A Lightweight AV System for Providing a Faithful and Spatially Manipulable Visual Hand Representation

Andreas Pusch, Olivier Martin, Sabine Coquillart

A Lightweight AV System for Providing a Faithful and Spatially Manipulable Visual Hand Representation

Andreas PUSCH[a, 1], Olivier MARTIN[b, 2], and Sabine COQUILLART[b]
[a] *IIHM – UJF – LIG, France*
[b] *INRIA Grenoble Rhône-Alpes – LIG, France*

Abstract. This paper introduces the technical foundations of a system designed to embed a lightweight, faithful and spatially manipulable representation of the user's hand into an otherwise virtual world (aka Augmented Virtuality, AV). A highly intuitive control during pointing-like near space interaction can be provided to the user as well as a very flexible means to experimenters in a variety of non-medical and medical contexts. Our approach essentially relies on stereoscopic video see-through Augmented Reality (AR) technology and a generic, extendible framework for managing 3D visual hand displacements. Research from human-computer interaction, perception and motor control has contributed to the elaboration of our proposal which combines a) acting in co-location, b) avoiding occlusion violations by assuring a correct scene depth ordering and c) providing a convincing visual feedback of the user's hand. We further present two cases in which this system has already successfully been used and then line out some other applications that we think are promising, for instance, in the fields of neuromotor rehabilitation and experimental neuroscience.

Keywords. Augmented Virtuality, video see-through head-mounted display, co-location, visuo-proprioceptive sensory conflict, visual hand shift framework

Introduction

Experimental setups typically used to manipulate the perceived location of the own hand are often very restrictive for subjects or patients who are forced to take specific artificial fixed postures and have only little space to move the interacting limb around. Moreover, the visual hand feedback is often reduced to ordinary cursors or otherwise oversimplified while it is known that the hand representation quality can have a strong impact on the feeling of limb ownership, self-action recognition, thus on the reliability in the display and so on the performed actions. Different types, amounts and variations of visuo-proprioceptive conflicts (VPC) can rarely be simulated by the same system.

Independent of the actual purpose or application context, be it the investigation of human multimodal perception, the study of certain symptoms of neuromotor disorders or the development of novel rehabilitation methods for motor skill recovery, it seems that the range of possibilities and the richness of the tools can, at fairly low cost, still be increased. To approach such a multipurpose solution is what we aim at in this paper. We also think that our system which does not require complex 3D hand reconstruction techniques can be of great interest to human-computer interaction.

1 Corresponding author: IIHM – UJF – LIG, 110, av. de la Chimie – BP 53, 38041 Grenoble Cedex 9, France; E-mail: andreas.pusch@imag.fr
2 Currently at: GIPSA-lab – UJF, 961, rue de la Houille Blanche – BP 46, 38402 Grenoble Cedex, France

1. Related work

Vision of the own hand affects efficiency and comfort during direct manipulation even of virtual objects. A virtual hand-substitute is perceived as faithful and better attributed to the own body [1], if it reflects biologically plausible motion [2, 3] and looks natural or familiar [4, 5]. High fidelity visual hand feedback has further been demonstrated to produce a strong visuo-proprioceptive integration [6], shorter reaction and movement times compared to a lower fidelity hand feedback [7] and less position estimation errors [8]. The rubber hand illusion [9] has recently been replicated in Virtual Reality (VR) and it seems that its effectiveness depends also on visual limb faithfulness [10].

Co-location refers to the visuo-proprioceptive alignment of an interacting limb. Qualitative and quantitative benefits could be shown [11, 12] as well as a stronger sense of presence [13]. All this is supporting the self attribution process [9] and thus favours a continuous natural observation and correction loop that is responsible for the compensation of pointing errors or target modifications [14], also in VR settings [15].

Whenever VPC have been produced, for instance, by visually simulating haptic constraints [16, 17] or inducing force illusions [18], it might be useful to return to "normal" visuomotor conditions (i.e., no VPC). To limit both a potential performance decline and VPC aftereffects (e.g., perceptual recalibration [19]), there are approaches to reduce VPC [16, 20]. The latter aims at a fast, unconscious offset reduction.

2. Technical system foundations

Our AV [21] systems represents a generalisation of the system used in [18]. It merges the advantages of a) acting in co-location, b) avoiding occlusion violations (i.e., the hand can occlude and be occluded by virtual objects) and c) providing a convincing visual feedback of the user's hand. We further aim at spatial employment flexibility and system integration ease.

The system is composed of four network-enabled building blocks: 1) The video see-through [22] head-mounted display (HMD) with two built-in (stereo) cameras, 2) the video acquisition and post processing unit also used for image correction and background segmentation, 3) the tracking system that tracks head and hand at 6 degrees of freedom and 4) the compositing unit performing realtime simulation, interaction and rendering.

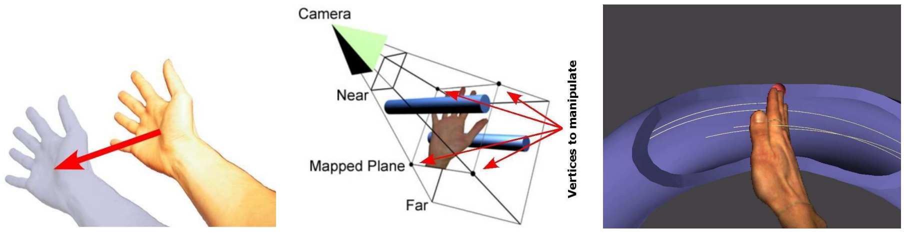

Figure 1. Here we illustrate all key elements of our approach. Left: The overall concept of displacing the hand visually in 3D space using live video data captured in stereo by the HMD's built-in cameras. Middle: The mixing approach generalising [18]. We provide an extendible framework for creating and managing VPC that operates on the vertices of the carrier geometry mapped with the segmented live video hand. Virtual objects in front occlude the hand correctly as virtual objects behind the hand are correctly occluded. Right: A screenshot from a study on perceptual illusions [18] we conducted using an earlier version of the presented system. Subjects had to expose their hand (viewed through the video see-through HMD) to a virtual force field that attracted the "visual hand" and shifted it progressively away from its real counterpart. The triggered motor reaction in the pectoralis major produced the illusion of a flow that pushes the hand / arm to the side.

On top of this basic infrastructure resides our VPC generation and management framework (henceforth: VPC framework) that offers a highly configurable interface to the top level application layer. The VPC framework does all the work, including the dynamic control (i.e., increase or reduce) of visual hand shifts, performing feasibility tests and storing intermediate runtime data for post-hoc analyses. In Figure 1, the essential steps are summarised. As the hardware and software platforms are similar to our configuration presented in [18], we will, in the next section, primarily focus our attention on the novel platform-independent VPC framework.

3. VPC generation and management

The presented lightweight solution for embedding the user's hand into a virtual scene can easily be exploited for static and dynamic visual hand displacements. Since the hand texture carrier objects consist of simple OpenGL quads redefined at runtime, their spatial attributes are available for additional computations. 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. A visual repositioning of the real hand can thus be achieved by shifting these objects or rather their geometry within the different frames of reference.

To manage static offsets as well as dynamic displacements in either the local viewing or the global world frame, we provide a data flow concept and a set of basic control functions. At the lowest level of the AR / AV mixing [23], the only thing we need is access to the carrier quads. Note that enabling this access represents the only change to be made to the kernel, whatever the actual AR / VR platform is.

A concrete application can trigger a fixed or a continuous shift 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. Parameters are pushed to the 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 computes m_{HS} assuming a linear offset development. Using the hand feedback convergence (HFC) flag, applications can activate an automatic process as MACBETH [20] or any other (e.g., [16]) to reduce a given offset while assuring the best possible motion coherence. Such a reduction helps return to "normal" visuomotor conditions and limit the risk of placing the visual hand outside the display space over time. $\vec{o}$ would be updated in this case before it is assigned to the hand texture carrier objects. Through a convergence amplifier, it is possible to control the impact of the HFC process. 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, risk of perceptual stress). The test is based on the estimated visual-to-real hand deviation angle and returns an error message, if predefined constraints have been violated. Alternatively, the hand representation can fall back to any static 3D hand model [1-2]. Offset thresholds are declared in a configuration file. It is further recommended to experimentally determine the true device properties for an optimised VPC bandwidth. m_{HS} will finally be applied to the viewing or world 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 analyses.

Generality and extendibility are important characteristics of the VPC framework. Beside the open framework entries, also m_{HS} is generic in terms of its applicability.

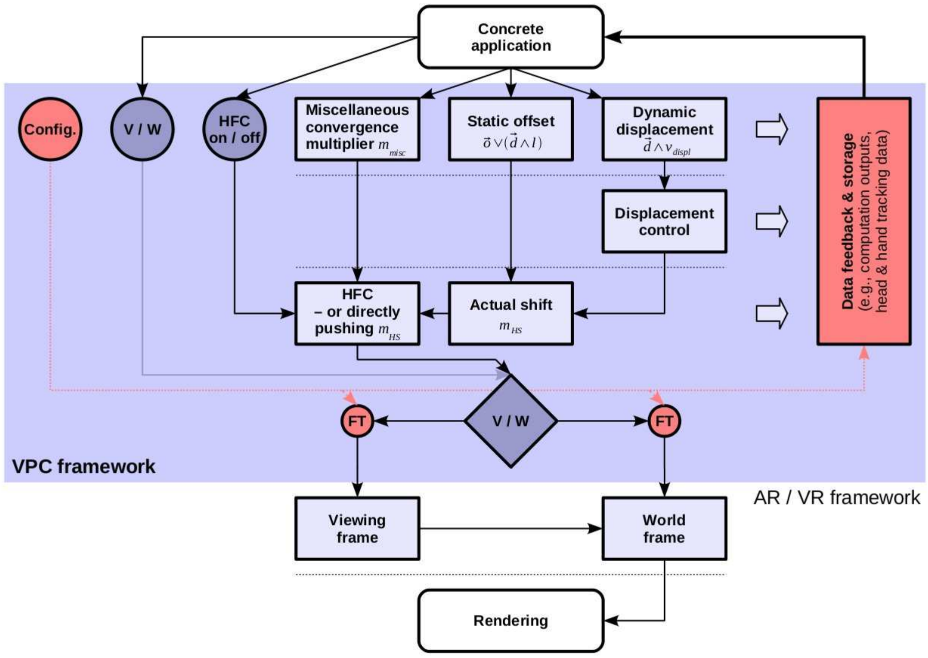

Figure 2. The internal architecture of the VPC framework. For details, please refer to the text.

Even virtual hand avatars can be managed using this matrix. Although the upper two 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.

4. Conclusion and future work

We have designed and implemented a novel lightweight near space interaction system generalising [18] that aims at a better stimulation of the processes involved in natural visuomotor coordination. The system includes a generic, extendible framework capable of inducing static and dynamic spatial VPC at hand level. We have recently used this setup to investigate effects of hand feedback fidelity on motor performance and user acceptance in a virtual object touching task [8]. Other promising fields of applications can be found where controlled VPC at hand level are to be combined with hand / finger visualisation and motion naturalness (e.g., studies of multimodal perception and online manipulation of visually guided actions, also for early diagnostics and rehabilitation of specific neuromotor disorders). We wish to perform conformity evaluations in order to compare our system to well-established, though mostly much less flexible, tools.

Further work is required to overcome technical limitations as incomplete stereo for very near objects due to insufficient viewing frustum overlaps in this region (possible solution: [24]), linear cutting edges on virtual object intersections due to using planar texture carrier objects (possible solution: Simplified 3D proxies for the computation of intersection effects; could additionally cast shadows) or the impression of having one's arm cut at the display boundary, if the hand is shifted away into the opposite direction (possible solution: Texture extension or repetition). Moreover, multimodal stimulation (e.g., acoustic and vibrotactile) does either partially exist or can easily be added.

Acknowledgments

This work has partially been supported by the French Research Agency ANR, by the region Rhône-Alpes and by the Intuition NoE EU project. Andreas Pusch has been funded by a grant from the EU under the Marie-Curie actions within the VISITOR project and has partially been supported by the Minalogic project NOMAD.

References

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

[2] 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.

[3] 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.

[4] 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 Proc. of IEEE VR, 2003.

[5] M. Tsakiris and P. Haggard, *The Rubber Hand Illusion Revisited: Visuo-Tactile Integration and Self-Attribution*, In J. of Exp. Psych.: Human Perception and Performance, 31, pp. 80 – 91, 2005.

[6] 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.

[7] P. J. Durlach, J. Fowlkes, and C. J. Metevier, *Effect of Variations in Sensory Feedback on Performance in a Virtual Reaching Task*, In Presence: Teleoperators & Virtual Env., 14 (4), pp. 450 – 462, 2005.

[8] A. Pusch, O. Martin, and S. Coquillart, *Effects of Hand Feedback Fidelity on Near Space Pointing Performance and User Acceptance*, to appear in Proc. of ISVRI, 2011.

[9] M. Botvinick and J. Cohen, *Rubber Hands "Feel" Touch that Eyes See*, In Nature, 391 (756), 1998.

[10] Y. Yuan and Anthony Steed, *Is the Rubber Hand Illusion Induced by Immersive Virtual Reality?*, In Proc. of IEEE VR, 2010.

[11] M. R. Mine, F. P. Brooks Jr., and C. H. Sequin, *Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction*, In Proc. of SIGGRAPH, 1997.

[12] 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.

[13] M. Slater and A. Steed, *A Virtual Presence Counter*, In Presence: Teleoperators & Virtual Environments, 9 (5), pp. 413 – 434, 2000.

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

[15] O. Martin, B. Julian, L. Boissieux, J.-D. Gascuel, and Claude Prablanc, *Evaluating online control of goal-directed arm movement while standing in virtual visual environment*, In The J. of Visualization and Computer Animation, 14, pp. 253 – 260, 2003.

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

[17] 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 Proc. of IEEE VR, 2005.

[18] A. Pusch, O. Martin, and S. Coquillart, *HEMP – Hand-Displacement-Based Pseudo-Haptics: A Study of a Force Field Application*, In Proc. of IEEE 3DUI, 2008.

[19] 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.

[20] 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 Proc. of IEEE VR, 2007.

[21] P. Milgram and F. Kishino, *A Taxonomy of Mixed Reality Visual Displays*, In IEICE Transact. on Inf. Systems, E77-D (12), 1994.

[22] E. K. Edwards, J. P. Rolland, and K. P. Keller, *Video See-Through Design for Merging of Real and Virtual Environments*, In Proc. of VRAIS, 1993.

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

[24] 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 Proc. of IEEE & ACM ISAR, 2001.