Lessons from digital puppetry - Updating a design framework for a perceptual user interface

Ferguson, J.

This is a copy of the author's accepted version of a paper subsequently published in the proceedings of the *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing - The International Workshop on Virtual Environments and Advanced Interfaces* Liverpool, UK, 26 - 28 Oct 2015, IEEE.

It is available online at:

https://dx.doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.239

Lessons from digital puppetry

Updating a design framework for a perceptual user interface

Jeffrey Ferguson
Faculty of Science and Technology
University of Westminster
London, UK
J.Ferguson@Westminster.ac.uk

***Abstract*—While digital puppeteering is largely used just to augment full body motion capture in digital production, its technology and traditional concepts could inform a more naturalized multi-modal human computer interaction than is currently used with the new perceptual systems such as Kinect. Emerging immersive social media networks with their fully live virtual or augmented environments and largely inexperienced users would benefit the most from this strategy. This paper intends to define digital puppeteering as it is currently understood, and summarize its broad shortcomings based on expert evaluation. Based on this evaluation it will suggest updates and experiments using current perceptual technology and concepts in cognitive processing for existing human computer interaction taxonomy. This updated framework may be more intuitive and suitable in developing extensions to an emerging perceptual user interface for the general public.**

Keywords—digital puppeteering; perceptual user interface; motion capture; cognitive; multi-modal; immersive; social media

I. Introduction

With the inevitability of hardware and software development, complex 3D avatars and games environments are merging with those of learning media, social networks, and consumer interfaces that are traditionally used by the general public. The paradigm of non-perceptual control devices (WIMP) for standard GUI is transitioning to that of perceptual controllers including computer vision, gesture recognition, and voice input for a Perceptual User Interface (PUI) [1]. This transition is in part due to the wider range of tasks and parameters required to be manipulated in complex immersive environments.

While they are a strong first step into multi-modal control, perceptual controllers such as the Kinect have not reached full intended use perhaps due to complexity, effort involved in interaction [2], and accuracy. Its interaction locks the user into a tracking space and expects work in a way that no other game controller does, and non-perceptual controllers such as remotes, keyboards or gamepads are still the preferred interaction method due to a basic level of cognitive ease of use [2].

We can see a parallel for this shift of control systems in the attempt by the digital media industry to translate physical puppeteering concepts to a digital platform.

This paper sets out to review some human-computer interaction lessons from digital puppetry, and to show how these could be applied to a more general purpose user interface for the public. This user interface would rely on perceptual sensors and look to an emerging Immersive Social Network (ISN), such as developed by REVERIE [3], as its final platform. We will finally suggest extensions to a Perceptual User Interface (PUI), [1, 7, 16], and further experiments that could be based on this.

II. Traditional and Digital Puppeteering

To understand the appeal of incorporating traditional puppeteering (TP) techniques as digital puppetry (DP) in production and other mediums, we can look back to the fundamental cognitive concepts of puppeteering through the lens of human computer interaction (HCI).

Traditional puppetry (TP) is naturally embodied and multi-modal. It is not abstracted, though there can be some offsetting of movement by different techniques such as use of rods or strings to control arms or legs [6, 9, 10] Basic categories of traditional puppetry include Glove, Rod, Marionette, and Shadow [4].

The goal of TP is to create the movements of a target character based on the performer's movements [5, 6]. Many concepts based on traditional techniques are utilized in digital key frame animation. Software rigs for digital characters are roughly analogous with the physical puppet and their manipulation: Glove puppetry has a parallel in soft body animation, Rod and Marionette in manipulating end effectors of an inverse kinematics (IK) rig for the digital character, and Shadow in Flash or other 2D animation techniques.

It is important to distinguish digital puppetry from human motion capture as we will be looking at DP concepts to help our interaction, and we need to separate those out.

Human motion capture (HMC), is kinematic mapping that is retargeted, or adapted for a digital character output. This is also called direct-drive or master/slave [7]. It can be used in whole or part (body, hand) and as defined the input is largely kept unchanged apart from error correction, necessary offsetting and calibration to virtual space [6, 7].

Digital Puppeteering (DP) is cognitive mapping that allows human motions to represent other motions or animations in the output to a variable degree of abstraction. This is the core concept of DP to distinguish it from HMC, and allows the segmentation of various motions for a more granular and flexible use of movement. Degrees of freedom, gestures, and discrete or continuous timings are all available for offsetting, coordinate space remapping, and abstraction [7].

Both DP and HMC introduce multimodal inputs and the possibility of live rather than offline production. Digital puppeteering works in the same way as motion capture for input, but the difference is that instead of always being used for direct drive output, the puppeteering input defines a more flexible use range.

With HMC and DP both defined as data-driven animation, there are 3 categories of animation to consider for modern digital production [19]:

A. Data Driven Animation (Offline or Live)

- DP is indirect drive human input, with remapping, retargeting, and trigger output. This can play stored animation, with or without adjustments.
- HMC is direct drive human input, with kinematic remapping and retargeting output.

B. Key Frame (Offline)

- Digital animator working offline manipulating degree of freedom function curves.

C. Procedural (Offline or Live)

- Interpolations/Simulated movement based on poses was and is not fully satisfying.
- Physics simulations enhance characters secondary animation - tails, hair, clothes.

The primary drive of the digital production industry to incorporate DP technology was to cut down the cost and time bottleneck of offline key frame animation (Procedural animation is not sufficiently advanced to assist, as interpolation based on poses does not appear natural).

The main desired result was the migration of large parts of production from offline to live, which has obvious benefits when learning lessons from this to relate to a new PUI. A secondary effect was to import the talents of the traditional artists to the new medium and expand the talent base.

Another result of the movement of production from offline to use of live motion capture data was the movement from non-perceptual controllers (WIMP) used for key framing to perceptual controllers for HMC and DP [1].

Moving away from non-perceptual input used by offline key framers was key in dealing with the complexity of motion input. Original perceptual inputs were based on one and two dimensional devices such as potentiometers and also hardware triggers to make devices based on the Waldo concept by Henson [21]. Later, joysticks with tilt sensors, individual 2dof and 3dof (degree of freedom, or X,Y,Z axes of movement) single sensors, data gloves, and full and part body motion capture systems were widely available as commercial products.

Commercial digital production which has both traditional non-perceptive inputs and off-the-shelf perceptive inputs in rigorous use by interface experts should yield some insights for future PUI [1, 5].

III. Available Control Inputs

It is useful to categorize what is available to the commercial digital production industry, what they have chosen to use and why, and what we can extrapolate from this and other research into suggestions for a PUI [1, 21, 22].

The following table shows typical inputs for the industry, with the perceptual inputs in bold. This is by its nature incomplete, but suggests common usage.

TABLE I. Non-Perceptual and Perceptual control Inputs

device	**Control Input Detail**		
	Device Sensor Type	***Capture (Human)***	***Control (Computer) Remapping, Retargeting***
Keyboard, 2dof mouse	Single point position and orientation, button, key presses	virtual space (expanded with clutching), triggers	Custom interactions, reference node translation in 2d cursor, selection, navigation, scroller
Triggers	button, key presses	analogue and digital trigger values	lip synching, expressions, vertex cluster animation, effects, custom interactions, POV, cursor, selection, navigation
Joystick	Button Presses **Tilt sensor**, Analogue sliders	orientation, axis, slider values, analogue and digital trigger values,	lip synching, expressions, vertex cluster animation, effects, custom interactions, POV, cursor, selection, navigation
Microphone	Audio, Voice	Phonemes, Volume, frequency, location	lip synching, expressions, vertex cluster animation, effects, custom interactions, location

device	Control Input Detail		
	Device Sensor Type	***Capture (Human)***	***Control (Computer) Remapping, Retargeting***
1dof potentiometer/ tape sensor	joint	segmented limb movements, triggers, gestures, end joint positions	full or part body direct drive segmented retarget/remap
3dof inertial/ accelerometer	stand alone or mobile orientation device	Segmented limb movements, triggers, gestures, End joint positions (calculated)	reference node orientation – body/part reference, POV, cursor, selection, navigation
6dof magnetic sensor	single point position and orientation hand, head, root reference	Segmented limb movements, triggers, gestures, End joint positions	reference node translation and orientation – body/part reference, POV, cursor, selection, navigation
Dataglove	full hand motion	Segmented limb movements, triggers, gestures, End joint positions	hand or part hand direct drive segmented retarget/remap
Webcam/ Computer vision	full body motion face tracking, capture	Full or individual limb movements, End joint positions, Root position/orientation	full or part body direct drive segmented retarget/remap
Depth camera	full body motion	Full or individual limb movements, End joint positions, Root position/orientation	full or part body direct drive segmented retarget/remap
FullBody Motion Capture System	full body motion	Full or individual movements, End Joint Positions, Root Position and Orientation	full or part body direct drive segmented retarget/remap

The above table of use is based on personal experience in installing and maintaining custom motion capture and interactive systems based on commercially available hardware and software.

IV. Expert Evaluation

Between 2004 and 2008, as head of motion capture at Inition Ltd., I conducted a series of three exploratory trials with three separate production companies and their puppeteers.

The trials used off-the-shelf hardware and software with interactions tweaked to the performance style of Ragdoll, Henson and Videobaza puppeteers. These performance styles proved similar, as were the hardware and software setups configured for them.

As the expert evaluations for Ragdoll, Redvision and VideoBaza between 2004-2008 were largely similar, I will blend them for the purposes of overall study and reference the Ragdoll configuration primarily.

Working with the puppeteer for Ragdoll, a starting point for the control experiment was set up. Commercially available off-the-shelf products, hardware control inputs, and software were used in the configuration detailed below.

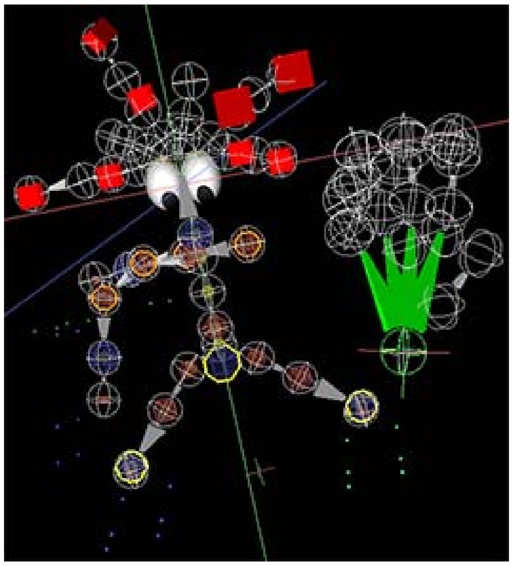

Fig. 1. Multi-modal inputs configured for Ragdoll.

A. Digital Character

The start to any DP is the character. A digital character committed to motion capture or puppetry has some required elements. Basic manipulation is handled by a hierarchy of digital bones, which are merely reference nodes for position and orientation and parented together to simulate a human or other figure. The bones serve to manipulate the mesh envelope, which is the digital character's textured skin and is deformed in relation to bone manipulation. Bones may be translated, and also rotated be either Euler or quaternion methods. They are best further manipulated by a software rig.

Motionbuilder, formerly Kaydara Filmbox, is the DP software platform of choice due to its ability to receive input from all commercial motion capture and control devices. From the beginning, Filmbox was a high end, open, and extensible system. Due to the variety of 3D software formats and hierarchal rigs using either Euler or Quaternion rotations, It was first and still is the only method for conversion of human motion across all 3D software packages and formats.

B. Software, Rigs, Constraints and Relations

Software rigs for digital characters are complex setups of digital bones and constraints in a unique hierarchy called a rig. Filmbox, later Motionbuilder, set out a software template named Character. If a rig had a base collection of bones in the correct hierarchy, Motionbuilder is able to "Characterize" it in a software template. This character template confers on the rig additional features, including the option to create a blended

FK/IK control rig for more procedural manipulation, and advanced interaction with software constraints and relations

Relations and constraints for these rigs can be connected to hardware input devices and also internal software parameters. Motionbuilder has a built in object oriented scripting language to set up and adjust these relations in panes within the constraints module.

Adjustments that make the software ideal for DP setup include isolation of all dof for incoming hardware inputs, vector to number conversion and back, draggable and drop mathematical and logical functions, drag linking and unlinking, value input, triggers, 2D and 3D filters, and Vector manipulation, all functioning in real time.

Motionbuilder’s flexibility in custom setup of a framework of software modules makes it ideal for final DP production work or prototyping for more custom hard coded options.

C. General Setup

There was a joystick of the commercial 13 button variety (Macally Airstick) with tilt sensor and thumb hat in the left or non-dominant hand of the puppeteer. On the right or dominant hand was a data glove. The Cyberglove and 5DT models with 22 or 14 sensors respectively were both tried. Attached to the glove was a 6dof magnetic sensor. Both Ascension flock of birds and Polhemus Fastrak models were trialed. The Puppeteer was seated with a full range of movement for the right hand with 6dof magnetic sensor and in range of a production microphone to PC jack for voice recognition.

D. Joystick Control – Eye Gaze, Animation Triggers

The joystick after setup and testing resulted in the following configuration; Thumb hat for eye gaze control, trigger button to animate eye blink, and 2dof tilt mechanism to control head orientation. Additional triggers would control the head cartoon scaling and some shape animation on the skin mesh, which was damped for ease-in and out. Various other buttons and triggers could all play pre-stored animations such as ‘surprise’ and ‘jump’ in the same way.

E. Data Glove Control – FK/IK Rig

The glove was used in a variety of ways over the shoot to get the best control method. Sometimes it would be used upside down with the index and middle finger mapped to the characters legs, with varying results. At other times they would be assigned per finger or phalange as desired to be retargeted to limb motions in other coordinate spaces with scaling, or to procedural animations for vertex clusters (mesh shapes) to make expressions. Abduction/adduction was used as well to trigger the ‘surprise and ‘jump’. Abstract gestures could have been assigned for triggers instead, but this resulted in a slower cognitive puppeteering process.

F. 6dof Magnetic Sensor Control – Root Reference Pos, Ori

The magnetic tracker controlled the root reference of the character for position and orientation, with appropriate coordinate space swap, and damping and scaling as desired by the puppeteer. Procedural elements on the rig such as hair and tail physical simulations would react to the movement input and generate secondary animation live.

G. Microphone Control – Phoneme Recognition, Lip Synch

The microphone was connected to a software driver module in Motionbuilder that is designed for flexibility in building up a framework of recognized phonemes. These were live mapped in the relations pane to lip synch shapes on the digital character.

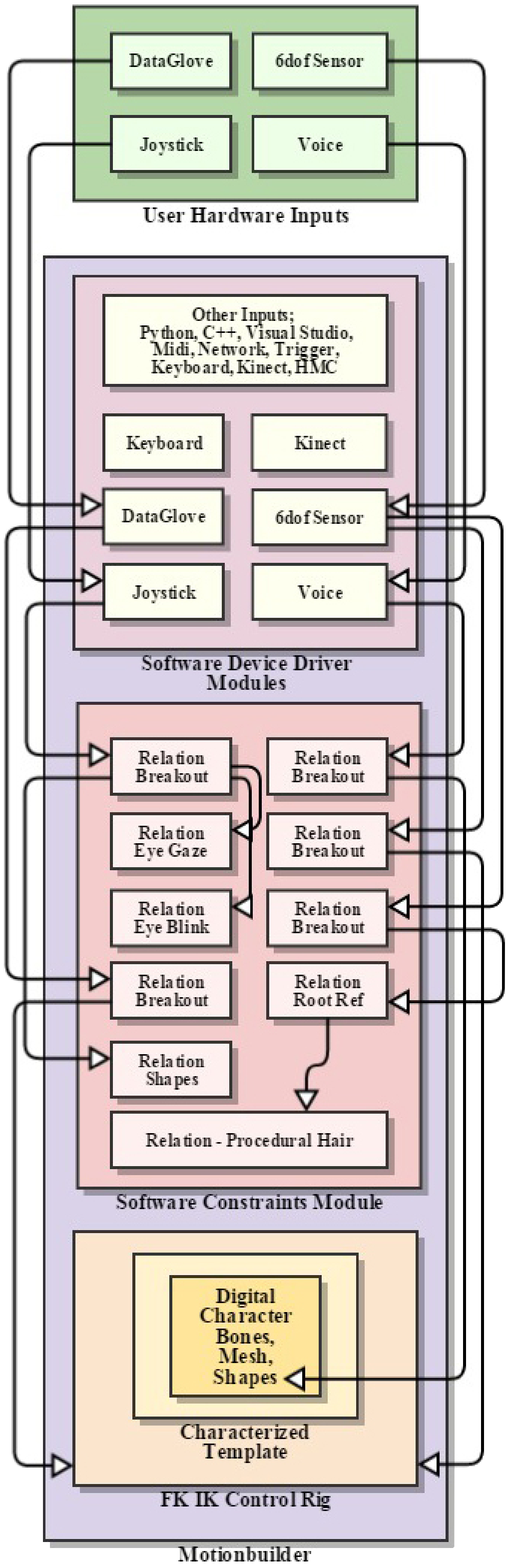

Fig. 2. Multi-modal input flowchart for Ragdoll

V. DISCUSSION

The attempts above were made to get a one person multi-modal system for live operation, but gave largely unsatisfactory results. While the systems were made to work well, and could run concurrently, the following issues were noted in an expert live evaluation: (1) inaccuracy and unrepeatability of data streams, (2) long setup time and constraint of devices, (3) conceptual disconnects, (4) perceptual-motor coordination issues, (5) motoric constraints in operation, and (6) operation fatigue.

While the magnetic sensors offered ideal accuracy and repeatability, the same was not the case for the data gloves. They were awkward for the user to put on, wear and remove. As well, the sensors moved during replacement or capture, making calibration to the digital character difficult, unreliable and hard to reproduce. For all devices, there was too much abstraction in control structure between the control input and task structure, even with more direct mapping. [10, 16] Rather than learning an interaction technique, it would be better to match the task with the input in a conceptually sound manner.

There was cognitive disconnect in the relation of physical movements to DP results [13, 15]. This eroded the concept of self in the DP, and would not translate well to a PUI.

Hand and finger operation to perform these movements also required complex coordination [13, 15], and therefore also served to distance the user from the DP.

In the end, the system configurations were too bespoke and inaccurate, and the performers were reliant on rare talent to rise above the inconveniences of one person DP control.

With further development, a system was settled based upon direct drive from HMC, with augmentations from DP for lip sync, facial expressions, and additional effects in a second pass for offline or live performance and recording. This was used to good effect on the Gadget Show for a live Suzi Perry Avatar broadcast from our office to the show's studio in Birmingham.

This is the system settled on in a more advanced form with the same concepts by Henson Digital Performance studio for its Sid the Science Kid series [21], and VideoBaza for its series of childrens' shows with animator Boris Linnikov. All rely on HMC for direct drive of the main body of the Digital Character and a variety of hardware triggers for expressions driven by a second pass or second skilled operator.

This is less abstracted than the previous full DP expert evaluations discussed which require two operators, and fits the above commercial brief very well. It is an accurate system, but not very adaptable as a concept for a public PUI, as it is still based around tiring HMC performance and abstract triggers.

We are moving on from the complete abstraction of non-perceptual gamepads, mouse, and keyboards, (WIMP) which rely on learned responses by dedicated users and do not translate well to a perceptual structure for generalists in an immersive world [1]. A perceptual structure that marries the input to the task with correct and ideal mapping in both directions is the desired interaction [1, 16].

The Kinect is a prime example of a generalized PUI device, with depth camera full body capture for HMC, hand and finger capture for navigation and interaction. With facial expression recognition, voice recognition and possible DP controls, this is a good starting point for building a PUI framework, as REVERIE has done for elements of their platform.[8]

However, the Kinect still roots the user in place and also can tire them out or lose their interest. This proved to be true in some custom installations [2].

It is therefore the conclusion of this evaluation to move away from a continuous HMC model and also from the hands as main continuous input for UI and interaction. Hands have built in physical motor and coordination limitations [13, 15], and proved difficult to remap effectively for precise control. Precision movement with direct mapping can be tiring, and there is too much disconnect if used for indirect mapping or abstracting to another animation.

Low abstraction is sensible for better cognition. Less disconnect avoids the mismatch of perceptual structures. [16], and ease of use is paramount to general users [17]

From expert evaluation, the sensible fragmentation of very basic dof movements [17] as started for DP gives a good direction to making a lower impact generalized interface that is more accurate with higher cognition. Research [7] also points the way for non-continuous and discrete motions for triggers.

Using lessons learned from this DP evaluation and associated research, this paper will suggest experiments that could lead to additions to a framework for PUI.

VI. CONCLUSIONS

In this paper we presented a combination of commercially available perceptual input devices for digital puppetry. With expert evaluation from traditional puppeteers and motion capture experts, we have considered the viability of a range of multi-modal input to a digital character. From this we suggest extensions to a proposed PUI to be used by the general public for immersive social networks, and future experiments to investigate their suitability.

From the expert evaluations, HMC devices should be replaced with computer vision or depth sensing devices for full and part body capture for the general public. There is no present commercial value in developing wearable hardware HMC devices.

Due to the inaccuracies of most motion capture devices and the fatigue from continuous operation, any full body capture should be of discontinuous gestures and poses for enhancement only [7, 17].

Reducing the number and scale of movements will also work to improve hand input. The key is to not tie them up for navigation or direct drive capture, but for more casual poses for emphasis or input.[7, 17] gesture recognition of thumbs up and down, hand flat, direction pointing and recognition of holding a coffee cup or holding a mobile will enhance interaction in a non-intrusive way.

Eye gaze and voice commands blend very effectively [10,21] with low level hand movements for a more natural interface [22].

Voice capture was good for lip synching, which did not come across as essential. However, it also doubles up for user location, identification, and can reinforce and emotional states being built up for the digital character. Additionally, Face capture is suitable for low impact cognitive triggers. As long as there is no imperative to make expressions, these can be a rich source of context for the digital character.

The granularity lesson of DP gives rise to accessing movements that are simple to capture, reduce the impact of errors, and are more cognitively connected to a perceptual structure. A range of experiments could be done with various body movements to determine ideal perceptual structures for interfaces, and best coordinate space remapping and scaling up. The goal is effortless navigation and communication in ISM. As evidenced by LazyNav [17] for REVERIE, this can be tested and achieved with careful consideration of framework control choices and procedural events.

Finally, having the sensing device give its machine perception of the environment for context would be valuable [1], and real objects could be introduced as physical/virtual elements for a more natural and engaging physical interaction. [18].

VII. FUTURE WORK

An ideal use of DP in reducing the overhead of controlling a digital character in an immersive environment would be a variation on a computer vision version of a direct drive physical marionette control [12]. By utilizing small range finger pointing from the reduced capture region of the users torso, a small capture space could be defined where the fingertips define the effector end points of a digital character's arms.[12,22]

Eye tracking engagement is being used by REVERIE [3], but would be an ideal system if used for highlighting. As a preselection, gaze could highlight or bring up information on objects, reinforcing a strong sense of self in the UI, but leaving actual selection and manipulation only until voice commands or finger gestures are used in conjunction.[21,22]

Disengaging, adjusting, and re-engaging an input device, or clutching [10] is common in 2d input, but would be a good concept to test with a PUI to extend navigational space with low cognitive and motor impact on the user, and could be applied by either the fingertip capture in torso space described above or eye gaze over virtual navigation points set in the virtual environment.

Blue sky or stretch research might include the recognition of users holding or looking at smartphones, and context rich interactions with these devices including eyegaze, access to camera and sensors for virtual POV, or social media information and links.

REFERENCES

[1] M. Turk, "Perceptual user interfaces," workshop available at: http://www.cs.ucsb.edu/~mturk/pubs/TurkEC-NSFWorshop.pdf , 2004, [Accessed 18 August 2015].

[2] M. Aarts, "Using kinect to interact with presentation software," [Accessed 18 August 2015], Available at: http://www.diva-portal.org/smash/get/diva2:639336/ FULLTEXT01.pdf

[3] F. Kuijk, K.C. Apostolakis, P. Daras, B. Ravenet, H. Wei, and D. S. Monaghan. "Autonomous agents and avatars in REVERIE's virtual environment," In proceedings of the 20th International Conference on 3D Web Technology (Web3D '15). New York, NY, 2015, pp. 279-287.

[4] The Ballard Institute and Museum of Puppetry, Available at: http://www.sp.uconn.edu/~wwwsfa/library_puppettypes.htm , [Accessed 18 August 2015].

[5] D. J . Sturman, "Computer puppetry," IEEE Comput. Graph. Appl. 18, 1, 1998, pp. 38–45.

[6] H.J. Shin, J. Lee, S.Y. Shin, M. Gleicher, "Computer puppetry: an importance-based approach," ACM Transactions on Graphics (TOG), vol. 20 no. 2, 2001, pp. 67-94.

[7] D. Sturman, D. Zeltzer, "A design method for whole-hand human-computer interaction," ACM Transactions on Information Systems, vol. 11 no. 3, 1993, pp. 219-238.

[8] REVERIE Deliverable D2.3 "Full technical specification of the REVERIE framework architecture", October 2014, avaliable at: http://www.reveriefp7.eu/resources/deliverables/ , [Accessed 18 August 2015].

[9] L. Leite, "Virtual marionette," IUI '12 In proceedings of the ACM international conference on Intelligent User Interfaces (ACM IUI '12). New York, NY, 2012, pp. 363-366.

[10] I. S. MacKenzie, Input devices and interaction techniques for advanced computing. In W. Barfield, & T. A. Furness III (Eds.), Virtual environments and advanced interface design, pp. 437-470. Oxford, UK: Oxford University Press 1995.

[11] Z. Luo, I. Chen, S.H. Yeo, C. Lin, , T. Li, "Building hand motion-based character animation: the case of puppetry," In proceedings of the IEEE international conference on Cyberworlds (IEEE CW), 2010, pp. 46-52.

[12] A. Mazalek, S. Chandrasekharan, M. Nitsche, T. Welsh, P. Clifton, A. Quitmeyer, F. Peer, F. Kirschner, D. Athreya, "I'm in the game: embodied puppet interface improves avatar control," In proceedings of the ACM international conference on Tangible, Embedded, and Embodied Interaction (ACM TEI '11), 2010, pp. 129-136.

[13] A. Mazalek, S. Chandrasekharan, M. Nitsche, T. Welsh, P. Clifton, A. Quitmeyer, F. Peer, F. Kirschner, "Recognizing self in puppet controlled virtual avatars," In proceedings of the 3rd international conference on Fun and Games, 2010, pp. 66-73.

[14] A. Mazalek, M. Nitsche, "Tangible interfaces for real-time 3D virtual environments," In proceedings of the international conference on Advances in Computer Entertainment (ACE '07), 2007, pp. 155-162.

[15] R. Chua, D.J. Weeks, D. Goodman, "Perceptual-motor interaction: some implications for human-computer interaction," The human-computer interaction handbook, L. Erlbaum Associates Inc. Hillsdale, NJ, USA, 2003, pp. 23-34.

[16] R.J.K. Jacob, L.E. Sibert, "The perceptual structure of multidimensional input device selection," In proceedings of the ACM Human Factors in Computing Systems Conference(ACM CHI '92), 1992, pp. 211-218.

[17] E. Guy, P. Punpongsanon, D. Iwai, K. Sato, T. Boubekeur, "LazyNav: 3D ground navigation with non-critical body parts," In proceedings of the 2015 IEEE Symposium on 3D User Interfaces, 2015, pp. 43-50.

[18] R. Held, A. Gupta, B. Curless, M. Agrawala, "3D puppetry: a kinect-based interface for 3D animation," In proceedings of the ACM

Symposium on User Interface Software and Technology (ACM UIST'12), 2012, pp. 423-434

[19] S.H. Huang, M.T. Chi, T.Y. Li, "Physically-Based Virtual Glove Puppet," In proceedings of Edutainment'11, 2011.

[20] Jim Henson's Creature Shop, "Henson digital performance studio, " Hensons Digital Puppetry Wiki, Available at: http://hdps.wikia.com/wiki/Henson_Digital_Performance_Studio, [Accessed 18 August 2015].

[21] R.A. Bolt, "Put-That-There: voice and gesture at the graphics interface," In proceedings of the ACM SIGGRAPH computer graphics conference, 1980, Vol.14 (3), pp. 262-270.

[22] R.A. Bolt, E. Harranz, "Two-handed gesture in multi-modal natural dialog," In proceedings of the ACM Symposium on User Interface Software and Technology (ACM UIST'92), 1992, pp. 7-14.