduction methods covered by the first experimental factor are RB and HFC. As a pilot study revealed, the currently applied HFC parameters in conjunction with the envisaged initial VPC of 10 cm (see Section 6.6.3.1) would not have allowed for a complete offset elimination. Neither tuning the parameters appeared to be a solution, since awareness and perturbation probabilities would have been raised, nor smaller and therefore potentially marginal VPC. Placing the targets at a sufficiently far distance from the starting position so that a full reduction can be guaranteed would have led to two undesirable effects. First, the near target conditions become redundant and second, the closest possible D would be rather inappropriate regarding the VST HMD pointing ergonomics. Suppose a maximised attention (i.e. $m_{H2V}=0.25$) and a minimised δ_{H20} (i.e. $m_{H20}=0.75$), then $v_{conv}=0.1875\, v_{RH} \cdot m_{misc}$. Removing the more illustrative time component and simplifying the equation by excluding m_{misc}, one obtains for the convergence distance: $d_{conv}=0.1875\, d_{RH}$. Letting $d_{conv}=10\, cm$ (i.e. imposed VPC, see above) finally yields for the minimum D: $d_{RH}=53.\bar{3}\, cm$. These calculations imply ideal, perfectly straight hand movements. However, an equivalent physical or motor distance can alternatively be achieved by adapting the given D_{near} and D_{far} to the remaining VPC for the associated unmodified visual D. Estimated offset residuals are 6.8 cm for D_{near} and 4 cm for D_{far}. Informal tests showed actual residuals of about 7 and 3 cm suggesting an increased hand trajectory noise and / or bigger δ_{H2V} variations with a longer target distance. The concluding hand offset reduction methods are thus:

3. RB.
4. HFC-V (i.e. HFC with target distances equivalent to RB at visual level).
5. HFC-M (i.e. HFC with target distances equivalent to RB at motor level: $D_{near}+7\, cm$, $D_{far}+3\, cm$).

The experimental conditions according to the 3 x 4 factorial design were presented in a balanced random order (i.e. 12 x 12 random latin square). With 12 repetitions per condition, subjects had to accomplish 144 trials.

6.6.3 Procedure

General information concerning age, gender, handedness, prior AR / VR experiences etc. was requested before the begin of the experiment. The main part comprised the interaction task and the questionnaire (see next sections). One trial of the former had a duration of 12 seconds. In total, subjects passed two blocks of 14:24 minutes each with a short break in between.

Due to immanent system lags and the need for the smallest possible visual feedback delays, the video hand paradigm could not be employed. As a trade-off and to provide additional high precision cues (see also Section 4.7), the VPC framework was configured as to present an arrow object instead. The inclination of the arrow coincided with the index finger and and top match the fingertip (see Fig. 6-7).

Figure 6-7: Pointing gesture, with virtual arrow overlay.

The visual field thresholds used for the computation of m_{H2V} and m_{H2O} were set to 180 degrees horizontally and 60 degrees vertically. An application of the more realistic values proposed in Section 6.4 followed later.

6.6.3.1 Interaction task

The task setting was chosen in order to simulate a simple interaction scenario as it could be met in near space object manipulation or the control of a 3D graphical interface. Subjects had first to read an introductory note explaining that whenever they would observe something "'strange' or unusual beside changes in target / sphere sizes and distances", they should immediately report it. Keywords would be sufficient. The preparation was rounded off by demonstrating the rest position and the tracking device grasp gesture (see Fig. 4-5).

The interaction task itself consisted of two subtasks (see Fig. 6-8 for the whole sequence): cube pushing (i.e. hidden VPC creation) and sphere popping (i.e. Fitts'-Law-conform rapid pointing). The base scene contained one upright and parallel to the viewing direction oriented red square of 6 cm edge length. It was floating 45 cm in front, at shoulder height and 13 cm left of the subject's sagittal plane.

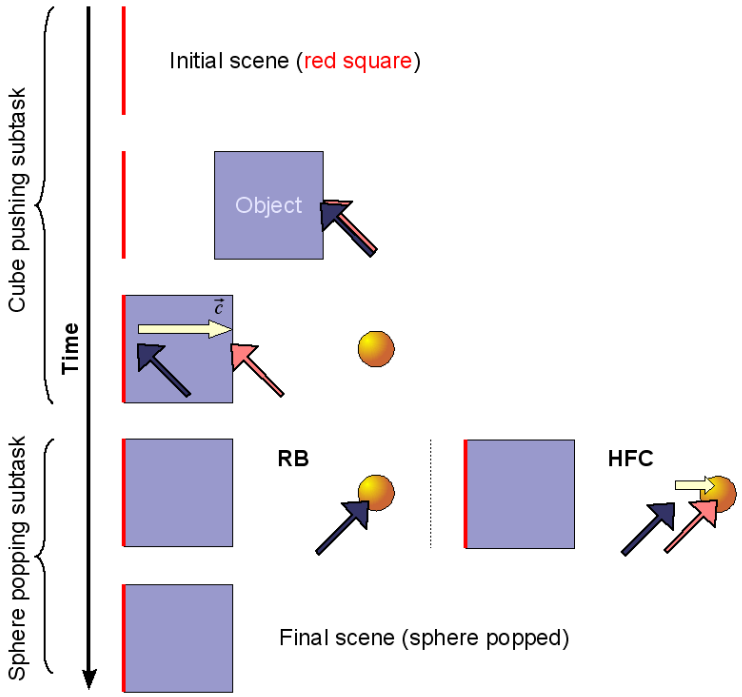


Figure 6-8: Interaction task sequence (dark blue: real hand / arrow position, salmon: visual hand / arrow position).

A trial was initiated by an acoustic notification (i.e. first beep). One second before the beep, a cube of the same edge length as the square appeared at 5 cm distance right of the latter. Participants had to touch the right side of the cube with the top of the arrow and push it leftwards. Once the cube had touched the square, a tapping sound was played back to confirm this essential event. During pushing, a VPC of 10 cm was generated. The cube moved visually for about 5 cm. With the contact, the second subtask began and the final pointing target (i.e. yellow sphere) emerged at the same depth as the cube. Subjects were instructed to perform a fast, straight and direct movement towards this target and put the top of the arrow inside the sphere until it would pop. The intersection had to last for about one second to make sure that the hand did not simply pass through the target. Then, the hand or rather the arrow had to be returned to the rest position. Another beep, 8 seconds after the first, signalled the end of the active trial period and all objects added to the scene were removed (i.e. the cube and the target sphere). In cases of unfinished trials, subjects had to come back to the rest position and were asked to do a little bit faster next time. The procedure was practised until participants felt comfortable handling both subtasks and their control was objectively acceptable (e.g. good timing, smooth movements, stable trial completion rates).

6.6.3.2 Questionnaire

The questionnaire (see Appendix 3) was handed out just after the interaction task was done. Its design focused on indicators for testing the statements of *Hypothesis 1*. Moreover, the supposed pseudo-haptic side effect was allusively addressed.

Each of the successive questionnaire sheets opened with a phase-wise “recall of the task”:

(a) “Starting from the rest position, touching the right side of the cube”.

(b) “Pushing the cube until it got in contact with the red square”.

(c) “Once the red square was reached, touching the sphere”.

(d) “After the sphere was touched, returning to the rest position”.

There were 6 main questions. The three first emphasised global phase-related sensations, “'strange' or unusual” effects during the experiment and the perceived differences other than varying target distances and sizes. If anything odd was remembered, it had to be explained and the temporal occurrence had to be specified. The questions 4 and 5 concentrated on phase (c) in which the offset reduction methods eventually took effect. Detected differences as well as “possibly perturbing or uncomfortable situations” had to be reported and to be described. The last main point refocused again on the experiment as a whole by asking for any other condition or moment of discomfort. General remarks could be given in the end.

6.6.4 Data acquisition and analysis

The analysis resource was composed of the qualitative questionnaire responses (see Section 6.7.1) and quantitative task completion times for the sphere popping subtask. Additionally, the remaining offsets for the HFC conditions and head and hand tracking information were recorded. Standard statistics (i.e. first descriptive, then comparative using repeated measures ANOVA, Pearson's product moment correlation and, if suitable, Bonferroni-corrected LSD) have been performed in a cascading manner to study the following aspects:

1. Global method comparison (i.e. overall effects of the offset reduction procedures and the targets on the task performance, see Section 6.7.2).
2. Task completion time regression analysis (i.e. test for the Fitts' law conformity of the results, see Section 6.7.3).
3. Target-characteristics-based method assessment (i.e. method effects dependent on the target submatrix elements, see Table 6-1 and Section 6.7.4).

The task completion times were subtracted by the required target intersection or pointing stabilisation duration.

6.7 Results

Questionnaire and performance statistics results will be presented in the next three sections.

Two subjects had to be excluded from the analysis of the quantitative data because of too high error rates (i.e. about 18 and 23%). In most cases, trials were not completed in time, although the practice block was successfully passed. Both subjects showed ongoing problems in estimating the targets' depths. This might have been due to a lack of stereo vision either caused by an imprecisely adjusted VST HMD or limited personal visual abilities. The data loss among the remaining 13 participants was at 2.72% (i.e. approx. 4 trial per individual).

In Section 6.6.1 it was stated that subject fell into three categories. An informal test did neither reveal a group effect on the task performance ($F(2, 10) = 0.46$; $p > 0.6$) nor an interaction between the subject category and the offset reduction procedure ($F(4, 20) = 0.55$; $p > 0.7$). The population was hence not divided.

Another pre-analysis aimed at the verification of the preset motor equivalence parameters for HFC-M. To balance ID at motor level, the near targets were shifted for 7 cm and the far targets for 3 cm away from their normal positions. The experimentally determined average VPC residuals for HFC-V were 7.05 and 2.91 cm and confirmed this way the appropriateness of the corresponding corrections.

6.7.1 Qualitative questionnaire results

Participants were asked to report their sensations during the interaction task phases (i.e. approach, cube pushing, sphere popping, return), any “'strange' or unusual” observation and the differences they perceived between trials except for changes in target sizes and distances.

An incidental “stickiness effect” was mentioned by more than 73% of the subjects (i.e. 11 out of 15). Without the two subjects excluded from the quantitative analysis, it would have been almost 85%. But only 6 out of 11 (i.e. approx. 54%) considered this effect to be a “difference between trials”. In four cases (i.e. approx. 36%), the “stickiness” was experienced as being perturbing or hindering in performing the task. One subject felt that “the pointer was constrained in about 20% of the trials”. Another one thought that “sometimes there was a lag when the ball was close” and a last subject commented that the “arrow tends to stick on the

cube when a big and near sphere appears”. On the other hand, 2 out of 15 participants (i.e. approx. 13%, without the two excluded subjects: slightly more than 15%) reported a “feeling that the top of the arrow was shifted compared to my finger” or an “impression of moving the hand further than what the visual display showed”. In the former case, the subject was not sure about the direction of the “shift”. After specifically asking for the potential displacement dimension, the subject guessed it would have been a shift in depth. The second candidate experienced “different 'real distances' compared to 'visual distances'”. None of them found the apparent VPC uncomfortable or perturbing. Only one subject remarked “lags”, later identified as the visual RB effect, in the course of the experiment.

Mainly in the sensation description section and once declared as an observed difference, at least 6 out of 15 subjects (i.e. 40%, revised: approx. 31%) perceived a certain “heaviness” of the cube while pushing it. The impressions were specified as “seems like you need to apply some force to move the cube”, “a slight heaviness feeling when pushing the cube”, “the cube was resisting” or “varying 'force' to apply in order to make the cube moving”.

Further interesting statements concerned various facets of the experiment, for instance, “not easy to estimate depth, since the own hand cannot be seen”, “impressive precision achievable handling the cursor” and “natural handling, also when pushing the cube”. One person who found it “difficult to stabilise the hand on the far sphere” and asked whether the target sizes actually differed: “I did not observe different target sizes!”

General discomfort was mostly related to the VST HMD. Remarks included: “device was hurting on nose and forehead”, “blurred view in the end of each block”, “slight dizziness”, “fairly poor ergonomics”, “eye pressure, accommodation problems and perturbing luminosity changes”, “heavy HMD”, “visual fatiguing”, “HMD too bright causing a headache” and “from time to time, slight disequilibrium”. However, critical or dangerous situations occurred at no time.

6.7.2 Global method comparison

With respect to the employed 3 x 4 factorial design (see Section 6.6.2), a global method effect was found ($F(2, 24) = 27.28$; $p < 0.0001$, see Fig. 6-9) as well as a general target influence on the completion time of the sphere popping subtask ($F(3, 36) = 158.58$; $p < 0.0001$). The method-target interaction reached significance ($F(6, 72) = 2.89$; $p < 0.015$). A Bonferroni-corrected LSD revealed timing differences between RB and HFC-V ($p < 0.0001$), HFC-V and HFC-M ($p < 0.0001$), but not between RB and HFC-M ($p > 0.3$). HFC-V performed always better than either of the other two offset reduction methods.

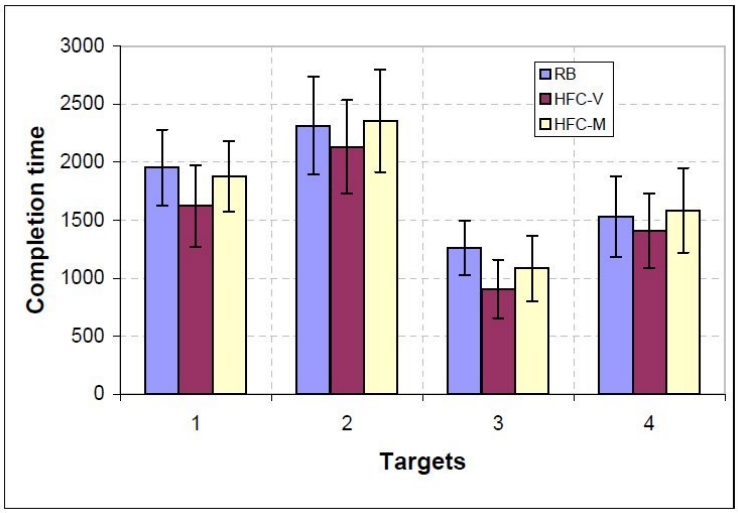


Figure 6-9: Global method and target effects
on the task completion time (means, SDs).

Regarding the target factor separately, results differed from another at $p < 0.0001$ (for a detailed assessment, see Section 6.7.4). When considering the 2 x 2 target property submatrix, both the target distance D and the target size W showed a significant impact on the task performance ($F(1, 12) = 128.93$; $p < 0.0001$ and $F(1, 12) = 221.36$; $p < 0.0001$, resp.). An interaction could only be found between the reduction procedure and D ($F(2, 24) = 6.45$; $p < 0.006$). Within each subfactor, a pairwise comparison yielded significant differences among levels ($p < 0.0001$).

6.7.3 Task completion time regression analysis

In this section, the suitability of the chosen motor performance investigation scenario and the predictability of the results of each VPC minimisation procedure will be verified using the following regression function (i.e. refined Fitts' law, see MacKenzie, 1992):

$$MT = a + b\log_2\left(\frac{D}{W} + 1\right) \qquad \text{(Eq. 6-6)}$$

MT describes the movement time (i.e. completion time of the second subtask), and a and b the regression coefficients. A global linear fitting was found ($r^2 = 0.933$; $t = 6.54$; $p < 0.023$, see Fig. 6-10, top). But for RB alone, the correlation did not prove significant ($r^2 = 0.83$; $t = 3.96$; $p > 0.05$) whereas it did for HFC-V ($r^2 = 0.98$; $t = 10.81$; $p < 0.009$) and HFC-M ($r^2 = 0.95$; $t = 7.63$; $p < 0.017$). RB tends thus to violate the fundamentals of Fitts' law. An extra load for searching targets outside the HMD's field of view appears not to exist.

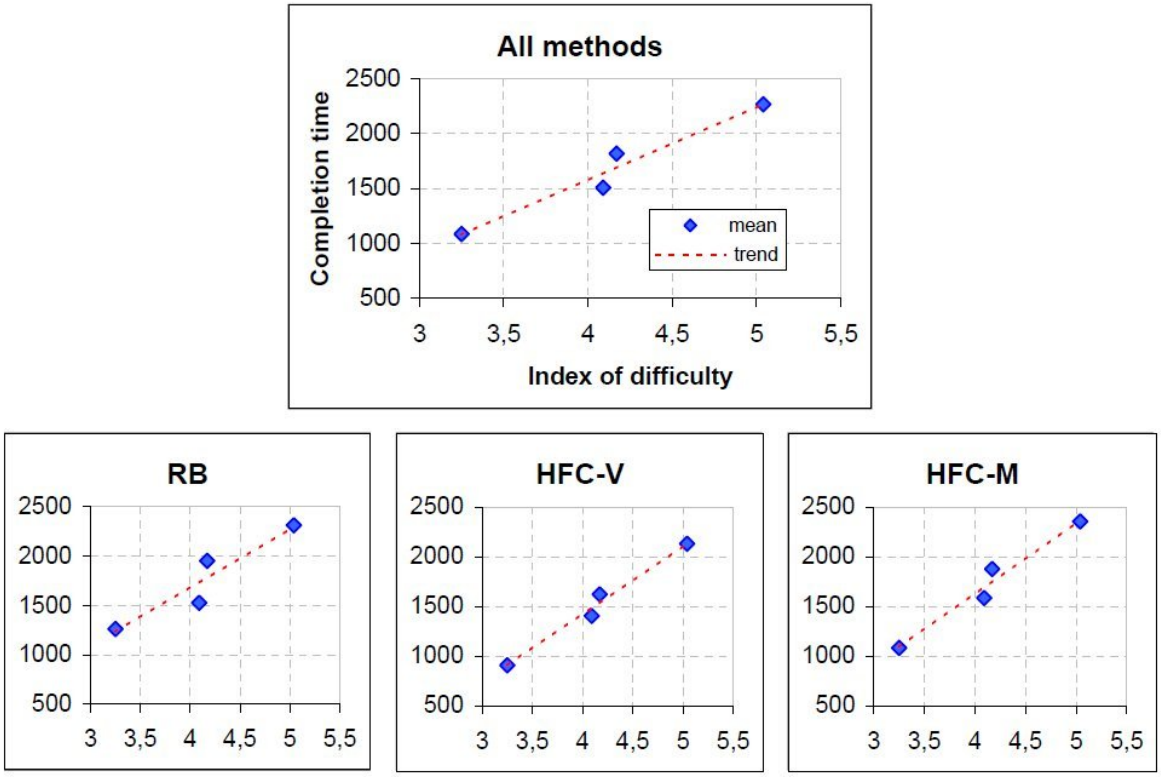


Figure 6-10: Target-ID-to-mean-movement-time correlation at global level (top) and for each offset reduction method.

6.7.4 Target-characteristics-based method assessment

This part of the analysis emphasises the influence of the target properties (i.e. distances and sizes) on the results of the compared offset reduction methods. Since a method-distance interaction was found, first, the effect of D was tested. For the near and far targets, the differences between methods are listed in Table 6-4 (i.e. mean completion time difference: $\text{mean}_{\text{method1}} - \text{mean}_{\text{method2}}$). It is important to note that in the "near RB versus HFC-M" case, an advantage for the motor-equivalent HFC version was observed (see Fig. 6-11), even with a successfully adjusted ID (see above).

Table 6-3: Pairwise comparison of the offset reduction methods I: near and far.

	near			far		
Methods	mean diff.	critical diff.	p-value	mean diff.	critical diff.	p-value
RB vs. HFC-V	342.33	118.37	**< 0.0001**	151.62	121.48	**< 0.0001**
RB vs. HFC-M	126.37	118.37	**< 0.012**	-50.31	121.48	**> 0.2**
HFC-V vs. HFC-M	-215.97	118.37	**< 0.0001**	-201.92	121.48	**< 0.0001**

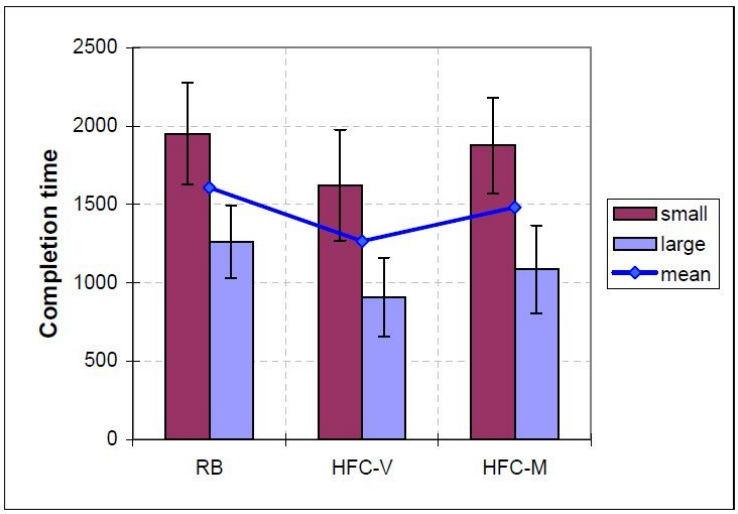


Figure 6-11: Influence of the offset reduction method on the task completion time for near targets (means, SDs).

The general benefit of HFC-V continued also for the second target variable W (see Table 6-3). Apparently, in the comparison of RB versus HFC-M, the size of the sphere did not matter statistically.

Table 6-4: Pairwise comparison of the offset reduction methods II: small and large.

	small			large		
Methods	mean diff.	critical diff.	p-value	mean diff.	critical diff.	p-value
RB vs. HFC-V	256.2	125.61	**< 0.0001**	237.75	111.1	**< 0.0001**
RB vs. HFC-M	15.82	125.61	**> 0.7**	60.24	111.1	**> 0.1**
HFC-V vs. HFC-M	-240.39	125.61	**< 0.0001**	-177.51	111.1	**< 0.0005**

A combined analysis of D and W helped encircling specific target-related performance advantages and disadvantages for each method. Table 6-5 summarises all effects according to the target characteristics defined in Section 6.6.2.

Table 6-5: Pairwise comparison of the offset reduction methods III: crossed target properties.

		near			far		
	Methods	mean diff.	critical diff.	p-value	mean diff.	critical diff.	p-value
small	RB vs. HFC-V	330.27	154.66	**< 0.0001**	182.14	178.85	**< 0.016**
	RB vs. HFC-M	74.55	154.66	**> 0.2**	-42.92	178.85	**> 0.5**
	HFC-V vs. HFC-M	-255.71	154.66	**< 0.0004**	-255.06	178.85	**< 0.004**
large	RB vs. HFC-V	354.4	164.59	**< 0.0001**	121.09	110.78	**< 0.01**
	RB vs. HFC-M	178.18	164.59	**< 0.011**	-57.69	110.78	**> 0.1**
	HFC-V vs. HFC-M	-176.23	164.59	**< 0.012**	-178.79	110.78	**< 0.0005**

The near large target T_3 represents therefore the condition in which HFC-M outperformed RB (see Fig. 6-12), despite the forced ID harmonisation and an average VPC residual of still 5.29 cm. Due to the results shown in Table 6-4, one could have thought to find a similar outcome also for the near small target T_1. But the gain is far from being significant.

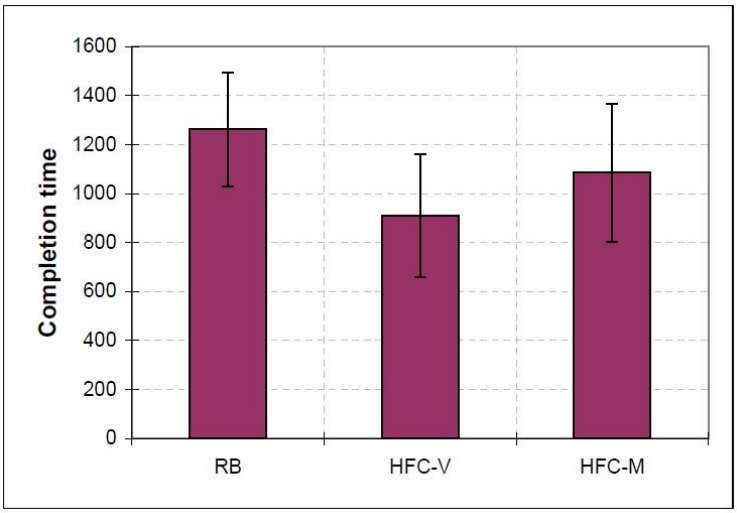


Figure 6-12: Influence of the offset reduction method
on the task completion time for near large targets (means, SDs).

6.8 Discussion of the experiment

The study focused on the novel HFC approach from a conscious perception and a motor performance point of view. Subjects had to accomplish an interaction task and to fill in a sequential questionnaire.

In *Hypothesis 1*, it was stated that the hand motion incoherence caused by RB would be easier detectable than the spatial sensory discrepancy maintained, albeit reduced, by HFC. Two thirds of the interaction task trials were run under HFC (i.e. HFC-V and HFC-M). But about 85% of the subjects noticed some kind of “stickiness” effect compared to only 15% reporting on a, relative to their real hand, spatially “shifted” arrow. Thus, the majority clearly perceived the RB motion feedback artifact. Such strict constraints of intended actions are probably immediately interfering with the incoming proprioceptive signals of a moving limb and the anticipated visual reflexion of the triggered motor activity (i.e. the internal model of forward dynamics, see Wise et al., 2002). For 54% of the participants who have remarked

“stickiness”, this was a potential experimental aspect. They may have noticed it repeatedly, although only three of them have specified it this way (i.e. approx. 27%). Once the impression was linked to near targets and another time to the “big and near sphere”. Obviously, the alertness was raised, if attention rested at proximity. However, 36% described the effect as perturbing or otherwise inconvenient. To conclude, RB conveyed the expected more artificial feeling of control and was frequently detected. The manipulations of HFC were discovered at a considerably lower rate. It is possible to criticise that RB performed a full offset reduction whereas HFC did not and that the convergence parameters for HFC were maybe chosen too conservative. Given the fact that a few participants nevertheless perceived the positionally discrepant visual representation of the hand, it might rather be the case that the resulting convergence velocity was still too high and the multipliers not conservative enough. By and large, even though qualitative advantages seem to exist, adjustments of the underlying HFC model and its parameters are necessary to fulfil all aspects of the initial hypothesis.

It has been assumed in *Hypothesis 2* that RB would profit from co-location during pointing. In fact, there was no condition found in which this constellation occurred, neither at global level nor when focusing on D, W or any T in particular. The only observation made is that with an increased target distance, the mean difference between RB and the concurrent method HFC-M was reduced. In two cases, this has led to an annulment of previous effects (see Tables 6-4 and 6-5). The method-target and the method-target-distance interactions could be explained by means of these observations. But why was RB outperformed, specially in the presence of motor equivalence and a dynamically changing, not completely removed VPC? To recall the situation, the near target appeared 7 cm right of the cube after it had been pushed to its final position. This relatively short distance was likely to trigger an attentive supervision of the whole sphere popping subtask. Since RB inhibits motions perpendicular to an active virtual surface constraint, about 59% of the reaching movement were not displayed (i.e. cube penetration and initial VPC of 10 cm). It can probably be assumed that subjects noticed the “stickiness” mainly here. The visual immobility may have rendered the perceived action unrealistic and compelled the motor system to invest an unexpected and alerting effort (see above). The lacking applicability of Fitts' law seems to confirm this artificial character of RB. A performance decline compared to the potentially more natural HFC-M was only found, when the described unfavourable influences actually became relevant. That is, whenever manipulations in near space required smaller attentional shifts, as it is usually the case, there

will be an increased risk for the user to face perturbing events. In techniques which rely on a stable perceptual involvement (e.g. in pseudo-haptics), suchlike could heavily reduce, if not break the desired effect. Finally, the surprisingly high performance advantage of HFC-V shows that, under normal conditions, RB would drop far behind the proposed method – in most respects. This suggests first, that the induced VPC did not exceed the boundaries of an efficient co-located interaction. Second, the HFC model and its convergence dynamics may represent a good starting point for future refinements. Third and consequently, the initial hypothesis has to be corrected by putting additional emphases on motion coherence and attention and by including an interaction relocation distance factor for a more accurate method eligibility prediction. However, there is still one general benefit of the RB procedure: it provides the fastest hand offset reduction practically achievable. One could try to push HFC closer to this performance by employing nonlinear multiplier functions (see next section).

As a side effect of the modified C/D ratio during cube pushing, the elicitation of a pseudo-haptic impression has been prognosticated. Indeed, about 31% of the participants reported a “heaviness” feeling. But, unlike in the FF experiment, they had to conduct a task not primarily related to a force percept and the question design did not imply a similar response scheme. One reason for the limited occurrence and effect strength could have been that an abstract interaction tool was used in place of the real hand. So, an exciting question for future studies would be to determine how far the degree of limb realism can alter the impact of hand-displacement-based pseudo-haptics (HEMP).

The properties of the VST HMD were once more criticised by the majority of the subjects.

6.9 Conclusion and future work

The goal of this chapter was to introduce a generic method which allows for the convergence of decoupled visual and proprioceptive hand representations. HFC should work unnoticeably, even in the absence of supporting virtual surfaces. It tries to continuously reduce all existing VPC as soon and as fast as possible while preserving interaction naturalness. To this end, the visual-hand-to-viewing-centre and hand-movement-direction-to-offset deviation multipliers are computed. Each has its own configurable scaling range. Different perceptual aspects and an estimated user state are contributing to the final convergence computation. Due to its general purpose, the proposed HFC procedure has been integrated as a new module into the VPC framework. Concrete applications can thus automatically make use of it. Further, the resulting effect can be controlled at runtime (i.e. activation, amplification, attenuation).

A first experimental assessment consisted of a partial comparison to the only alternative method (i.e. RB). “Partial” basically means that, because of RB limitations (i.e. virtual object shapes required for a reasonable functioning), a full contrast could not be afforded. Subjects had to perform a task in which a certain VPC was created during pushing a virtual cube. After, a Fitts'-law-conform pointing movement had to be executed (i.e. sphere touching / popping subtask). In the end of the experiment, a sequential questionnaire was given. Results revealed substantial qualitative and quantitative advantages of HFC over RB. Although a special adjustment permitted harmonising the targets' ID between RB and HFC, there was no case found among the presented conditions with RB performing better than HFC. Instead, the corrected HFC proved again beneficial for near or, more precisely, near large targets. These findings in conjunction with the questionnaire responses and the task completion time regression analysis suggest an increased conspicuity of the RB treatments as well as a less intuitive, sometimes hindering or perturbing control under its influence. HFC overcomes most of these drawbacks.

An interesting supplementary detail was observed in the course of the main experiment. It was planned to test the novel approach in an interaction setting which should not be too artificial. The cube pushing subtask seemed to be a good choice for gradually inducing the VPC. Based on what is known and has been presented about pseudo-haptics in this thesis, it was speculated that a respective sensation would be provoked. It actually emerged, but not in all subjects and, of course, not in a manner directly qualifying it to be useful.

Most of the parameters incorporated into the HFC model should be regarded as an initial estimate for upcoming optimisation iterations. To improve the method's overall efficiency, a number of essential future steps may be undertaken, for instance:

- Verification of the existing empirical best practice measures.
- Refinement of the “implicit” multiplier weights.
- Study of exponential versus linear functional dependencies.
- Potential of a reduction method exchange on interaction events.
- Incorporation of a method to handle rotational VPC.
- Investigation of VPC fundamentals as stated in previous chapters.

Multifarious scenarios are imaginable, and some have already been demonstrated, in which controlled visuomotor discrepancies at hand level can simplify or add value to VE (see also Chapter 8). With the technique elaborated in this chapter, a first version of a generic tool can be provided that rounds off the conflict management cycle established by the original VPC framework.

7 General conclusion

This thesis has introduced and explored a novel form of 3D UI which is based on a visual repositioning of the user's real hand in space. Video see-through (VST) AR technology constitutes the preferred system ground, since it offers possibilities to perform visuomotor manipulations directly to the real limb rather than to an avatar. In previous research (e.g. Snijders et al., 2006), it has been indicated that seeing the own hand can raise the influence of vision on multisensory interaction. But not only the visual fidelity is important. Plausibly reflected actions (e.g. Vercher, 2006), here at arm or hand level, can again strengthen visual dominance and therefore the degree to which the user's perception can effectively be deceived by and adapt to the imposed visual conditions. Mainly for these reasons did the presented approaches rely on the embedding of stereoscopic real hand video feedback.

The framework designed for the simulation of static and dynamic visuo-proprioceptive conflicts (VPC) does also support occlusive VR (i.e. VE without any view of the real limb). Apart from this, it serves three principal purposes. Suppose it has successfully been linked to the underlying software system, then it can at first generate VPC by applying three-dimensional visual hand shifts in the vein of a model view transformation. Second, a generic API provides control to concrete applications. The internal module and data flow structure of the framework as well as specific displacement boundaries can be adapted to particular needs using configuration files. Third, the VPC management has been extended by a hand offset minimisation method which allows for an uninterrupted interaction, even in the advent of varying sensorimotor conflicts. The hand feedback convergence (HFC) computation considers the user's action state and a number of presets meant to make the treatment neither noticeable nor perturbing. It was shown in an experiment that the convergence parameters are not yet optimal. Amongst others, the offset reduction functions have to be adjusted. In spite of this, one might see the VPC framework as a comprehensive basis for interaction techniques aiming at the exploitation of spatial sensory discrepancies between the real hand and its visually perceived counterpart.

A virtual surface touching paradigm was developed on top of the above-described system. It should convey an intuitive feeling of control to the user by merging different beneficial near space interaction properties. These include a) acting in co-located space, b) using only visual movement constraints to prevent hand-scene intersections, c) preserving realistic occlusions and d) adding real hand feedback. The synthesised paradigm was experimentally tested for user acceptance and its impact on behaviour during a standard goal-directed pointing task. Analyses revealed that participants preferred the presumed most realistic hand representation (i.e. real hand video) over static 3D hand models of descending shape naturalness. But if the level of detail of the 3D hand substitute was high, then acceptance did not differ significantly compared to an embedded video feedback. Behavioural results yielded solely a larger lateral target overshooting in the case of the most artificial hand representation (i.e. ordinary 3D pointer arrow) suggesting a reduced quality of the perceptual link to the interaction tool. In summary, the subjective feeling of control can be improved with an increased limb realism. However, interaction performance is not affected a lot, if the hand gesture, like in the studied pointing scenario, does not really change. Hence, VE construction and implementation costs, for instance, for employing VST AR, should be traded against the final interaction purpose.

The main advantage of the proposed hand-related VPC is probably their potential to bias perception. As long as a few basic requirements are respected (e.g. principles of multisensory processing, adaptation and attribution, thresholds for sensory discrepancies and displacement dynamics), various novel methods and applications can be conceived for 3D UI and many other fields (see Chapter 8). Hand-displacement-based pseudo-haptics (HEMP) is, in this sense, a first attempt to exploit the involved perceptual mechanisms. The goal was to evoke an illusory sensation of force in the absence of any passive or active haptic device. It seems to be crucial for pseudo-haptics to deliver appropriate and actually integrable information about the force emitting or receiving, possibly purely virtual environmental property. That is, there must be an internal image or model of a perceived or applied force which can then be altered by vision. This brief theoretical explanation of the pseudo-haptics principles does apply to all known approaches (e.g. pseudo-haptic stiffness, torque and material resistance), perhaps even to the mouse-based desktop techniques. The force field (FF) illusion and response model (FIRE) was designed to trigger a base muscle activity and visually modulate the induced sensation. At the moment the visual hand is put into the virtual FF, it will be attracted and shifted along with the particle flow. A compensatory postural adjustment (CPA) at arm level

serves as the primary muscle response initiator. Voluntary sweeping movements are also qualified to produce a sufficient activity. To account for several perceptual constraints and VST HMD device limitations, the displacement velocity is progressively faded out. An experiment has been conducted in order to evaluate the HEMP FF concept. Indeed, the majority of the participants reported an active “stream-like” or “pushing” force. It could further be shown that the mean activity in the lateral arm flexor correlated with the simulated F. FIRE was thus able to generate varying contraction intensities. All presented F could be discriminated, but a significant accuracy decline was observed for smaller force differences and lower forces in general. The output of FIRE and the SNR in the sensory supply were most likely worst in these cases. After all, it should be repeated that this early HEMP example, although working under the described conditions, is not intended to replace comparable real haptic rendering.

To conclude, well-controlled VPC of the real hand employed in AR VE can evidently enrich existing 3D UI and add completely new methodological and application aspects. Nonetheless, as illustrated in the Sections 3.7, 4.8, 5.9 and 6.9, there remains still a lot of work. Improving the approaches, updating technologies and clarifying the questions raised would be essential next steps to finally come to productivity. The last chapter will venture some more future prospects of what else could be done starting from the contributions of this thesis.

8 FUTURE PERSPECTIVES

It has been shown that introducing visuo-proprioceptive conflicts (VPC) at hand level can add value to 3D UI in AR. The purpose of this chapter is to highlight further possibilities, not only focusing on the origin of the approach.

Visuomotor conflicts could be used for online hand trajectory modifications, to correct or stabilise real hand movements (see Section 8.1). It is also imaginable to operate on articular substructures of the hand or even other limbs (see Section 8.2).

8.1 Online hand trajectory manipulations

There are various situations in which it would be useful to visually manipulate the trajectory of the hand. In experimental psychology and neuroscience, for instance, this kind of planned perturbation has been applied to investigate possible interdependencies between vision and proprioception or the associated brain processes (e.g. “critical role of visual perception in trajectory formation” [Wo95b], modular decomposition of visuomotor maps [Gh97]). The setups used are often constrained in terms of action space, interactivity and immersion. The proposed framework based on AR would eliminate most of these drawbacks while providing a high level of perceptual reliance.

However, an interesting point in reconfiguring the hand position or trajectory perception is that the real hand's movement could be controlled exploiting the following fact. The natural hand transport from one location to another is usually performed straight. Visual guidance basically takes care of observed reaching error and corrects motor plans in order to achieve the desired goal. Given an optimal hand movement path in a specific scenario (e.g. hidden machine part assembly, surgery training or the remote control of robots). Then all deviations from it can be determined (e.g. using hand tracking), up to drawing near to dangerous zones. The question is, what to do with the retrieved error information? A classical user notification could be generated (e.g. visual, acoustic, haptic). But this would be likely to distract attention,

prolong the completion time or, in the worst case, disrupt the task completely. An alternative lower level intervention could profit from the automatic correction mechanism described above. Amplifying the actual drift based on application-dependent rules has therefore the potential to actuate a fast and unconscious compensatory reaction (see Fig. 8-1). Actual amplification functions may consider deviation risk categories (i.e. VPC development from linear to exponential).

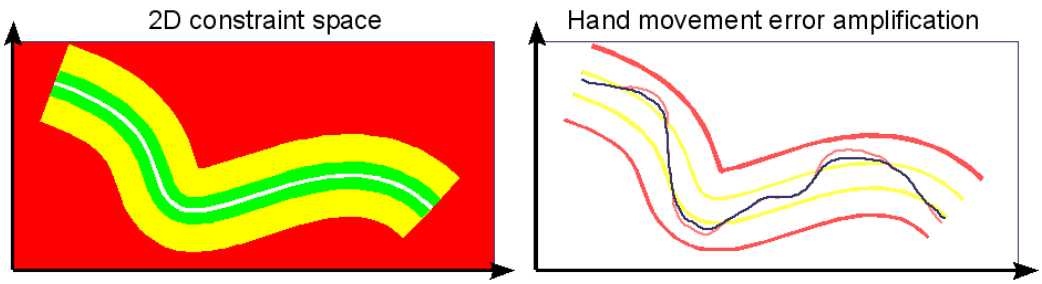


Figure 8-1: Hand movement error amplification,
left: optimal path (white), "good" (green), potentially critical (yellow) and dangerous (red) areas,
right: real hand movement (dark blue), visual hand movement with error amplification (salmon)

The same technique applies also to the opposite purpose: smoothing or stabilising the hand trajectory. At least two fields of applications are envisaged:

1. *Interaction acceleration.*

   If a rapid performance would have the priority (e.g. processing huge amounts of data, gaming), then it seems to be reasonable to reduce the pointing targets' ID (see Fitts, 1954). This could be achieved by decreasing the target distance and / or increasing the target size (see Eq. 6-5). The latter is often used to make graphical interfaces easier accessible by enlarging selection or menu items. Another opportunity would be to permit coarse, somewhat imprecise and only roughly co-located actions. Estimating the goal, eventually a hard task for itself, and displaying the hand or any substitute moving accordingly could offer a substantial performance gain. Whether this treatment can still be called VPC might be to debate, because the interaction spaces can become truly decoupled.

2. *Apparent tremor reduction.*

   Muscle tremor in arms and hands can be a pathological symptom of different disorders ranging from stress to Parkinson's Disease (PD) and cerebral dysfunctions or injuries. It causes limb shaking at varying intensities. Patients may not be able to perform well-directed actions or maintain limb posture. In real life, grasping objects, writing a letter or lifting a glass of water can be highly demanding. In VE, the system could support performing actions and reacquiring motor skills at a satisfying level. To this end, the particular tremor should be modelled as an oscillating error. A tremor learning function could facilitate hand movement prediction and enable the final error compensation (see Fig. 8-2). As a result, the patient would visually perceive less or no shaking at all. Their mental condition could improve and former activities be recovered.

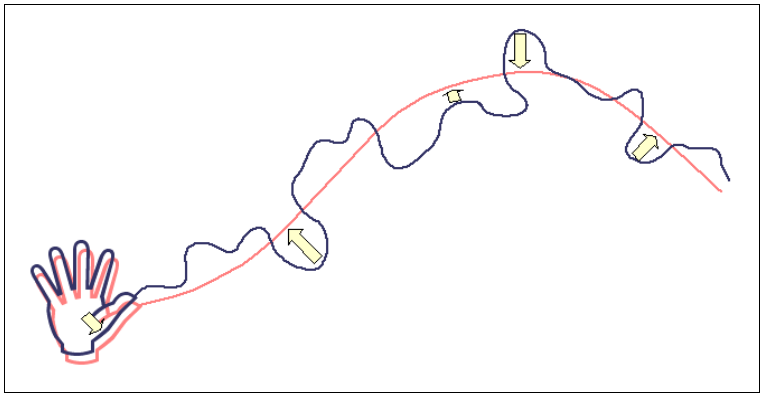


Figure 8-2: Hand movement stabilisation
(dark blue: real hand, salmon: visual hand)

8.2 Beyond positional visuo-proprioceptive conflicts of the hand

The current VPC framework can easily be extended to support angular hand representation discrepancies as they are used, for instance, in the rubber hand illusion (see Botvinick et al., 1998). Suchlike has been skipped so far because of the lightweight video embedding strategy. But it is, in fact, not sure whether real hand feedback is actually needed in all cases. Assumed it is not and perceptual reliance can be guaranteed, then a CyberGlove® could be used to track finger movements as well. These can directly be mapped to a high fidelity 3D hand model, perhaps textured with images of the user's real hand. A lot more manipulations would become possible this way.

First, all presented approaches can still be covered (i.e. intuitive virtual surface contacts, hand-displacement-based pseudo-haptics, HEMP). Second, rotational VPC at hand level would additionally allow for a deviceless simulation of active and passive torque (see also Paljic et al., 2004). As in the force field (FF) application, *active* refers to that the virtual phenomenon seems to exert a force on the user's limb. Once a full hand articulation capturing can be afforded, the degree of interaction realism could drastically be raised thanks to finger-based VPC. Touching, grasping and deforming virtual objects, the evocation of a compliance or stiffness feeling (see also Lécuyer et al., 2000) are just a few examples. To extend the simulation range, digital foam, an experimental isotonic input material, could serve as a support device [Sm08] (see Fig. 8-3).

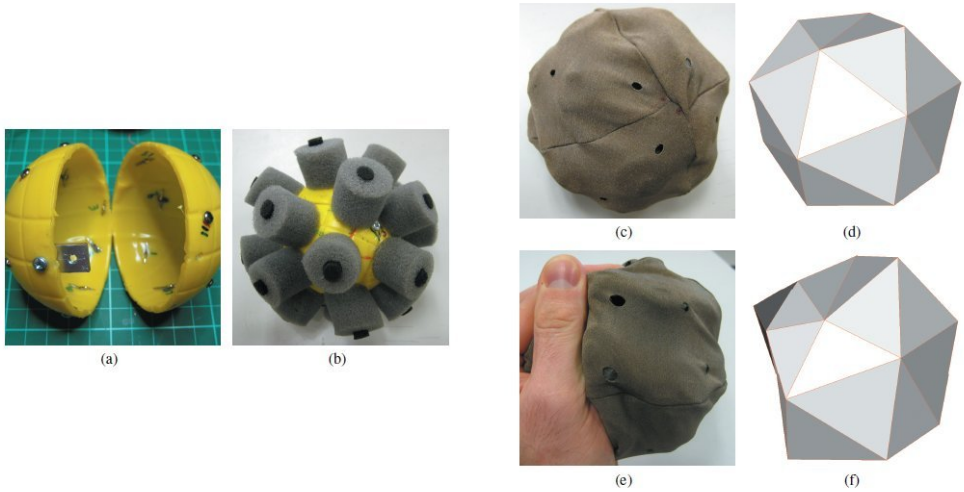


Figure 8-3: Digital foam (see Smith, 2008),
"(a) Plastic inner skeleton with sensor terminals. (b) Foam sensors attached to spherical prop. (c) Spherical prop with conductive fabric outer in place. (d) Geometry representation of sphere prop. (e) User squeezing part of the prop. (f) Geometry captured while user is squeezing the prop."

Only little is known about how such manipulation would affect multisensory interaction and behaviour. But what if people suffering from partial palsy or spasticity of hand or arm could again trigger apparent hand and finger movements, though they are incapable of performing them? "Noisy", limited or uncoordinated hand and finger movements could be translated into "normal" and smooth movements, grasping actions or tool usage. Here is certainly a big potential of controlled VPC.

Beside visually modulating proprioceptive information of arms, hands and fingers, another opportunity would be to alter the spatial appearance of legs and feet. This could possibly be helpful for psychophysical studies or interaction scenarios in which the lower extremities are visible to the user (e.g. walking through mud, wearing heavy shoes or performing impossible movements, see above).

Finally, and if the presumed theoretical background of pseudo-haptics can be confirmed, one might think of inducing force sensations, at which body level soever, in people who have lost their natural ability to perceive passive or active force feedback of natural phenomena. A summary of pathologies for which limb-displacement-based pseudo-haptics would be suitable could be the starting point for another multidisciplinary effort subserving human.

LIST OF FIGURES

Figure 1-1: Basic hand displacement principle.....9

Figure 1-2: Compositing of a user touching a virtual object, with the visual hand constrained on the object's surface and the real hand entering it.....11

Figure 1-3: Compositing of a user reaching into a virtual force field, including the visualised hand displacement to the right.....11

Figure 1-4: Gradual convergence effect until spatial alignment (dark blue: real hand, salmon: visual hand).....12

Figure 2-1: Psychometric functions for different cases of visuo-haptic sensory integration....17

Figure 2-2: Co-located space as the spheric tolerance surrounding "perfect co-location" (i.e. tip of the real index finger).....20

Figure 2-3: AR-VR system continuum (see Milgram et al., 1994).....23

Figure 2-4: Occlusion violation problem, top: given depth ordering, bottom: incorrect visual impression with the real hand always in front.....24

Figure 2-5: Real hand integration approach (see Ortega, 2007) using a VST HMD with built-in cameras (top right).....25

Figure 3-1: Existing system-subsystems scheme.....30

Figure 3-2: Overview of the existing software platform.....32

Figure 3-3: Foreground-background generation and separation.....33

Figure 3-4: Mixing approach indicating the vertices to manipulate of one hand texture carrier object (see Ortega, 2007).....36

Figure 3-5: Disparity distortion due to invariant relative camera viewpoints.....36

Figure 3-6: VPC framework model, including AR framework interaction.....38

Figure 4-1: Stringed Haptic Workbench, compositing of a user touching a virtual cube (see Tarrin et al., 2003).....44

Figure 4-2: RB-IM process on surface entering and release (dark blue: real hand, salmon: visual hand).....46

Figure 4-3: Hand representations used during the experiment (i.e. upper left: video feedback, upper right: detailed 3D model, lower left: simplified 3D model, lower right: 3D pointer model).....47

Figure 4-4: Virtual scene (left), with target locations (right).....49

Figure 4-5: Rest position (left), grasp and pointing gesture (right).....51

Figure 4-6: Application screenshot of a user touching the target (i.e. VST HMD view).....51

Figure 4-7: Effect of the hand representation on the coarse hand oscillation (means, SDs)....54

Figure 4-8: Effect of the hand representation on the target entering depth (means, SDs).....55

Figure 4-9: Effect of the hand representation on the hand movement duration (means, SDs). 56

Figure 4-10: Effect of the hand representation on the hand trajectory length (means, SDs)....57

Figure 4-11: Ranking of the hand representation as a function of the task aspects (means, SDs), with “marks” ranging from 1: best to 5: worst.....59

Figure 5-1: Size-weight illusion experiment.....66

Figure 5-2: Pseudo-haptic spring stiffness experiment (see Lécuyer et al., 2000), device (left) and visual feedback (right).....67

Figure 5-3: Pseudo-haptic torque experiment (see Paljic et al., 2004), device (top left, bottom) and visual feedback (top right).....67

Figure 5-4: Virtual Technical Trainer (see Crison et al., 2004), device (left) and visual feedback (right).....68

Figure 5-5: Virtual stream tube object, with the sensitive region highlighted (red).....69

Figure 5-6: Hand stabilisation within the flow (dark blue: real hand, salmon: visual hand, stabilisation area indicated by salmon triangle).....71

Figure 5-7: VST deviation angles for a just visible pointing.....74

Figure 5-8: FIRE (i.e. FF illusion and response model), with all internal state levels and the displacement velocity development with level correspondences (bottom left).....75

Figure 5-9: Rest position (left), hand tracking body fixation (right).....81

Figure 5-10: Close-up of the stream tube object, with particle flow and the hand stabilisation region indicator (i.e. red sphere).....81

Figure 5-11: Application screenshot of a user reaching into the FF (i.e. VST HMD view).....82

Figure 5-12: Musculus pectoralis major and the EMG sensor positions chosen for the experiment.....83

Figure 5-13: Influence of the simulated FF levels on the mean pectoralis major activity.....86

Figure 5-14: Influence of the force combination differences on the comparison performance (means, SDs).....87

Figure 5-15: Influence of the force combination zones on the comparison performance (means, SDs).....88

Figure 5-16: Influence of the force combination senses (red: greater force in the first trial, blue: in the second trial) on the comparison performance (means, SDs).....89

Figure 6-1: Potentially problematic VPC cumulation due to repeated surface contacts or touching events (dark blue: real hand, salmon: visual hand).....98

Figure 6-2: CALM / MACBETH situated between RB and IM (see Burns et al., 2005).....99

Figure 6-3: HFC embedded into the VPC framework.....101

Figure 6-4: Human visual field (see Werner, 1991).....103

Figure 6-5: HFC model (i.e. hand feedback convergence model).....104

Figure 6-6: Visual and motor equivalence (dark blue: real hand, salmon: visual hand).....106

Figure 6-7: Pointing gesture, with virtual arrow overlay.....109

Figure 6-8: Interaction task sequence (dark blue: real hand / arrow position, salmon: visual hand / arrow position).....110

Figure 6-9: Global method and target effects on the task completion time (means, SDs)..…115

Figure 6-10: Target-ID-to-mean-movement-time correlation at global level (top) and for each offset reduction method.....116

Figure 6-11: Influence of the offset reduction method on the task completion time for near targets (means, SDs).....117

Figure 6-12: Influence of the offset reduction method on the task completion time for near large targets (means, SDs).....119

Figure 8-1: Hand movement error amplification, left: optimal path (white), "good" (green), potentially critical (yellow) and dangerous (red) areas, right: real hand movement (dark blue), visual hand movement with error amplification (salmon).....130

Figure 8-2: Hand movement stabilisation (dark blue: real hand, salmon: visual hand).....131

Figure 8-3: Digital foam (see Smith, 2008), "(a) Plastic inner skeleton with sensor terminals. (b) Foam sensors attached to spherical prop. (c) Spherical prop with conductive fabric outer in place. (d) Geometry representation of sphere prop. (e) User squeezing part of the prop. (f) Geometry captured while user is squeezing the prop.".....132

LIST OF TABLES

Table 4-1: Overall assessment means for the hand representations used.....58

Table 5-1: Condition square for the comparison task: greater force was either presented in the first trial (condition 8 to 14) or in the second trial (condition 1 to 7).....80

Table 5-2: Condition ranking and overall results. Each rank shows: the condition with respect to the condition square (cond.), the condition-dependent force level difference (diff.), zone and sense as well as the related mean success rate and SD.....87

Table 6-1: Target conditions.....107

Table 6-2: The targets' IDs.....107

Table 6-3: Pairwise comparison of the offset reduction methods I: near and far.....117

Table 6-4: Pairwise comparison of the offset reduction methods II: small and large.....118

Table 6-5: Pairwise comparison of the offset reduction methods III: crossed target properties. 118

LIST OF ABBREVIATIONS

3D UI.....3D user interface(s)

ANOVA.....Analysis of Variance

API.....Application programming interface

AR.....Augmented Reality

AV.....Augmented Virtuality

C/D ratio.....Control-to-display ratio

CAD.....Computer Aided Design

CALM.....Credible Avatar Limb Motion technique

CNS.....Central nervous system

CPA.....Compensatory postural adjustment

EMG.....Electromyography

FF.....Force field

FIRE.....FF illusion and response model

fMRI.....Functional magnetic resonance imaging

FT.....Feasibility test

HCI.....Human-computer interaction

HEMP.....Hand-displacement-based pseudo-haptics

HFC.....Hand feedback convergence

HFC-M.....HFC – motor equivalence

HFC-V.....HFC – visual equivalence

HMD.....Head-mounted display

HSL.....Hue Saturation Luminance colour scheeme

IM.....Incremental motion method

JND.....Just Noticeable Difference

LCD.....Liquid crystal display(s)

LSD.....Least Significant Difference

MACBETH.....Management of Avatar Conflict By Employment of Technique Hybrid

miniOSG.....Acronym for the laboratory's own AR / VR platform

MLE.....Maximum Likelihood Estimate

MR.....Mixed Reality

PET.....Positron Emission Tomography

PSE.....Point of Subjective Equality

RB.....Rubber band method

RGB.....Red Green Blue colour scheeme

SD.....Standard deviation

SNR.....Signal-to-noise ratio

VE.....Virtual Environments

VPC.....Visuo-proprioceptive conflict(s)

VR.....Virtual Reality

VRPN.....VR peripheral network

VST.....Video see-through

REFERENCES

[**Bl03**] S.-J. Blakemore and C. D. Frith, “Self-Awareness and Action”, In Current Opinion in Neurobiology. 13 (2), pp. 219 - 224, 2003.

[**Bo98**] M. Botvinick and J. Cohen, “Rubber Hands "Feel" Touch that Eyes See”, In Nature 391, 756, 1998.

[**Bu03**] G. Burdea and P. Coiffet, “Virtual Reality Technology, Second Edition”, Wiley-IEEE Press, 2003.

[**Bu05**] E. Burns, S. Razzaque, A. T. Panter, M. C. Whitton, M. R. McCallus, and F. P. Brooks, Jr., “The Hand is Slower than the Eye: A Quantitative Exploration of Visual Dominance Over Proprioception”, In Proceedings of IEEE VR 2005, Bonn, Germany, 2005.

[**Bu07**] E. Burns, S. Razzaque, M. C. Whitton, and F. P. Brooks, Jr., “MACBETH: The Avatar which I See Before Me and its Movement Toward My Hand”, In Proceedings of IEEE VR 2007, Charlotte, USA, 2007.

[**Cr92**] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, “The CAVE: Audio Visual Experience Automatic Virtual Environment”, In Proceedings of SIGGRAPH 1992, New York, 1992.

[**Cr04**] F. Crison, A. Lécuyer, A. Savary, D. Mellet-d'Huart, J. M. Burkhardt, and J. L. Dautin, “The Use of Haptic and Pseudo-Haptic Feedback for the Technical Training of Milling”, EuroHaptics Conference, Munich, Germany, 2004.

[**Ed93**] E. K. Edwards, J. P. Rolland, and K. P. Keller, “Video See-Through Design for Merging of Real and Virtual Environments”, In Proceedings of VRAIS 1993, Seattle, USA, 1993.

[**Eh04**] H. H. Ehrsson, C. Spence, and R. E. Passingham, “That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb”, In Science MagazineVol. 305 (5685), pp. 875 - 877, 2004.

[**Er02**] M. O. Ernst and M. S. Banks, “Humans Integrate Visual and Haptic Iinformation in a Statistically Optimal Fashion”, In Nature 415, 429 - 433, 2002.

[**Er04**] M. O. Ernst and H. H. Bülthoff, “Merging the Senses into a Robust Percept”, In Trends in Cognitive Sciences Vol. 8 (4), 2004.

[**Fi54**] P. M. Fitts, “The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement”, In Journal of Experimental Psychology 47, pp. 381 - 391, 1954.

[**Gh97**] Z. Ghahramani and D. M. Wolpert, “Modular Decomposition in Visuomotor Learning”, In Nature 386, pp. 392 - 395, 1997.

[**Ha06**] M. Havryliv, G. Schiemer, and F. Naghdy, “Haptic Carillon: Sensing and Control in Musical Instruments”, In Proceedings of the Australasian Computer Music Conference, 2006.

[**Ki00**] K. Kiyokawa, Y. Kurata, and H. Ohno, “An Optical See-Through Display for Mutual Occlusion of Real and Virtual Environments”, In Proceedings of IEEE & ACM ISAR 2000, Munich, 2000.

[**Kr95**] W. Krüger, C. A. Bohn, B. Fröhlich, H. Schüth, W. Strauss, and Gerold Wesche, “The Responsive Workbench: A Virtual Work Environment”, In IEEE Computer Vol. 28, 1995.

[**Le00**] S. M. Lephart and F. H. Hu (eds.), “Proprioception and Neuromuscular Control in Joint Stability”, Human Kinetics, 2000.

[**Lé00**] A. Lécuyer, S. Coquillart, A. Kheddar, P. Richard, and P. Coiffet, “Pseudo-Haptic Feedback: Can Isometric Input Devices Simulate Force Feedback?”, In Proceedings of IEEE VR 2000, Washington, DC, USA, 2000.

[**Lé01**] A. Lécuyer, J.-M. Burkhardt, S. Coquillart, and P. Coiffet, “'Boundary of Illusion': an Experiment of Sensory Integration with a Pseudo-Haptic System”, In Proceedings of IEEE VR 2001, Yokohama, Japan, 2001.

[**Lo03**] B. Lok, S. Naik, M. Whitton, and F. P. Brooks, Jr., “Effects of Handling Real Objects and Avatar Fidelity on Cognitive Task Performance in Virtual Environments”, In Proceedings of IEEE VR 2003, Washington, DC, USA, 2003.

[**Me02**] K. van Mensvoort, “What You See is What You Feel.”, In Proceedings of Designing Interactive Systems 2002, 2002.

[**Me08**] K. van Mensvoort, D. J. Hermes, and M. van Montfort, “Usability of Optically Simulated Haptic Feedback”, In International Journal of Human-Computer Studies 66, pp. 438 - 451, 2008.

[**Mi94**] P. Milgram and F. Kishino, “A Taxonomy of Mixed Reality Visual Displays”, IEICE Transactions on Information Systems, Vol. E77-D (12), 1994.

[**Mi97**] M. R. Mine, F. P. Brooks Jr., and C. H. Sequin, “Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction”, In Proceedings of SIGGRAPH '97, 1997.

[**MK92**] I. S. MacKenzie, “Fitts' Law as a Research and Design Tool in Human-Computer Interaction”, In Human-Computer Interaction Vol. 7, pp. 91 - 139, 1992.

[**Mu99**] D. J. Murray, R. R. Ellis, C. A. Bandomir, and H. E. Ross, “Charpentier (1891) on the Size-Weight Illusion”, In Perception & Psychophysics 61 (8), 1999.

[**OCV**] OpenCV, “Open Source Computer Vision Library, Intel Corporation”, http://www.intel.com/technology/computing/opencv, visited: May 2008.

[**OGL**] OpenGL, “Open Graphics Library, SGI”, http://www.opengl.org, visited: May 2008.

[**Oh99**] Y. Ohta and H. Tamura (ed.), “Mixed Reality: Merging Real and Virtual Worlds”, Springer-Verlag New York Inc., New York, 1999.

[**Or06**] M. Ortega, S. Redon, and S. Coquillart, “A Six Degree-of-Freedom God-Object Method for Haptic Display of Rigid Bodies”, In Proceedings of IEEE VR 2006, 2006.

[**Or07**] M. Ortega, “Visuo-Haptic Solutions for Virtual Prototyping: Automotive Applications”, Ph.D. dissertation, Université Claude Bernard Lyon 1, 2007.

[**OSG**] OpenSG, “Open Source Scene Graph”, http://opensg.vrsource.org/trac, visited: May 2008.

[**Pa96**] J. Paillard, “Fast and Slow Feedback Loops for the Visual Correction of Spatial Errors in a Pointing Task: A Reappraisal”, In Canadian Journal of Physiology and Pharmacology 74, pp. 401 - 417, 1996.

[**Pa02**] A. Paljic, S. Coquillart, J.-M. Burkhardt, and P. Richard, “A Study of Distance of Manipulation on the Responsive Workbench(tm)”, In Proceedings of IPT Symposium 2002, Orlando, USA, 2002.

[**Pa04**] A. Paljic, J.-M. Burkhardt, and S. Coquillart, “Evaluation of Pseudo-Haptic Feedback for Simulating Torque: A Comparison between Isometric and Elastic Input Devices”, HAPTICS '04, Chicago, USA, 2004.

[**Pe01**] D. Perani, F. Fazio, N. A. Borghese, M. Tettamanti, S. Ferrari, J. Decety, and M. C. Gilardi, “Different Brain Correlates for Watching Real and Virtual Hand Actions”, In Neuroimage 14 (3), 2001.

[**Sc05**] R. A. Scheidt, M A. Conditt, E. L. Secco, and F. A. Mussa-Ivaldi, “Interaction of Visual and Proprioceptive Feedback During Adaptation of Human Reaching Movements”, In Journal of Neurophysiology 93 (6), 2005.

[**Sm08**] R. T. Smith, B. H. Thomas, and W. Piekarski, “Digital Foam”, In Proceedings of IEEE 3DUI 2008, Reno, USA, 2008.

[**Sn06**] H. J. Snijders, N. P. Holmes, and C. Spence, “Direction-Dependent Integration of Vision and Proprioception in Reaching Under the Influence of the Mirror Illusion”, In Neuropsychologia (ePub), 2006.

[**St01**] A. State, J. Ackerman, G. Hirota, J. Lee, and H. Fuchs, “Dynamic Virtual Convergence for Video See-through Head-Mounted Displays: Maintaining Maximum Stereo Overlap throughout a Close-Range Work Space”, In Proceedings of IEEE & ACM ISAR 2001, 2001.

[**Ta03**] N. Tarrin, S. Coquillart, S. Hasegawa, L. Bouguila, and M. Sato, “The Stringed Haptic Workbench: A New Haptic Workbench Solution”, In Proceedings of EUROGRAPHICS, 2003.

[**Ve06**] J.-L. Vercher, “Perception and Synthesis of Biologically Plausible Motion: From Human Physiology to Virtual Reality.”, In Lecture Notes in Computer Science 3881, pp. 1 - 12, 2006.

[**We91**] E. B. Werner, “Manual of Visual Fields”, Churchill Livingstone, New York, 1991.

[**We08**] R. B. Welch and A. C. Sampanes, “Adapting to Virtual Environments: Visual-Motor Skill Acquisition Versus Perceptual Recalibration”, In Displays 29, pp. 152 - 158, 2008.

[**Wi98**] B. G. Witmer and M. J. Singer, “Measuring Presence in Virtual Environments: A Prescence Questionnaire”, In Teleoperators and Virtual Environments 7 (3), 1998.

[**Wi02**] S. P. Wise and R. Shadmehr, “Motor Control”, In Encyclopedia of Human Brain Vol. 3, 2002.

[**Wo95a**] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, “An Internal Model for Sensorimotor Integration”, In Science Vol. 269 (5232), pp. 1880 - 1882, 1995.

[**Wo95b**] D. M. Wolpert, Z. Ghahramani, and M. I. Jordan, “Are Arm Trajectories Planned in Kinematic or Dynamic Coordinates? An adpation study”, In Experimental Brain Reserach, 1995.

[**Za01**] G. Zachmann and A. Rettig, “Natural and Robust Interaction in Virtual Assembly Simulation”, In Eighth ISPE CE: Research and Applications 2001, 2001.

APPENDICES

Subjects were generally not allowed to return to previous questionnaire pages. ***Page breaks*** indicate thus actual processing steps.

A.1 Questionnaire of the experiment of Chapter 4

1. Did you perceive any difference between the experimental trials?
2. If you perceived differences between the the experimental trials, how many and which differences did you perceive?
3. Which experimental condition(s) do you prefer from a global comfort point of view? WHY?
4. What helped most to perform the tasks (from a general point of view)?
5. What perturbed most when performing the tasks (from a general point of view)?

<Page break>

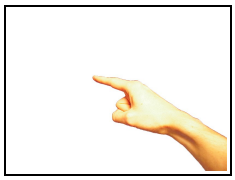
6. How many different visual hand representations did you perceive?
7. Please draw (sketch) each visual hand representation that you saw.
8. Which visual hand presentation(s) did you prefer? WHY?

<Page break>

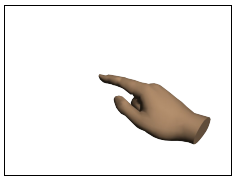
Assessment section (please give marks)

The following four visual hand representations were used during the experiment

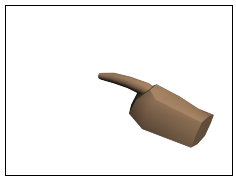
(1)



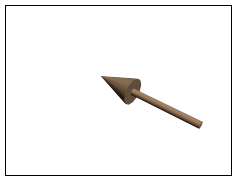
(2)



(3)



(4)



9. Visual hand representations used for pointing

   (1: very good => 5: bad)

Hand representation	1	2	3	4	5
(1)					
(2)					
(3)					
(4)					

10. Final pointing accuracy on the target as a function of the visual hand representation

    (1: very good => 5: bad)

Hand representation	1	2	3	4	5
(1)					
(2)					
(3)					
(4)					

<Page break>

11. Naturalness of the hand movement / transport towards the target as a function of the visual hand representation
    (1: very good / intuitive => 5: bad / abstract)

Hand representation	1	2	3	4	5
(1)					
(2)					
(3)					
(4)					

12. Overall comfort while performing the tasks as a function of the visual hand representation
    (1. comfortable => 5: uncomfortable)

Hand representation	1	2	3	4	5
(1)					
(2)					
(3)					
(4)					

Explanation of the reasons for the best AND the worst assessment

13. General remarks

A.2 Questionnaire of the experiment of Chapter 5

1. Please describe your sensation when exposing your hand to the visual flow.
2. If the sensation has changed over time, in WHICH way / HOW did it change?
3. If you have perceived differences between the trials, HOW MANY and WHICH differences did you perceive? Please describe each difference.
4. How much did your experiences in the virtual environment seem consistent with your real-world experiences? (1: inconsistent => 7: consistent)

1	2	3	4	5	6	7

Explanation of the score

5. In case you felt confused or disoriented at some point during the experiment, WHEN, HOW STRONG and WHY did this happen?
6. General remarks

<Page break>

7. Please describe the indicator(s) you used to determine which of the two consecutive trials to compare was actually the “stronger” one.

A.3 Questionnaire of the experiment of Chapter 6

Recall of the task (hand movement sequence):

Each trial consisted of 4 main phases:

(a) Starting from the rest position, touching the right side of the cube.

(b) Pushing the cube until it got in contact with the red square.

(c) Once the red square was reached, touching the sphere.

(d) After the sphere was touched, returning to the rest position.

1. For EACH phase (a) to (d), please describe (globally) the sensation you had.
2. If something appeared to be "strange" or unusual during the experiment, please note WHAT you observed and WHEN it appeared (WHEN during a trial, in which phase).

<Page break>

Recall of the task (hand movement sequence):

Each trial consisted of 4 main phases:

(a) Starting from the rest position, touching the right side of the cube.

(b) Pushing the cube until it got in contact with the red square.

(c) Once the red square was reached, touching the sphere.

(d) After the sphere was touched, returning to the rest position.

3. Target distances and sizes differed between trials. Did you notice any other difference(s) (Y/N)? If so, please describe possible difference(s) and WHEN it / they appeared (e.g. WHEN during a trial or the experiment).

<Page break>

Recall of the task (hand movement sequence):

Each trial consisted of 4 main phases:

(a) Starting from the rest position, touching the right side of the cube.

(b) Pushing the cube until it got in contact with the red square.

(c) Once the red square was reached, touching the sphere.

(d) After the sphere was touched, returning to the rest position.

4. In phase (c) in particular, beside differing target distances and sizes, do you think there were any other difference(s) between trials (Y/N)? If so, please describe possible difference(s) and WHEN it / they appeared.

<Page break>

Recall of the task (hand movement sequence):

Each trial consisted of 4 main phases:

(a) Starting from the rest position, touching the right side of the cube.

(b) Pushing the cube until it got in contact with the red square.

(c) Once the red square was reached, touching the sphere.

(d) After the sphere was touched, returning to the rest position.

5. In case you noticed difference(s) during phase (c), except for target distances and sizes, do you remember any perturbing or uncomfortable situation(s) (Y/N)? If so, please describe the possibly perturbing situation(s) and WHEN it / they appeared.

<Page break>

6. In case you felt perturbed or uncomfortable at some point during the experiment, please describe WHAT you felt and WHEN this happened.
7. General remarks