Exploitation of Multiplayer Interaction and Development of Virtual Puppetry Storytelling Using Gesture Control and Stereoscopic Devices

Hui Liang[1], Jian Chang[1*], Shujie Deng[1], Can Chen[2], Ruofeng Tong[3], Jian Jun Zhang[1]

[1]National Centre for Computer Animation, Bournemouth University, Poole, BH12 5BB, UK

[2]Changzhou University, Changzhou, 213164, China

[3]Zhejiang University, Hangzhou, 310058, China

Abstract

With the rapid development of human-computer interaction (HCI) technologies, the new media generation demands novel learning experiences with natural interaction and immersive experience. Considering that digital storytelling is a powerful pedagogical tool for young children, in this paper, we design an immersive storytelling environment which allows multiple players to use naturally interactive hand gestures to manipulate virtual puppetry for assisting narration. A set of multimodal interaction techniques is presented for a hybrid user interface that integrates existing 3D visualization and interaction devices including head-mounted displays (HMDs) and depth motion sensor. In this system, the young players could intuitively use hand gestures to manipulate a virtual puppet to perform a story and interact with props in a virtual stereoscopic environment. We have conducted a user experiment with four young children for pedagogical evaluation, as well as system acceptability and interactivity evaluation by postgraduate students. The results show that our framework has great potential to stimulate learning abilities of young children through collaboration tasks. The stereoscopic HMD outperformed the traditional monoscopic display in a comparison between the two.

*Correspondence

Jian Chang, National Centre for Computer Animation, Bournemouth University, Poole, BH12 5BB, UK.
E-mail: JChang@bournemouth.ac.uk

1 INTRODUCTION

Storytelling is essentially one of the original teaching forms and commonly used as an effective method to teach ethics, values, and cultural norms and differences. As the modern form of traditional storytelling, digital storytelling systems have received increasing attention and demonstrated useful pedagogical functions. In the environment of digital storytelling system, young children can express themselves and cooperate with other players. There has been substantial amount of research undertaken on digital storytelling [1], which mainly focused on the ability of narration. Digital storytelling can also help disabled children with learning difficulties to remove the communication barriers by providing them with the opportunities to play active roles [2], as well as support the development of a wide spectrum of cognitive functions and skills [3].

Virtual Reality (VR) technology is currently used in a broad range of fields, such as games, movies, military simulations and healthcare. The advent of novel human-computer interaction (HCI) technologies, such as Head-mounted displays (HMDs) and motion sensors provide users with improved virtual reality experiences and more natural and intuitive interaction methods. Oculus Rift is an emerging HMDs device specially designed for VR applications. It provides high fidelity and immersive VR experience with wide visual view, high refresh rate, and low-persistence display. Kinect is widely used as a depth motion sensing input which enables users to use gestures for control and interaction without the need for a game controller. Leap Motion is another cost-effective depth motion controller that detects and recognises hands and fingers motions as input; it also requires no direct hand contact or touching.

Considered as one of the powerful and promising education methods, VR technology also enables a wide range of interaction options for digital storytelling and provides a variety of tools and techniques for exploiting and strengthening the adoption of digital storytelling. In the virtual storytelling environment, users get engaged, perform learning tasks and experience activities [4]. ShadowStory [5], for example, is a digital storytelling system inspired by traditional Chinese shadow puppetry. Children use handheld sensors to manipulate shadow puppetry to perform stories on a projection screen.

At present, the young children, also known as the "new media generation" or Digital Natives [6] were born in a rich media society and they started to interact with new technologies from an early age. Saturated with digital media in their daily lives, they need interactive learning environments, and multiple forms of feedback and meaningful learning experiences [7], which bring new challenges to current digital storytelling. The new generation of the digital storytelling system is expected to offer a novel and immersive approach, which is more engaging and interactive, to captivate learners' interests like never before and improve the quality of teaching and learning in the virtual story world.

One of our main concerns is how to provide this "new media generation" an engaging and immersive interaction to accelerate their learning progress by involving novel VR technologies. From academia to industry, a wide range of new VR technologies and novel interaction applications are available, which encourage us to exploit potential novel interactivity in the form of digital

storytelling for enhanced engagement and enjoyment. Various technologies and new hardware can facilitate a synthetic, highly interactive VR environment, for instance, motion sensing technology such as Microsoft's Kinect device.

It is important to captivate young players' interests in narrating with a vigorous representation. Interactive storytelling has long been synergetic with puppetry. Researchers have explored the possibilities in constructing a communicative experience through gestures and voice to provide multi-modal and cooperative interactions with virtual puppets. Studies in storytelling systems with user created and controlled digital characters have proved that they can improve creativity, collaboration and intimacy among children. Therefore, we consider digital puppetry as a proper medium to assist children's narration in terms of emotional stimulation, where virtual puppets act as avatars to perform stories in virtual storytelling environment.

Not only focusing on narration as what previous research did, our research also devotes to increasing children's abilities of cooperation with peers by introducing multiplayers interaction, as well as cognitive development and the ability of motor coordination during their storytelling performance through hand gesture control. Storytelling is also a powerful cognitive tool with great potential to support learners in the process of meaning construction.

In this paper, a prototype of an interactive multiplayer gesture-based digital puppetry storytelling system is developed for young children. Two digital puppets are animated to assist storytelling in VR environment, which makes the system visually and interactively appealing. The depth motion sensing technology is utilised as the gesture-based HCI. By hand tracking and gesture recognition, young children could manipulate virtual puppets to perform story in an intuitive and natural way. Addition to the development of children's narrative skills, we also devote to improving young children's other abilities including cognitive development, motor coordination ability and collaboration ability through their storytelling performance.

Various VR technologies, such as motion sensing technology of gestural-based control, computer animation for content synthesis and cloud-based multi-player interaction, are integrated in the system. In summary, this work has four main contributions:

- Implemented a prototype of the digital storytelling system by developing and integrating the novel VR technologies;
- Introduced a novel multimodal interaction that integrated hand gesture control and stereoscopic display into digital storytelling;
- Designed virtual puppets and their animations for digital storytelling to assist children's narration;
- Introduced the cloud-based network connection for multi-player narration in virtual environment with low latency.
- Conducted a pilot study that compares the differences between the use of a traditional monoscopic display and a stereoscopic HMD.

The remainder of this paper is organised as follows: Section 2 briefly reviews the related research works; Section 3 presents our system design as well as the main technologies involved; result is discussed and analysed in Section 4; Section 5 concludes the paper.

2 RELATED WORKS

There has been a considerable growth of VR technologies over the past decade. VR has been applied in diverse domains and for a variety of purposes, including pedagogical practices, simulation and training, virtual experiments, and game development. One of its key features is interactivity, i.e., the ability to detect user inputs and modify virtual world instantaneously. Interaction interfaces such as head tracking, motion detection and more immersive or intuitive technologies using hand gestures as input have attracted growing interests.

With the advantages of the HCI development, virtual reality technologies are considered to be one of the most powerful tools for supporting learning process [8,9] by providing a synthetic, highly interactive 3D virtual environment and allowing students to experience virtual learning environments in a natural way. With the help of modern VR technologies, digital storytelling emerged as a modern and challenging evolution of traditional storytelling over the last few years for effective teaching and learning. Examples such as *FaTe2* [10] and *MyStoryMaker* [11] have utilised 3D virtual environments in digital storytelling.

Lots of VR interactive technologies are involved in digital storytelling, e.g., the tangible interfaces, mobile-based authoring system, haptic feedback and wireless handheld orientation sensors. *Toontastic* [12] can be considered as a collaborative and constructive digital animation creator that was designed to help children capture and share their stories with other children around the world. *Storytelling Alice* [13] introduced computer programming to learners using the approach of 3D animated stories. *Wayang Authoring* [14] enabled diverse storytelling styles under various cultural backgrounds for the children to create digital stories by using virtual puppets.

Picture This! [15] is a HCI device which embedded a camera and a gesture sensor in children's toys for video composition. It worked as a video storytelling performance system that allowed children to capture stories with physical artefacts at different levels of interaction among different age groups. Children could craft videos about character toys while their gestures and play patterns was analysed by the system. It could benefit young children from developing the spatial and visual coordination while playing with their favourite toys and telling stories.

ShadowStory [5] is a digital storytelling system inspired by traditional Chinese shadow puppetry. It was designed for children to use a Tablet to create digital shadow puppets and perform story cooperatively on a projection screen. To simplify the control mechanism of the shadow puppetry, a pair of handheld wireless 3D orientation sensors (one in each hand) is used to manipulate the character: tilting the first sensor to move the character left, right, up, and down, and tilting the second sensor leftwards or rightwards to let the character bend the body in the corresponding direction, as if bowing or looking up. This kind of interaction seems not intuitive or natural enough because of the direct manipulation on the puppet using handheld controllers.

Leap Motion controller was released in July 2013 as a non-invasive solution using infrared sensing technology for hand tracking and gesture recognition. It is designed to track high fidelity information of fingers' positions and orientations through a small depth sensor and translate hand movements into computer commands. There already exist several applications that have

integrated the Leap Motion devices, such as the *Virtual Piano For Beginners* [16] which allows users to interact with the Leap Motion as a musical instrument. Leap Motion is integrated in our system to offer natural gesture inputs and control the virtual puppets during storytelling.

The Oculus Rift is a virtual reality headset commonly used for video gaming. As a new head-mounted display technology, Oculus Rift features high refresh rate and low-persistence display that work together with its custom optics system to provide incredible visual fidelity and an immersive, wide field of view, which enables the sensation of presence - "the feeling as though you're actually there" [17]. Oculus Rift has opened up new and exciting ways of increasing applications. *ASCENT* [18] is a first person mountain climbing game on Oculus Rift. The VR headset is used for camera control and ice axe aiming. *Zero Latency* [19] sets up a wireless multiplayer shooting game environment, which integrates multiple input devices, such as Oculus Rift, a PlayStation Move controller, and an array of cameras. *Holodeck* [20] is a virtual reality platform using an Oculus Rift headset and the PlayStation Move controllers mounted on a helmet for positional tracking, and Razer Hydra controllers to facilitate interaction.

3 System Design

To create an engaging and interactive storytelling environment for the "new media generation", we needed to resolve two main challenges in our system design: how to provide a novel interaction to young children, and how to design and utilise virtual puppets to assist children's narration. To solve these problems, various technologies are integrated in our system, including the motion sensing technology of gesture-based control, computer animation for content synthesis and cloud solution for multiple players supporting.
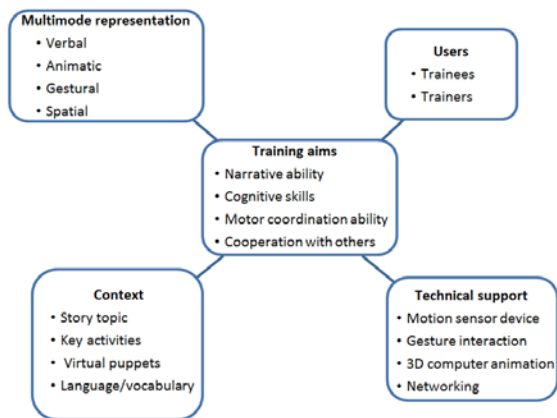
3.1 Conceptual Model

Figure 1: *Components of our conceptual framework*

The components of the conceptual framework of our system are illustrated in Figure 1. "Training aims" is the core component of our conceptual framework. Just like the aims of the majority of current digital storytelling systems, narrative ability training is one of our main considerations in system design. Lots of researchers in the psychology field deal with learning and using technology in educational process, e.g., behaviour analysis [21,

22] and teaching technology in earlier ages [23]. Besides, some researchers have engaged in the development of some other important skills of the young children, such as cognitive ability development, motor coordination training and the ability of cooperation, which might shed insights on the development of children's other abilities in our system. For example, the abilities of constructing mental representation of spatial-temporal environment and spatial visualisation are important to achieve success in science, technology, engineering and mathematics [24]. Also pre-schoolers' visual spatial attention ability predicts the reading skill in the future [25]. With poor motor coordination ability, children tend to underachieve educationally and develop difficulties in building peer relationships [26]. Cooperation is the children's ability to work or act together with others for achieving synergy, as opposed to working in competition.

The puppets actions, such as flying, grasping and talking, can only be successfully performed by specific hand movements and predefined hand gestures. Using hand gestures for controlling the avatar, their motor coordination ability can be trained. In interaction with virtual items that have different properties and roles in the story, the children's space and object recognition capability will also be benefited.

3.2 Storyline

Aesop's story "the Fox and the Crow" is a great fable that teaches children an important lesson in life - about flattery and deception. Considering its popularity among young children and its positive pedagogical meaning, we chose this story to implement and to provide children an exciting and educational interactive storytelling experience with our proposed system. The sketch of the plot outline is presented in Figure 2 which is also provided as a useful hint when children are telling the story and interacting with the virtual puppets.
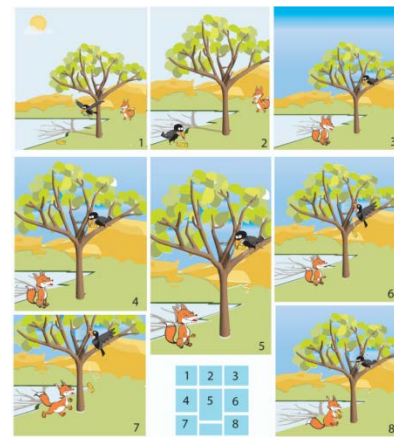

Figure 2: *Sketch of the storyline*

In our interactive storytelling system, players are not required to follow the direct causality pattern. The structure of the entire storytelling can be considered as a graph consisting of multiple story "nodes". The connections between the key "nodes" push the story forward. We listed the key "nodes" of the story "the Fox and the Crow" as shown below:
- Control the puppet fox and the puppet crow to search for food
- Perform how the fox flatters the crow to defraud the cheese

- Help the crow reveal the fox's attempts to trick her
- Describe how the crow finally gives in and drops the cheese
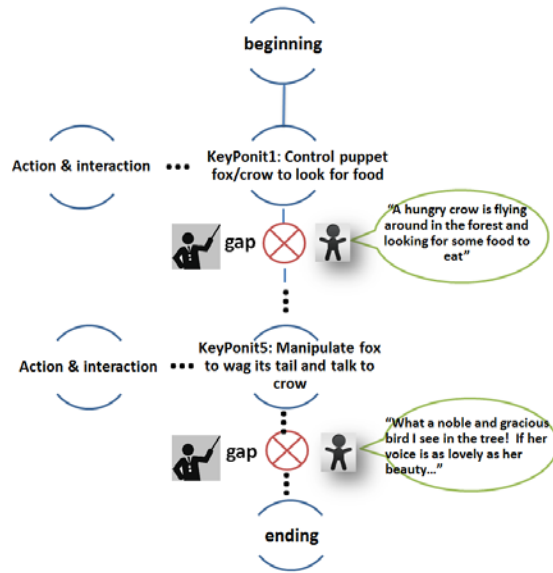- Present how regretful the crow is

Figure 3: *Concept model of the storyline*

The concept model of the system storyline is illustrated in Figure 3. The main storyline consists of multiple key plot points accompanied by players' hand movement and interaction. At the key story points, a tutor would gently interrupt and remind the young players to stop and take their time to narrate at the "gap". The players then could organise their words and describe the key points from their own understanding of the story.

3.3 System Overview

To offer a better interaction and cooperation, our system provides multi-player support which enables two children to play together (with/without supervision) or one on one attention (one tutor and one student) cooperatively. The prototype involves two virtual characters, which provides young children with both the opportunity of manipulating the puppet and the opportunity of playing with peers.

We applied Leap Motion as the depth sensor in this system for its specialties in hand tracking accuracy and gesture recognition advantages. It is more cost effective and space saving than other sensors, such as Kinect. However, multiple Leap Motion Controllers cannot be running on one computer. To solve this problem we added networking and multiplayer support using cloud resource to construct a distributed architecture. The concept of the distributed system architecture is shown in Figure 4. In this system, players interact with their own Leap Motion controller at their client computer locally. The control command is uploaded and exchanged between the local client computers via the remote server, and the whole game scenario is synthesised and displayed on each client's screen. It enables players to participate simultaneously via internet connection just like the experience of Massively Multiplayer Online Game (MMOG).

Technique framework is illustrated in Figure 5. The Photon Cloud [27] is used as the under layer cloud computing infrastructure to provide multiple players supporting. Unity 3D Game Engine is the core component of the system. As the system implementation platform, Unity integrates the Phone Cloud framework with digital puppetry model, audio, animation, user's hand gesture interaction, and Oculus Rift Development Kit to provide technical support for our training aims at the top level.
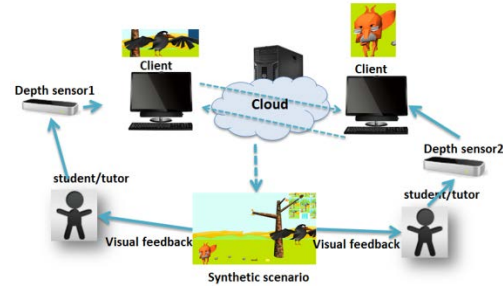

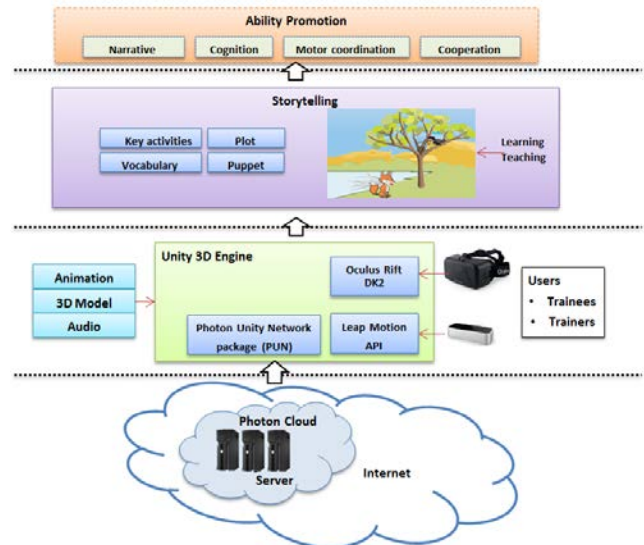**Figure 4:** *Concept model of the distributed architecture*

Figure 5: *An overview of the technique framework*

3.4 Virtual Puppetry Design

Puppetry presents a highly developed form of expressive interaction in our implementation. It depends on motion gestures as the "input" to control a wide range of character's animation. The player and the puppet reflect a relationship of actor and avatar.
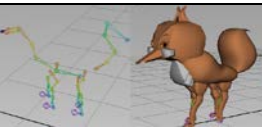
The benefits of creating and manipulating digital characters for storytelling have been explored by previous work [5] that they can improve children's creativity and collaboration skills. The digital puppets are developed to make our system more appealing for young children, due to puppetry's positive educational influence as a lively presentation.

As shown in Table 1, stylised cartoon characters of a cute fox and a doltish crow are designed in Maya 2014 according to children's favour, including their sketches, skeletons and 3D models.

To offer more engagement and interaction to captivate students' learning interests, we designed characteristic actions for the two puppets, including flying, walking and talking, and these actions could be triggered and performed correspondingly to players' hand gestures. To provide real-time response and live

visualisation, these actions were integrated as a set of pre-recorded animation clips. Furthermore, these pre-recorded animations associated with puppets' particular actions can only be triggered by specific events of hand movement and hand gestures. The mapping between animation and hand gestures is illustrated in Table 2 in section 3.5.

Table 1: *Sketch & 3D model of virtual puppetry*

	Fox	Crow
sketch		
3D model		

3.5 Hand Gesture-based Interaction

Many input methods are widely used in HCI, such as keyboard, mouse, touch screen, voice control and hand gestural interaction. We used Leap Motion to capture hand motion. It provides new possibilities for deriving educational benefits of storytelling by creating new modalities of interaction through hand gestures, which can provide more intuitive and natural user experience.

Once the Leap Motion controller has been mounted, a connection between the controller and our system will be automatically established and then the motion tracking data will be passed through. The motion tracking data is presented to our application as a series of snapshots, which is called as the Frames. By programming with the API provided by Leap Motion, the hand positions and orientation as well as the fingertip positions can be accessed.

The motion tracking data accessed from the sensor controller can be represented as a tetrad illustrated as follows:

$$TrackingData = \langle F, HL, PL, GL \rangle \qquad (1)$$

F indicates the frame information provided by the Leap Motion controller. Each frame contains the measured positions and the following information about each entity detected in that snapshot recorded by the controller. HL represents the list of detected hands, PL represents the list of pointables (such as fingers or finger-like tools), and GL indicates the list of recognized gestures. The Leap Motion device recognises certain movement patterns as gestures which indicate user's intent or command.

The tracking device provides model of hands with bone structures of five fingers and each finger is made of four bone joints. We could design specific interactions by using the positions and rotations information of each finger bone provided by the Bone API. The data model of hand motion can be represented as follows:

$$HandMotion = \langle Movement, Gestures \rangle \qquad (2)$$

where $Movement = \langle m_1, m_2, \dots m_i, \dots m_n \rangle$ including different hand movements, such as moving up and down; $Gestures = \langle g_1, g_2, \dots g_i, \dots g_m \rangle$ which may vary with different puppets; each

gesture is composed by a set of joints and their motions.

How to design hand gestures for the purposes of ease of use and intuition is of particular importance for HCI, which involves various considerations, such as the natural language, young players' physical and memorial limitation, recognition accuracy of the input device, and the assumption of the task in storyline. To provide a natural and intuitive way for young children to interact with the virtual toys, we designed our hand gestures vocabulary elaborately. There are in total two sets of hand gestures ("Steering" gestures and "Trigger" gestures) and each set is attached with one hand: "Steering" gestures are assigned for navigation tasks and "Trigger" gestures are allotted for action performance. As illustrated in Figure 6, the Leap Motion controller is placed in front of the player whose hands are held out straight above each side of the device.

Figure 6: *Hands position relative to Leap Motion*

The most straightforward way for the young children to control puppets' movement is mapping the translation of children's hand position in the gesture controller's coordinate system to the puppet's position in virtual environment, which is called the navigation task. This task is designed to be accomplished by a set of "Steering" gestures performed by one hand (the red hand), as illustrated in Table 2(a). In our definition, such hand is in the state of "Steering Mode". When young players try to control virtual puppets' movement in the virtual environment, they just need to move their stretched palms intuitively. The virtual puppet's position in virtual workspace will move accordingly as shown in Equation (3):

$$C_{avatar} = T \cdot S \cdot R_{hand} + C'_{avatar} \qquad (3)$$

where C_{avatar} represents the virtual puppet's position coordinate in current frame, R_{hand} is the player's relative hand movement in real world, i.e. the displacement between the hand positions in current frame and the previous frame, which can be obtained in Equation (1), T is the transformation matrix between the player's workspace and the avatar's coordinate system, S is the scaling matrix, and C'_{avatar} represents avatar's position in the previous frame. For more details of hands movement control mechanism and gesture recognizing, please refer to our previous paper [28].

Except for the navigation, virtual puppets' actions are controlled by another set of predefined "Trigger" gestures performed by another hand (the blue hand), as illustrated in Table 2(b). This hand is defined in the state of "Trigger Mode". The assignment of these "Trigger" gestures to computer commands are designed by following the principles of natural language as well as the young players' limitations of hand postures. Each "Trigger" gesture representing different puppet's action is associated with different pre-recorded animation respectively, such as grasping to talk.

Table 2(a): *Predefined "Steering" hand gestures*

Gestures	Illustration	Description
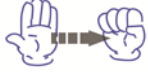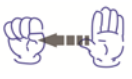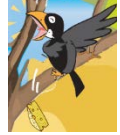
1. Leftward		Fly/run to the left
2. Rightward		Fly/run to the right
3.Upward		Fly/jump up
4.Downward		Fly/jump down
5. Hover/stay		Hover in the air/stay unmoved

Table 2(b): *Predefined "Trigger" hand gestures*

Gestures	Illustration	Description
6. Grasp	Hand open to close	
7. Drop	Hand close to open	
8.Sing/talk	single finger tap	
9.Swipe	rapid multiple finger taps	Waging tail, flattering

Figure 7: *Concept model of state transition*

3.6 Immersive stereoscopic vision experience using HMDs

Head-mounted displays (HMDs) technology enables the players to naturally interact with the virtual world by providing stereoscopic imagery, which has a wide variety of applications, such as entertainment, training and education. The personalized monitoring technology offers an ultra-low latency when tracking the head in a 360° environment and creates stereoscopic vision with excellent sense of the depth.

To present a higher immersive degree in the virtual story world to young players by using Oculus Rift HMD, a stereoscopic vision experience is provided in our system (as shown in Figure 11). Wearing the device on the head, players could enjoy an immersive virtual environment. Unique and parallel images for each eye recreate a natural 3D experience. In this immersive environment, a slightly different image is received by the retina of each eye. These small differences are then processed in player's brain for them to perceive the distance to the objects [29]. Hence, the young players could perceive the place of the objects in a three dimensional space, and get a sense of depth or volume. Although this exciting technology allows the user to feel a higher degree of presence in the virtual environment and a true VR experience, users maybe also experience negative side effects from such immersion, e.g., motion sickness, such as disorientation, nausea, headaches and difficulties with vision [30].

During this VR immersion, a young player has to complete several different motion control tasks by moving the hands and interact with virtual props to follow the key plot points that make the main storyline flow. The scenarios of the key story points are illustrated in Table 4. In this immersive environment, the key points of the story flow are tightly coupled with the movements of different body parts of the player including the eyes, the head and the hands. Interaction with virtual items requires spatial and temporal coordination of the movements of these three parts, which is called "eye-head-hand coordination" [31, 32]. For example, to get the cheese, firstly, the player is required to use hand gestures to manipulate the puppet crow to hover over the grassland and evade the tree at the same time; secondly, the player moves the hand smoothly and locates the crow accurately to the position above the cheese. A good eye-head-hand coordination may be resulted from the creation of temporary synergies or

A state transition model is illustrated in Figure 7 to provide a better explanation of the relationship between the hand gestures and the depth sensor, as well as the possible inner hand motion transitions among different hand gestures. From the left side of the diagram we can see that, in this model, the depth motion sensor exists in three states: "Idle", "Active" and "Tracking". State transition is triggered by the change of input condition, based on the current state, such as losing image data or out of detective range. On the right side, the states of hand gestures are divided into two parts: "steering" and "trigger". Each part consists of several sub-hand-gestures, such as "Fly/run to the left" and "Fly/run to the right" etc. The states of these sub-hand-gestures switch with the change of players' hand motions.

Additionally, after modified for the specific cases, this kind of concept model could also be adapted in other applications which possibly have different set of hand motion or hand gesture transitions.
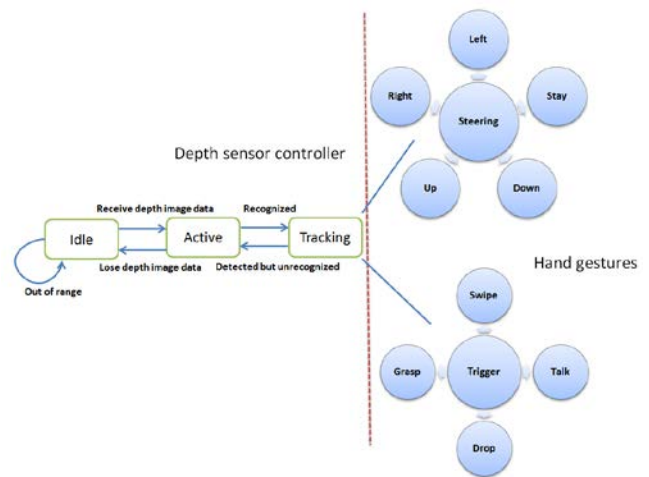
coordinative structures [32]. In another word, HMDs such as Oculus Rift combining with hand gestures can provide a good eye-head-hand coordination and thus a more immersive education environment. Except for narrative skills, young players' abilities such as cognitive skills and motor coordination are also expected to be nourished. The results of the immersive experience of using HMDs in the prototype will be discussed in Section 4.

3.7 Cloud Solution to Support Multi-user

Multiplayer supported storytelling will provide approaches and environments for young players to tell stories collaboratively through their actions and allow multiple means of problem solving. Since there were two characters involved in our demo - the "fox" and the "crow", it needed two players to play cooperatively. It requested the system having the ability to support multiplayers interaction. However unfortunately, the Leap Motion devices do not support multiple sensors connecting to a single computer, which has been a recurring question with no direct answer in Leap Motion development community.

We solved this problem by integrating a Cloud service provided by Photon Unity Networking to realise one gesture controller to one virtual character. This network has the ability to involve many real entities interacting in real time in the same virtual environment. It presents the solution for synchronising the local manipulation of the virtual puppets to other clients in real time. The storytelling facilitated with multiplayer network connections encourages the players to develop individual and collective strategies. The platform where our digital storytelling system ran on with features of multiplayer supporting over the cloud is illustrated in Figure 8.
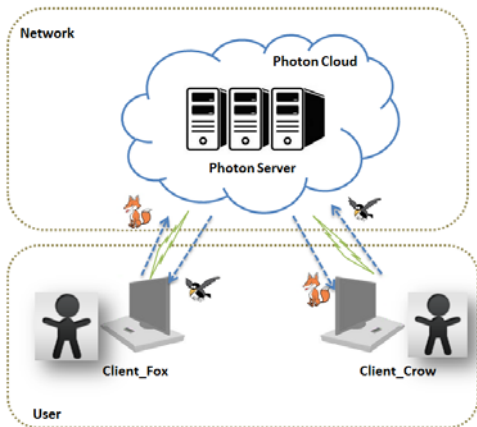

Figure 8: *Multi-user support over the cloud*

For data transmission, firstly, the puppet's information in each frame is serialised, which includes the vectors of the puppet's position, the quaternion of the puppet's rotation, and the animation state as well. For communication and easiness to be parsed, before transmission, this information is compressed and optimised by binary protocol. After encapsulation, the serialised data is transmitted from the client to the Photon Server using User Datagram Protocol (UDP).

Network Quality of Service (QoS) has a significant impact on our

Cloud-based cooperative storytelling system which requires real time interaction, especially with the virtual puppetry. The whole system response delay could be segmented into three components: the *network transmission delay*, the *server's processing time* and the *client's processing time*, as illustrated in Figure 9. The *network transmission delay* is the summation of the time consumed by delivering the local virtual puppet operation data from the client to the Phone Server and the time consumed by sending back the synthesised game information from the server to the client. It is usually referred to the network round-trip time (RTT). The *server's processing time* represents the time consumed by the server to handle the incoming data, which is the time interval between the time that the server receives the incoming information of the virtual puppet operation and the time it responds the client after processing the received information. The *client's processing time* represents the time that the client consumes to decode the feedback data from the server and display the incoming synthesised game information to the client.
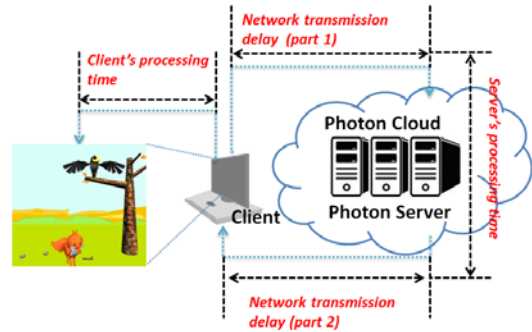

Figure 9: *Latency of the network platform*

To investigate the impact of delay on the cooperative storytelling, a latency test was conducted. Firstly, we measured the actual readings of the latency as the objective observation in different time periods (in the morning and in the afternoon). The latency mostly varied under 250ms. Further, we tried to find out whether or not players could notice this degradation. The result was that players seldom perceived a delay up to 200ms. However, if the delay exceeded 300ms, they reported the system became annoying and unacceptable. Our studies have shown that the cooperative digital storytelling system which actively requires interaction with other players is sensitive to the network delay when exceeding the tolerable threshold. In fact, according to our observation, in most circumstances, the Photon service could transfer data and messages synchronously in real-time. The latency is small compared to the whole loop of distributed game execution.

There exists some studies focusing on the impact of delay on real-time multiplayer games, but few have addressed the related issue in digital storytelling so far. We are the first that studied the influence of the delay on the cooperative digital narration. Although the result was only obtained from our own system and further analysis was expected, we believe that it is applicable to other distributed storytelling system (such as the mobile based storytelling and the web based storytelling) involving similar types of interaction between players.

4 Results & Discussion

4.1 Prototype of the Experiment

The experimental prototype of our system is presented in Figure

10. As shown in Figure 10(a), one young boy was controlling the puppetry fox, and on the right, an adult was manipulating the puppetry crow to assist the narration as a tutor. A scenario performed by two children cooperatively is shown in Figure 10(b). Figure 10(c-e) indicate how engaged the young player was when performing storytelling through this novel interaction. To offer improved immersive experiences, as a piolet study, the head-mounted display is also used as a different display method from the traditional desktop display, as shown in Figure 11.

(a)

(b)

(c)   (d)   (e)

(c) Fox is flattering Crow. (d) Fox is so glad because Crow drops the cheese! (e) Fox is so excited for getting the cheese!!
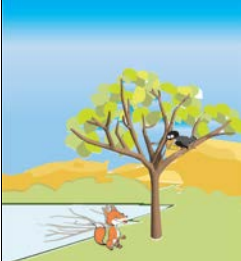
Figure 10: *Experimental prototype*

Figure 11: Storytelling *in Oculus Rift condition*

Each story point involves a set of hand gestures, which can only be successfully performed by specific hand movements to touch off puppets' related actions. Then the pre-defined animations associated with puppets' particular actions can be successfully triggered. Two examples of the scenarios using hand gestures in the virtual storytelling environment are illustrated in Table 3.
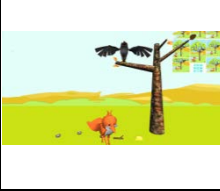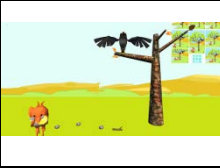
Table 3: *Example of Hand gestures used in two scenarios*

Script	Fox and crow search for food. Crow picks up cheese once he finds it.	
Scenario	**Involved gestures**	
Key point 1->2		Leftward, Rightward, Upward, Downward, Hover, Talk, Grasp
		Leftward, Rightward, Upward, Downward, Step, Talk
Script	Fox flatters the crow trying to defraud crow's cheese. Crow reveals the fox's attempts.	
Scenario	**Involved gestures**	
Key point 5->6		Talk
		Leftward, Rightward, Upward, Downward , Stay, Step, Swipe, Talk

The screenshots of the story cooperatively narrated by two young boys is illustrated in Table 4.

Table 4: *Example of key points performed by young players*

Interactive virtual scene	Narration
	(Aside): One day, a fox and a crow were looking for food in the forest… Fox: *I'm so hungry...* Crow: *I need some food... What's that? ... Woo!!! Cheese!!!*
	(Aside): The crow found a cheese… Fox: *Wait a minute! That's for me!!!...* Crow: *Got it!!!*
	(Aside): The fox tried flattery… Fox: *Good day, Dear Crow...What a noble and gracious bird...* Crow: ... Fox: *How glossy your feathers... How bright your eye...* Crow: ...

	(Aside): The fox tried again and again… Fox: *Let me hear but one song from you that I may greet you as the Queen of Birds...* (Aside): Tricked by the flattery, the crow opens her beak to sing… Crow: *Caw... Caw*
	(Aside): The cheese was dropped to the ground… Fox: *That's all I wanted! Thank you...* Crow: *My cheese...I'm wrong... I'm so foolish... I would not trust flatterers anymore…*

4.2 Pedagogical Evaluation

In this research, a pilot study participated by four 5-7 years old young children (mean 6 years and 4 months old) has been conducted for pedagogical evaluation with permission from their parents who were informed about the nature of the study and its purpose. The participants were carefully selected to cover the age group of Key Stage 1, where the "Key Stage" refers to the levels of the education system in UK, and Key Stage 1 typically represents pupils of age 5-7. Due to difficulty of working with young children, only 4 children participated in the study. However, experiment results of 4 participants can also be valid or have some significance. For example, in Arora's study, initial prototype of a narrative learning device "FunPi" has been tested with four children of age group of 6-8 years [33] and small group of participants can also deliver good results. Five adults volunteered in another test for system evaluation which is elaborated in the section 4.3. All participants were trained to familiarise themselves with the gesture controller after detailed explanation of the experiment. All children were interviewed regarding their subjective assessments when they completed the recorded experiments.

In the experience, the children were divided into two groups to narrate the story cooperatively. Each group carried out five trials of the "Fox and Crow" story. After the learning process, following the provided key story points, all of the children easily got engaged with the linguistic game and built their own story by interacting with the digital puppets with great interests. For the purposes of pedagogical evaluation, we primarily set up some metrics which were recorded as the result of each round of trial. The results are presented in Figure 12, which shows improvement in the following aspects:

Narration ability. As shown in Figure 12(a), the children tended to use more descriptive words for narration across the five rounds. For example, "A crow is looking for some food" in the first round was rephrased to "A hungry crow is flying around in the forest and looking for some food to eat" in the fifth round. Figure 12(b) showed a decrease in duration time followed by an increase. The decrease in completion time benefits from that the players' getting more familiar with the interaction technique. The increase indicated that once the young children got used to the control mechanism and the interaction interface, they tended to pay more attention on the story line and the improvisation, which is vital for storytelling and more suitable for pedagogical purposes as a learning tool for children.

Cognition ability. The decrease of times of average locating tries and wrong gestures performed shown in Figure 12(c, d) suggested

that both the abilities of spatial cognition and visual perception got practiced and improved.
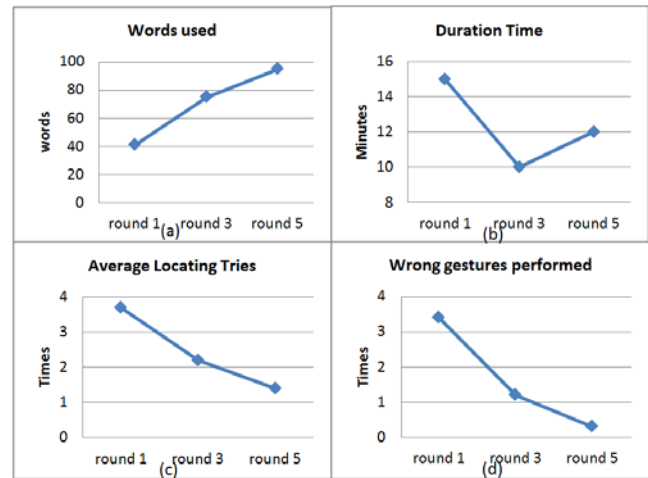

Figure 12: *Average values of metrics in the experiments*

Motor coordination. The decrease of times of average locating tries and wrong gestures also suggested that the children could coordinate the position of their hands with the moving puppet more smoothly and steadily with the predefined gestures.

Collaboration ability. The collaboration in our shared virtual environment is twofold: collaborative narration and collaborative manipulation. The collaborative narration involves dialogues between the two players which develops the storyline. In the meantime, to trigger a specific virtual event, it requires both virtual characters to satisfy certain movements, in terms of positions and orientations. For instance, the fox could only pick up the cheese when the crow was on the tree and had already dropped the cheese. Clearly the required movements are gesture-based control which demands collaborative manipulation between the two players to make sure each character is at the triggering position facing the designed direction.

It was reported by the children that the collaboration increased the task difficulty compared to a single player story, as in the pre-trial practice. The fluency of the storytelling was often interrupted when one of the players made wrong gestures or oriented to a wrong position. However, it was noticed that all the children were more engaged and excited in the collaborative storytelling. Some participants even came up with different tricks to defraud the cheese in each trial. It suggested that the collaboration demanded more attention because the participants needed to react according to their interlocutor's dialogues and actions in addition to the narration and manipulation of their own virtual characters.

4.3 System Evaluation

Considering young children's knowledge limitation on the technologies involved in the system, we asked postgraduate students to take part in the system evaluation instead. Five 22-28 years old postgraduate students who were naive of this experiment participated in the system evaluation test. Each student did six trials with the experimenter, in three of which they acted as the "fox" and in the other three they acted as the "crow". A questionnaire was given after the trials, regarding acceptability, interactivity as well as their suggestions for improvement. The

Likert scales were used to evaluate acceptability and interactivity. An open-ended question was given for suggestions to improve our experiment design.

4.3.1 Acceptability

There are seven criterions given for evaluating the acceptability, ease of use, precision, stability, naturalness, handiness, freedom of movement and effectiveness. The participants are asked to rate each criterion from 1 to 5, indicating "Poor satisfaction" to "Excellently satisfied". Figure 13 shows the results of the mean score of each criterion about the acceptability of our system. The ease of use, naturalness, freedom of movement, handiness and stability has returned positive feedback with score around four, especially the ease of use and naturalness have achieved the highest satisfaction. This is possibly due to that we used the elaborated hand gestures design as the method of natural interaction. However, the effectiveness and precision received the lowest satisfaction which could be induced by the recognition accuracy limitation of the depth sensor. For example, we have noticed that folding the fingers over the hand or overlapping the hands could lead to false detection. The hand position relative to where the sensor is also crucial. The perfect position of the tracked hand is on the direction that is perpendicular to the top surface of the device. Any other position may affect the accuracy of hand tracking.
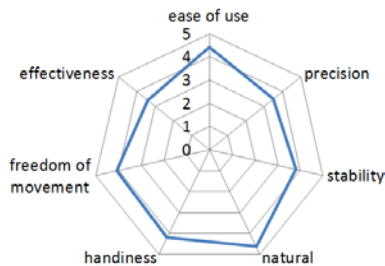

Figure 13: *Mean scores for each criterion of acceptability (1-Poor, 2-Fair, 3-Average, 4-Good, 5-Excellent)*

4.3.2 Interactivity

The evaluation of interactivity was rated in terms of Steering gestures and Trigger gestures with the following criterions: effectiveness, intuitiveness, simplicity, precision and quickness. Similar with the evaluation of the acceptability, participants rated each criterion scaled from 1 (Poor satisfaction) to 5 (Excellently satisfied) whose mean scores are shown in Figure 14.

Overall the Steering gestures returned better user experience than the Trigger gestures. It is probably due to the ease and intuitiveness of the directional navigation design and participants could understand the interaction approach without learning. The Steering has achieved almost perfect scores in effectiveness, intuitiveness and simplicity. Weaker but still satisfying scores were obtained in quickness and precision. Three participants noted that the steering is too sensitive and not very accurate over short distance and four participants mentioned excessive speed. This indicated the sensitivity and moving speed of the virtual characters needed to be tailored depending on users' habit or expectation. The relatively lower rates of the Trigger gestures revealed weak effectiveness and precision, as well as slightly better quickness, intuitiveness and simplicity. To improve the

Trigger gestures effectiveness, it requires most simplified gesture design for the ease of recognition by the device. This is a dilemma that complex gestures can enrich the control of virtual characters as well as the story, but it will also cause difficulty of recognition by the device. "Steering" gesture is probably considered more acceptable than the other hand gesture mode because it is the easiest and the most intuitive action for young children without any other complex manipulation.
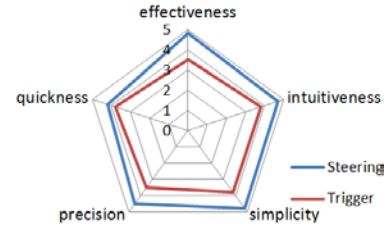

Figure 14: *Average scores for each criterion of the two modes of hand gestures (1-Poor, 2-Fair, 3-Average, 4-Good, 5-Excellent)*

4.3.3 Improvement Suggestions

Several participants mentioned the issue of arm fatigue when raising hands in the mid-air without support. This has been a common issue for all mid-air gesture-based applications especially when long periods of operation involved. There is yet no optimal solutions reported but it can be improved by gesture design and even assistance from multimodality. For instance, Song et al. [34] designed a CAD modelling approach combining hand gestures and gaze gestures in zooming tasks.

Although the interaction were reported intuitive, some participants still suggested extending the learning process or improving the hand gestures design, because current experiment settings can cause misoperation and accuracy issues. An ideal intuitive interaction approach is expected with no learning process or only mild practice. However, it is difficult to achieve because different users have different habits, it will require our system to be adaptive to associate users' intention and their gestures. A richer gesture database is necessary for the Trigger mode and we need to train the system to learn the users' habits for this purpose.

4.4 Comparison between monoscopic display and stereoscopic HMD

Oculus Rift has attracted a number of research interests, such as a VR cycling exergame [35], a human based motion interactive platform [36] and a simulator that provides flying experiences through a bird's perspective [37]. However, to the best of our knowledge, efforts to date are still primarily focused on exploring its technical possibilities. In particular, we are interested to examine the user experiences of the VR HMD in storytelling, comparing to that of the monoscopic display.

4.4.1 Method

There are some common characteristics of immersion, such as cognitive involvement, real world dissociation, emotional involvement, challenge and control, which were observed by previous research [38-40]. For the purpose of immersion study, based on these characteristics a well-validated Immersive Experience Questionnaire (IEQ) was created [41] and utilized to conduct user studies [42-43]. Considering that our system is a

kind of educational storytelling system instead of a recreational video game, as well as that the targeting players are young children, part of the questions in IEQ is carefully selected in this user study. These immersion questionnaires are used in the experiment to study players' personal experience of the game by rating to what extent they would agree with the statements, e.g. "I felt detached from the outside world.", "At the time the game was my only concern.", "I still felt as if I was in the real world whilst playing.", "To me it felt like only a very short amount of time had passed".

In our study, we compare user experiences under two conditions: performing the same story "the fox and the crow" using HMD versus the same tasks performed using traditional interfaces (on a desktop computer with monoscopic display). We adopted a mixed approach to explore players' experiences during storytelling. This involved collecting and analysing both qualitative and quantitative data of 4 participants under the aforementioned two experimental conditions.

4.4.2 More immersive experiences on the HMD

The participants reported more immersive experiences on the Oculus Rift than the monoscopic monitor, the traditional desktop monitor, when asked about their overall training experience after completing all the story narration sessions on the two different platforms. It appears that the stereoscopic vision experience of narrating using the Oculus HMD provided a lot more vivid and immersive feeling than using the traditional computer monitor. "It is more interesting on the Oculus than the monitor" reported by one of the young participants. The player felt more involved when the cheese was found in front of his eyes. Especially, in the scene of the fox flattering the crow by calling it a beauty, a pair of participants playing the role of the crow and the fox respectively turned their head and performed the dialog face to face, just like "I felt I was a crow on the tree and facing a real existing cunning fox". This experience using the HMD greatly helped enhance the immersion of the story plot.

The scores of the IEQ are shown in Figure 15. As shown on the IEQ scale, the significant difference of 14 points between the mean of the two display methods indicates that the participants are more immersed in the Oculus Rift condition than using the traditional desktop monitor.

Based on our investigation, one possible reason that caused better immersion with the Oculus Rift may be due to the degree of concentration. When wearing stereoscopic 3D HMDs, player's entire peripheral vision is covered by the headset, which actually strengthens the immersive visual effects and allows the player feels surrounded by the virtual world. This made the participants concentrate on the interaction in front of his eyes and encouraged a "real-world dissociation", which might explain the generally improvement of immersion in the Oculus Rift condition.

Wearing the Rift, the way to interact with the virtual environment is significantly changed, such as navigating and locating avatars. The Oculus Rift Kit uses a tracker for tracking the position of the headset in 3D space. That means the player is required to use head tracking as input and look around to find the avatar or virtual props in the virtual environment. At the beginning of the experiments, most young players reported that they needed to get used to this head locating mechanism, e.g., "It took me quite a long time to get used to the head movement and the hand gesture

at the same time". But once the player learnt the interaction techniques he showed great interest in it. He claimed "It felt more natural, finding cheese was fun just like being in right there."
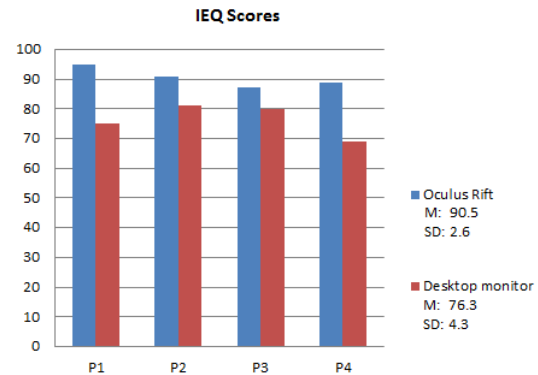

Figure 15: *Scores for participants in Oculus Rift and monoscopic display conditions*

During the investigation of players' immersive experiences, we also found that it was a common opinion that narration in the environment using Rift was more interesting and attractive. For example, one participant reported despite the novel hand gesture interaction, looking at the PC monitor was a little boring, "just a normal game". While wearing the head mounted device, the player was more engaged: "how amazing! It's totally a different world!" The players intended to spend more time on the storytelling process with more fun.

4.4.3 Passive effects of cybersickness with gameplay

Interviews after the trials show that beside the improvement of immersion and presence, users also can experience unwanted side effects. One of the young participants reported that he started to feel "a little bit sick" after the story went on and one boy felt "a bit of nauseous" when he wore the device for the first time. One of the authoritative explanations comes from the sensory conflict theory that the discomfort results from the incoherence between the visual input and the vestibular system [44]. A common effect due to the cybersickness was the decrease of players' performance when using the HMDs [40], e.g., users were slower in simple navigation and object interaction tasks wearing HMDs [45].

However, related studies were typically carried out in a competition game environment, such as in the first-person shooter (FPS) games, The Half-Life 2, which intensively involves time limitation and gun shooting [40]. Since we are in an interesting storytelling system, not like other studies, the task is relatively relaxed. Only one boy paused for a rest temporarily. Despite feeling sick quite early on, he could quickly adapt and still enjoy the immersive storytelling. In fact, cybersickness hardly affected immersion.

4.4.4 Insights for system design using HMDs

Based on our findings, despite the side effects of the cybersickness, all the participants reported enhanced experiences, a richer engagement and a deeper immersion in the HMD condition than on a desktop setup. Our findings provide some preliminary insights into the utilization of HMDs in digital

storytelling system, especially on the system design phase. It can also contribute to the application of VR Education.

1) Try to reduce the effect of cybersickness

At the same time of expecting new devices that could provide a better immersive experience and reduce the passive effects, cybersickness should be taken into account in the design process and the system designers will need to alleviate the passive effect. For example, the entire virtual game/story is better to be designed to consist of several short/simple key plot points as we discussed in our system design. And at some points (e.g., the key plot points), players are allowed to stop playing once they feel sick and then resume without affecting the performance. That means there should not be too much time restrictions, especially when the targeting users are young children.

2) Design different interaction types for different target users

System designers should be well aware of the limitations of the cognitive competence of different user levels. Both the HMDs and depth motion sensor based interactions involve the abilities of motor coordination, eye-head-hand coordination and spatial cognition etc. Some of them seem easy to achieve for adults while being difficult for the young children.

5 Conclusions

In this paper, we have developed a prototype of a novel multiplayer interactive digital storytelling system assisted with virtual puppetry to provide natural interaction/control and immersive experience when narrating stories. The system integrates useful VR techniques and novel hardware to fertilise development of young children's abilities, including narrative skills, cognitive skills, motor coordination and collaboration through storytelling. The development of the prototype demonstrates the possibility of combining novel HCI and digital storytelling to inspire the children's teaching and learning practices.

The prototype provides young players with cross-dimensional interactions, and multimodal technologies are involved: the depth motion sensing technology, HMDs, computer animation and cloud-based network solution. The use of multiple Leap Motion sensors is creatively implemented, where the multiple sensors' data are exchanged and synchronised through the Photon Cloud. Our method can wildly extend the potential usage of the devices and open up many new possibilities with the cloud-based network support. As a pilot study, the head-mounted display is validated to be able to offer improved virtual reality experiences. Despite minor cybersickness, participants experienced richer engagements and deeper immersions on the Oculus Rift than on a monoscopic display. Through our exploratory study of using the Oculus Rift in digital storytelling, our research would shed light on the development and applications of HMDs in VR Education.

The pedagogical evaluation of the system is preliminarily examined in our experiment and the results from the analysis are promising that the young children can be benefited from the using the system. During playing, some minor frustrations (or difficulties) are observed, such as dropping the cheese by mistake, having problem to locate the avatar, or using a second hand to support the main hand in operation etc. Getting over these difficulties also made the children feel rewarded and find the game interesting. Aside from the pedagogical effects, we also evaluated the system based on adult participants' feedback, which provided complementary information.

We can further improve the development and related research in the future. More diverse hand gestures, such as thumb lift, different swipes: one, two and more fingers in four independent directions (up, down, left, right), and different circles: one, two and more fingers in two independent directions (left, right), could also be investigated to provide more complex interaction.

There are limitations with our evaluation considering the limited data pool. Only a limited number of trials have been tested in our experiment, and also, the number of participants was limited by the difficulty of working with young children, because close supervision and support of parents were necessary. Therefore, a control group was missing in the experiment which may possibly cause ambiguity. We plan to exploit our system to conduct indicative experiments with more participants to identify the pros and cons with quantitative data and analysis. Both independent and dependent variables will be measured or observed to provide a clearer understanding. Also, a control group will be involved in the future serving for comparison evaluations. To improve the system evaluation, experiment should also involve young children and other age groups.

References

1. Psomos P, Kordaki M. Pedagogical analysis of educational digital storytelling environments of the last five years. *Procedia-Social and Behavioural Sciences* 2012; **46**: 1213-1218.

2. More C. Digital stories targeting social skills for children with disabilities multidimensional learning. *Intervention in School and Clinic* 2008; **43(3)**: 168-177.

3. Garzotto F, Paolini P, Sabiescu A. Interactive storytelling for children. In *Proceedings of the 9th International Conference on Interaction Design and Children,* ACM 2010; 356-359.

4. Hofer M, Owings SWAN K. Digital storytelling: Moving from promise to practice. In *Society for Information Technology & Teacher Education International Conference*, **2006(1)**: 679-684.

5. Lu F, Tian F, Jiang Y, Cao X, Luo W, Li G, Zhang X, Dai G, Wang H. ShadowStory: creative and collaborative digital storytelling inspired by cultural heritage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM 2011; 1919-1928.

6. Thompson P. The digital natives as learners: Technology use patterns and approaches to learning. *Computers & Education* 2013; **65**: 12-33.

7. Barnes K, Marateo RC, Ferris SP. Teaching and learning with the net generation. *Innovate: Journal of Online Education* 2007; **3(4)**: 1.

8. Jonassen DH. *Computers as mind tools for schools: Engaging critical thinking*, Prentice Hall, 2000.

9. Smeets E. Does ICT contribute to powerful learning environments in primary education? *Computers & Education* 2005; **44(3)**: 343-355.

10. Garzotto F, Forfori M. FaTe2: storytelling edutainment experiences in 2D and 3D collaborative spaces. In *Proceedings of the 2006 conference on Interaction design and children*, ACM 2006; 113-116.

11. Mckinley B, Lee YL. Mystorymaker. In *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, ACM 2008; 3219-3224.

12. Russell A. ToonTastic: a global storytelling network for kids, by kids. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, ACM 2010; 271-274.

13. Kelleher C. Motivating Programming: Using storytelling to make computer programming attractive to middle school girls. No. CMU-CS-06-171. CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 2006.

14. Widjajanto WA, Lund M, Schelhowe H. Wayang Authoring: a web-based authoring tool for visual storytelling for children. In *Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, ACM 2008; 464-467.

15. Vaucelle C, Ishii H. Picture this!: film assembly using toy gestures. In *Proceedings of the 10th international conference on Ubiquitous computing*, ACM 2008; 350-359.

16. https://apps.leapmotion.com/apps/virtual-piano-for-beginners/windows. Access date: 19, Jan, 2016.

17. https://www.oculus.com Access date: 19, Jan, 2016.

18. Dufour T, Pellarrey V, Chagnon P, Majdoubi A, Torregrossa T, Nachbaur V, Li C, Ibarra Cortes R, Clermont J, Dumas, F. ASCENT: a first person mountain climbing game on the oculus rift. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, ACM 2014; 335-338.

19. http://zerolatencymedia.com Access date: 19, Jan, 2016.

20. http://www.projectholodeck.com Access date: 1, Nov, 2015.

21. Greer, R. Douglas. Designing teaching strategies: An applied behavior analysis systems approach. Academic Press, 2002.

22. Vargas EA, Vargas JS. Programmed instruction: What it is and how to do it. *Journal of Behavioral Education* 1991; **1(2):** 235-251.

23. Skinner BF. The technology of teaching, *New York: Appleton-Century-Crofts*, 1968.

24. Wai J, Lubinski D, Benbow CP. Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology* 2009; **101(4)**: 817.

25. Franceschini S, Gori S, Ruffino M, Pedrolli K, Facoetti A. A causal link between visual spatial attention and reading acquisition. *Current Biology* 2012; **22(9)**: 814-819.

26. Skinner RA, Piek JP. Psychosocial implications of poor motor coordination in children and adolescents. *Human movement science* 2001; **20(1)**: 73-94.

27. https://www.photonengine.com. Access date: 19-01-2016

28. Liang H, Chang J, Kazmi IK, Zhang JJ, Jiao P. Puppet Narrator: utilizing motion sensing technology in storytelling for young children. In *Games and Virtual Worlds for Serious Applications (VS-Games), 2015 7th International Conference on*, IEEE 2015; 1-8.

29. Juanes JA, Gómez JJ, Peguero PD, Lagándara JG, Ruisoto P. Analysis of the oculus rift device as a technological resource in medical training through clinical practice. In *Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality*, ACM 2015; 19-23.

30. Davis S, Nesbitt K, Nalivaiko E. A Systematic Review of Cybersickness. In *Proceedings of 2014 Conference on Interactive Entertainment - IE2014*, ACM Press 2014; 1-9.

31. Yu C, Ballard DH. Understanding human behaviors based on eye-head-hand coordination. In *Biologically Motivated Computer Vision,* Springer Berlin Heidelberg 2002; 611-619.

32. Pelz J, Hayhoe M, Loeber R. The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research* 2001; **139(3)**: 266-277.

33. Arora, N., Agarwal, N., Reddy, S.R.N. FunPi: An Interactive Learning Experience Using Story Narration. In *Proceedings of the Sixth International Conference on Computer and Communication Technology 2015,* ACM 2015; 398-402.

34. Song J, Cho S, Baek SY, Lee K, Bang H. GaFinC: Gaze and Finger Control interface for 3D model manipulation in CAD application. *Computer-Aided Design* 2014; **46:** 239-245.

35. Bolton J, Lambert M, Lirette D, Unsworth B. PaperDude: a virtual reality cycling exergame. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, ACM 2014; 475-478.

36. Cheng LP, Lühne P, Lopes P, Sterz C, Baudisch P. Haptic turk: a motion platform based on people. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, ACM 2014; 3463-3472.

37. Rheiner M. Birdly an Attempt to Fly. In ACM SIGGRAPH 2014 Emerging Technologies, ACM 2014; 3.

38. Sweetser P, Wyeth P. GameFlow: a Model for Evaluating Player Enjoyment in Games. *Computers in Entertainment (CIE)* 2005; **3(3)**: 3-3.

39. Ijsselsteijn W, Kort YD, Bellotti F. Characterising and Measuring User Experiences in Digital Games. In *International conference on advances in computer entertainment technology* 2007; **2**: 27.

40. Tan CT, Leong TW, Shen S, Dubravs C, Si C. Exploring Gameplay Experiences on the Oculus Rift. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, ACM 2015; 253-263.

41. Jennett C, Cox AL, Cairns P, Dhoparee S, Epps A, Tijs T, Walton A. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 2008; **66(9)**: 641-661.

42. Cox A, Cairns P, Shah P, Carroll M. Not Doing But Thinking: The Role of Challenge in the Gaming Experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM 2012; 79-88.

43. Cairns P, Li J, Wang W, Nordin AI. The Influence of Controllers on Immersion in Mobile Games. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM 2014; 371-380.

44. Llorach G, Evans A, Blat J. Simulator sickness and presence using HMDs: comparing use of a game controller and a position estimation system. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, ACM 2014; 137-140.

45. Santos BS, Dias P, Pimentel A, Baggerman JW, Ferreira C, Silva S, Madeira J. Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimedia Tools and Applications* 2009; **41(1)**: 161-181.