Inter-Acting: Understanding interaction with performance-driven puppets using low-cost optical motion capture device

Luís Leite, & Veronica Orvalho

Manuscript
Received:
20,Oct., 2012
Revised:
12,Nov., 2012
Accepted:
20,Dec., 2012
Published:
15,Jan., 2013

Keywords
Digital Puppetry, Virtual Marionette, Performance animation

Abstract— **Puppets can be great storytellers when performed in a dramatic way. They create the illusion of life, making the audience believe in the story. But animating puppets using traditional key frame animation is not a trivial task taking too much time and practice, in particular for the non-expert artists. Digital puppetry presents performance-driven animation, making the puppet reactive to the motion of the performer in real-time. Motion capture methods makes the puppet animation fast and simple, based on the acting of the performer, but they are out of reach for the major consumers. We present a low-cost performance-driven technique that allows real-time puppet animation based on the inter-acting, which can be used in live storytelling. In this paper we study how users can interpret simple actions like walking, using their bodies as puppet controllers. The system was deployed using the Microsoft Kinect and by assuming the marionette aesthetics and constraints, showing how low-cost devices can provide a new mean for motion capture representation. We extend the previous study by presenting another method to interact with puppets using indirect mapping by connecting the puppet to the puppeteer with virtual strings. Last, we performed a pilot experiment animating silhouettes and 3D puppets, to better understand differences in the interaction. An audience had to identify by the output final animation, the actions performed by non-expert artists using their bodies to drive the puppets. We conclude, inter-acting with 2D puppets is similar to the marionette manipulation and needs more interpretation than with 3D puppets.**

1. Introduction

Digital animation provides an efficient environment for storytellers, helping them to bring the stories into real life. However, requires specific technical and artistic skills, which are difficult to achieve, especially when animating a character, thus, taking the production into a complex and time consuming process. Digital puppetry techniques can simplify the animation of puppets. By turning the production into real-time, storytellers with no experience in animation can build animated plays more easily, exploring their imagination by just pulling the "strings". By observing in real-time the feedback of their motion, they can practice the performance and search for the best way to manipulate. In this way, they can achieve manipulation skills to increase the quality of animation. The interaction becomes fluid and intuitive because it´s being driven by the performer, making the story come to life in an easier and spontaneous way compared with key-frame animation. In most cases, just moving the puppets one side to another with some physic features like gravity, can be enough to bring magic into the story. Thus, by exploring their interaction skills users with no animation experience become digital puppeteers, and by exploring their acting skills they become digital actors.

Although digital puppetry techniques, like motion capture, have been used for a long time [1], particular in the entertainment industry. They are inaccessible to the major consumers due to its cost. With the emergence of novel low-cost interfaces like the Microsoft Kinect, digital puppetry has become accessible to everyone. In this study we explore the Microsoft Kinect interface in particular, as a performance animation tool by mapping the user joints to puppet control points in a direct and indirect manner. And find out if users with no experience in animation can manipulate puppets like if they were using traditional marionettes. Furthermore, we compare the manipulation of 2D silhouettes and 3D puppets due to their differences in manipulation and expression. In both cases we use non-deformable meshes to obtain an articulated marionette style based on the users body motion as the “inter” acting controller. Finally, we expand our study going forward into the exploration of non-biped puppets interaction based on a virtual string controller, and into the virtual puppet play composition environment by animating the scenario.

Digital puppetry differs from computer conventional key frame character animation, as it involves performing characters and differs from conventional puppetry techniques because the virtual puppets are not tangible. Despite the differences, the building process of a virtual puppet is some how analogous to the creation of a traditional marionette or a digital character; Building the model, rigging the armature and mapping the control points

This work is partially funded by FCT Fundação para a Ciência e Tecnologia (BD/51799/2011), Instituto Telecomunicações, Universidade do Porto and the projects LIFEisGAME (Funded by FCT) and VERE (funded by EU).
Luís Leite: IT, DEI/FEUP – Instituto de Telecomunicações ✉ luis.grifu@gmail.com
Veronica Orvalho: IT, DCC/FCUP – Instituto de Telcomunicações ✉ veronica.orvalho@dcc.fc.up.pt

to the interface. This interface can be represented with a slider in a 3D software package, a string in a traditional marionette or a tracking point from a motion capture system.

In this article we present, Anim-Actor, a performance-driven full body motion capture interaction technique to manipulate 2D and 3D characters. Anim-Actor allows the control of digital characters in real-time by mapping the movement of the users body to the character. We described how this method was implemented using Microsoft Kinect. We also perform a pilot study that compares the interaction between 2D and 3D puppets using body movement. To evaluate the performance, we challenge an audience to identify actions made by the puppeteers by just observing the result animation. We choose simple actions like lifting weights to turn the performance as clear as possible.

2. Related Work

Digital puppetry is used for a long time in many environments [1]: from film and television to the game industry. The most used methods for performance animation are expensive optical motion capture systems that determinate constraint conditions for capturing the data that can be found in professional capturing studios. But, there are other interesting motion capture interfaces that help the digital puppeteer to control virtual marionettes. Waldo C. Graphic from Jim Henson is an example of how virtual marionettes can be manipulated in real-time for television using mechanical motion capture based on a traditional puppetry articulated arm [1]. Sid the Science Kid is a television cartoon series that shows how Henson studios continues exploring new digital puppetry possibilities by mixing optical motion capture to driven the puppets with other mechanical and digital interfaces for face expressions. The most common approaches to create animations using digital puppetry include: strings to control the puppet with a digital glove [2]; computer vision for tracking color marks in a object that controls the marionette movements [3]; using a multi-touch surface for direct manipulation of bi-dimensional shape puppets [4]; multi-modal interfaces to simulate different marionette techniques in a virtual puppet show [5]; motion capture system for performance-driven with gesture recognition to trigger behavior animation for virtual theater [6]; collaborative multimodal digital puppetry system based on hand tracking, exploring collaboration and interaction to approach the original spirit of puppetry [7]. Important contributions to digital puppetry studies have been made at Protozoa Pictures [8]. Challenged by the market requirements, they research deeply into real-time interactive characters with interesting approaches, like: applying human data to non-human characters; introducing artificial intelligence in characters; developing procedural animation techniques. Reverse Shadow Theatre [9] is an interactive installation developed by the Kitchen Budapest group. Based on their own open-source software, Animata, to map the users motion to a computer silhouette. This project is an example of how to use low-cost cameras with computer vision algorithms to track users motion.

These researches show different approaches for digital puppetry, exploring interfaces that connect puppets to puppeteers. Our proposal is to study the interaction between non-expert artists and 2D and 3D virtual puppets by using optical low-cost motion capture. And to understand if this novel interfaces can help to spread digital puppetry to new environments like in public spaces.

3. System Description

The system was built with the following features: capability for multi-user interaction with one depth camera device and other input interfaces; collaborative interaction using multiple devices in different computers connected to a network; performance-driven puppets in 2D and 3D.

The system captures the users input data (in this case, the body motion) and assigns it to the puppet´s joints, producing the motion-driven animation. It compromises three parts:

(1) Capturing and interpreting users motion producing a skeletal model; (2) Translating and sending the joint coordinates trough the network; (3) mapping the skeleton to the bone structure of the puppet making it move.

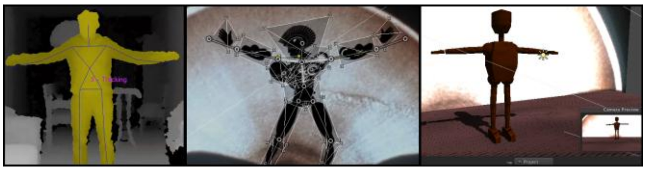

Fig. 1 (left) motion tracking and skeletal model; (center) silhouette figure in Animata; (right) 3D puppet in Unity.

1) In order to capture the performer's movements it was used the Microsoft Kinect depth camera device. This affordable optical interface presents a marker-less tracking with fast calibration process, which is ideal for working in non-prepared environments and with non-expert artists. This sensor is capable of capture the motion of two users in full body at the same time, bringing multi-user interaction into the play. Based on PrimeSense technology this device computes the user skeletal model in real-time providing coordinates of each joint (Figure 1). Then, the open natural interaction (OpenNI) framework provides the application programming interface (API) to access the hardware and interpret the data.

2) For translating and sending joint coordinates it was used OSCeleton from Sensebloom. This application interprets the data from OpenNI framework splitting it into joint coordinates, sending it through network via open sound control (OSC) protocol. This workflow provides multiple inputs at the same time from different machines making a collaborative interaction possible. By connecting through network a chain of computers with input sensors the number of users that can manipulate puppets will increase. OSCeleton also provides options to change the syntax of the messages with the joint coordinates,

increasing the communication compatibility with other applications.

3) For the mapping and animation process two applications were used: Animata for 2D representation and Unity for 3D puppets. Animata is an open source real-time animation editor that provides important features, like: rigging, inverse kinematics and simple physics; deformable and rigid skinning; mapping OSC messages to joints; multi-layer with depth compositing. These features make this multi-platform environment suitable for prototyping digital puppetry in 2D dimensions with 3D layer scenarios. In another hand, Unity is a commercial 3D game engine packaged for multi-platform providing all the important features needed for this project, like reading OSC messages sent from OSCeleton. It is possible to simulate the behaviour of 2D puppet interaction in Unity by using 2D planes with textures and constraining X-axis and Y-axis, but for this experiment in particular we obtained better results in 2D using Animata. It is possible to mix the two engines to get a 2D and 3D representation output at the same time using just one input. Figure 1 shows the framework from capturing the user and computing the skeletal model to the 2D and 3D animation environment.

We extend the previous developments [10] seeking new ways to interact with puppets that presents non-human like skeleton. The puppets were rigged using inverse kinematics (IK) technique, and the control points were attached to the IK end effectors. Puppets like a snake can be manipulated just with one or two hands, mapping the end effectors that are in the extremities (head and tail) to the hands of the puppeteer. Moving the IK end effector all the structure goes along, like in real puppetry. We can also achieve believable results (making the motion more natural and realistic) using physic forces like gravity. Although the process can be more complex to setup – describing the weight/mass for each part of the puppet, the puppeteer can benefit with this feature, working with the motion in a realistic manner. By exploring collisions between rigid bodies, makes the puppet react as if it was real. These are interesting features for the development of a virtual puppet environment.

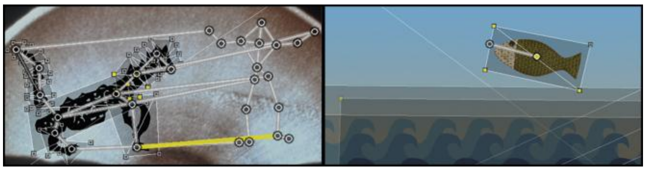

Fig. 2 (left) Virtual strings connecting the puppet to the puppeteer; (right) simple virtual play example using Animata features;

For complex figures we propose the marionette metaphor. Using virtual strings to manipulate the control points of the puppet (Figure 2). In this way, it is possible to map the human skeleton joints to any kind of structure. We also can map one joint from the puppeteer to several control points with different behaviours. To explore this dimension the strings must have physic features, like a network of springs with different tensions between them to achieve diverse reactions. In this first approach of virtual strings, we connect the puppets using a rod-like behaviour connected to the IK control points, which makes the control more direct and precise then with the strings simulating a physics environment. Although in complex figures like the dragon (Figure 2) we found some constrains in motion, as in traditional puppetry the puppets are built for certain particular actions and the puppeteer must search the suitable rigging and mapping solution to each puppet.

Finally we explore some other features from Animata that might contribute to create a virtual puppetry play environment, like: multiple z-depth layers creating the depth of field effect when moving a virtual camera or the automatic bone animation feature making possible to expand and compress a bone in a certain time.

As an example of Animata features implementation in virtual puppetry (right side Figure 2) we create an animated ocean scenario using animated bones with multiple z-depth layer composition. We include a simple rigged puppet to be performed with the hand of the puppeteer and could travel in z-depth.

4. Experiment

We prepared a survey based on a guessing game like Pictionary. Eight Olympic Sports were chosen for the experiment to be interpreted by the "users" and identified by the audience. The sports were weightlifting, boxing, karate, basketball, shooting, arch and hockey later removed from the list because it was not a clear action. The physical space for the experiment was organized to keep observers away from the performers preventing them to be influenced by the acting.

Fig. 3 (left) Kinect capturing the performer; (right) Observers identifying sports.

We made some cards with pictograms identifying the sport to give to the performers in a random order. As a performer the participant was called two times to interpret the sport with the 2D and 3D puppets. As an observer the participant sited about 3 meters away from the projection with a 5 meters wide screen projection, answering the survey, identifying each sport that the actor reproduce as shown in Figure 3. For this experiment we used a non-probability method with a convenience sampling. We chose a focus group with 54 volunteers (19 male and 35 female) with an average age of 19, ranging from 16 to 30. All of the participants were students, non-expert artists, and non-sportsman. None of them had used motion capture systems, and played with Microsoft Kinect before. The total number of answers to the identification survey was 459.

5. Results

We must underline the fact that the positive identifications depended on the acting of each performer: bad acting resulted in poorer identification results.

From our observation the participants when acting, searched a suitable pose to manipulate the puppet. Some participants obtained expressive animation by exaggerating the motion, moving their body members firmly and widely like in puppetry or cartoon animation.

The global comparison results were 72% of positive identifications for the silhouette figure and 84% for the 3D puppet, which was expect. But we found some interesting results when comparing sports, for instance: the weightlifting was better identified with silhouette having 100% of success than the 3D puppet with 85% of identifications (Figure 3), karate also had more success in 2D with 90% then in 3D with just 70%. A possible explanation for this result is the fact that the silhouette presented a simplified body structure with arms and legs more distant from the body, and with few moves the actor could operate the silhouette. But most of the movements had more success in the 3D puppet then in the silhouette puppet.

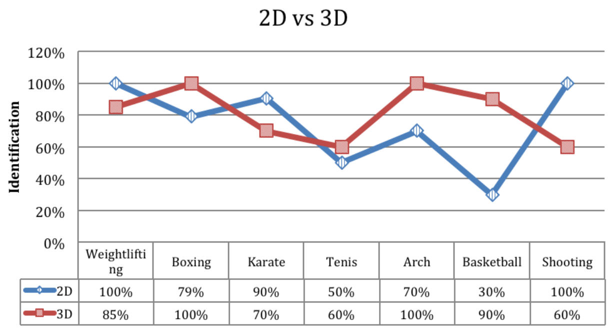

Fig. 4 2D vs. 3D identification graph.

The user instead of trying to interpret the natural position of the silhouette to follow, they searched new ways for interacting with puppet.

This project was presented in two public exhibitions to understand how users interact with puppets using Anim-Actor. For our observation, users took more time to understand how to perform with the silhouettes then with the 3D puppets. The silhouette manipulation required expressive poses, and the users prefer interacting with the 3D puppet, although, they found the silhouettes more dramatic.

6. Conclusions

We believe that the interaction with a two-dimensional puppet with full body motion capture resembles to the control of a real marionette, but instead of using rods, strings or gloves, we use our body as the controller. The manipulation of the virtual puppet can be resembled to the hand shadow performance where the performer searches for the best pose to create the illusion of a shape. In Anim-Actor the performer searches the best body pose to illustrate an action with the virtual puppet. By using our body as a controller the interaction is intuitive but relies on the acting of the performer. In this way, the user becomes part of the play, and not "just" the hidden puppeteer. Rather, the three-dimensional animation of the puppet is presented as a direct representation of body movement like in motion-driven games or classical motion capture animations. This explains the better positive results in the identification, with a total of 84% positive identifications for the 3D puppet against 72% for the silhouette based on the chart from figure 4.

In this exploratory experiment we found that the suggested affordable animation solution was well accepted by the participants, demonstrating that depth cameras are an efficient tool for digital puppetry, even in public spaces. The users were able to interpret the proposed actions without great effort and explored their bodies trying to fit in the puppet skeleton structure. In addition and to our surprise, some of them explored their bodies seeking expressive motion, making us believe the great potential of this interface for performance animation. We didn´t yet evaluate the virtual strings interaction for indirect mapping, but like traditional marionettes we think that this brings a more challenging manipulation.

The affordable motion capture solution that we propose, can be used by non-expert artists as an easy and intuitive way to animate characters for entertainment and artistic purposes, like storytelling.

This study opens new questions about how to interact with puppets using body motion based on digital puppetry that must be deeply researched, in particular: using different rigging methods and inverse kinematics; mapping the human body to different morphological puppets, or mapping multiple control points to just one joint or vice-versa; explore interaction methods with multiple users.

Acknowledgment

This work is partially funded by FCT Fundação para a Ciência e Tecnologia (BD/51799/2011), Instituto de Telecomunicações, Universidade do Porto and the projects LIFEisGAME (Funded by FCT) and VERE (funded by EU). Thanks to Luís Silva, Rui Espirito, Helena Santos and Jacqueline Fernandes.

References

[1] Sturman, D., "Computer puppetry," (1998) In Computer Graphics and Applications, IEEE, vol.18, no.1 , pp. 38--45.
[2] Bar-Lev, A., Bruckstein, A., Elber, G., "Virtual marionettes: a system and paradigm for real-time 3D animation," (2005) The Visual Computer, Vol. 21, pp. 488--501.
[3] Sirota, A., Sheinker, D., Yossef, O., "Controlling a Virtual Marionette using a Web Camera," (2003) Technion - Israel Institute of Technology, Israel.
[4] Takeo, I., Yuki, I., "Implementing As-Rigid-As-Possible Shape Manipulation and Surface Flattening," (2009) Journal of Graphics, GPU, and Game Tools, A.K.Peters, Vol. 14,

Number 1, pp. 17--30.
[5] Leite, L., "Marionetas Virtuais – Animação Interactiva em Tempo Real," (2006) Master Thesis in Multimedia Communication Technologies, Engineering Faculty of the University of Porto (FEUP), Porto, Portugal.
[6] Wu, Q.,Boulanger, P, Kazakevich, M., Taylor, R.,"A real-time performance system for virtual theater," (2010) Proceedings of the 2010 ACM workshop on surreal media and virtual cloning (SMVC '10). New York, USA. pp. 3--8.
[7] Bottoni, P., Faralli, S., Labella, A., Malizia, A., Pierro, M., Ryu, S., "CoPuppet: Collaborative Interaction in Virtual Puppetry," (2008) Communications in Computer and Information Science, Springer, Vol. 7, pp. 326—341.
[8] deGraf, B., Yilmaz, E., "Puppetology: Science or cult?" (1999) Animation World Magazine, Issue 3.11
[9] Reverse Shadow Theatre - Animata http://animata.kibu.hu/
[10] Leite, L., Orvalho,V., "Anim-actor: Understanding Interaction with Digital Puppetry Using Low-cost Motion Capture," (2011) In Proceedings of International Conference on Advances in Computer Entertainment Technology, ACM, pp. 65:1—65:2.

Luis Leite was born on carnival night in 1973 in Porto, and assumes Grifu as alias. Holds a M.Sc. in Multimedia from Faculdade de Engenharia da Universidade do Porto and his an assistant professor at Escola Superior de Música, Artes e Espetáculo (ESMAE) do Instituto Politécnico do Porto (IPP) teaching multimedia and animation classes at Audiovisual (TCAV) and Multimedia (TCM) degree´s.
He is currently a Digital Media Ph.D. student from UT Austin | Portugal, Faculdade de Engenharia da Universidade do Porto, and holds a scholarship from Fundação para a Ciência e Tecnologia (FCT) researching in the digital puppetry field. Member of Porto Interactive Group (PIC) and affiliated with Instituto Telecomunicações (IT / FEUP) as published papers like: "ShapeYour Body", "Virtual Marionette", "AnimActor" or "Marionetas Virtuais". Presented his work in conferences (CHI'12, IUI'12,ACE'11) and in events (Future Places, Imagens do Real Imaginado, Noite Europeia dos Investigadores).
Started as a computer programmer back in 1990 building Bulletin Board Services (BBS) and proceed exploring audiovisual and multimedia; Worked for companies like Animago and Miragem in the fields of animation, programming, post-production and music. Participated in the production and development of TV Series: Ace London (Live Action plus animation - Emmy award selection), Batatoon (children entertainment), Rita Catita (performance animation), Little Stars (children animated series); animated films: "História de um Caramelo" (3D short film) or "Zé dos Pássaros" (2D short film); performing arts: "Prometeu" (shadow puppet play) or "Frágil" (puppet play). The main interests are in interface design (HCI), 2D/3D digital animation, digital puppetry, post-production and computer programming. For more information please visit his Homepage: http://www.grifu.com, videos at Vimeo: https://vimeo.com/grifu/videos, research project at http://www.grifu.com/vm.

Verónica Costa Orvalho born in 1976, mother of a lovely boy. Holds a Ph.D. in Software Development (Computer Graphics) from Universitat Politècnica de Catalunya (2007), where her research centered on "Facial Animation for CG Films and Videogames" . She has been working in IT companies, such as IBM and Ericsson , and Film companies, like Patagonik Film Argentina since 1994. She has given many workshops and has international publications related to game design and facial animation in conferences like SIGGRAPH and Symposium in Computer Animation. She has received international awards for several projects: "Photorealistic facial animation and recognition", "Face Puppet" and "Face In Motion".
Now, she is a full time professor of Porto University, head of the Porto INteractive Center and Computer Graphics group, and co-founder and CTO of Face In Motion (http://www.faceinmotion.com). She is also a former research member at the Event Computational Lab (http://moving-event.org/) working on virtual reality and character animation. Current and past collaborations include several film and game companies (Blur Studios, Electronic Arts, Microsoft Portugal , Dygra Films), and research groups (Stanford University , Universitat Politècnica de Catalunya). Her current research focus on developing new methods related to motion capture, geometric modeling and deformation, and real time animation.