





Children's Intuitive Gestures in **VISION-BASED ACTION GAMES**

Novel computer vision-based game technologies aim to give players more immersive and physically challenging gaming experiences.

Video and computer games play an integral part in the lives of many children. However, some studies suggest that extended computer use may have

By **JOHANNA HÖYSNIEMI,**
PERTTU HÄMÄLÄINEN,
LAURA TURKKI, and
TEPPO ROUVI

negative effects on a child's physical development [3]. One factor causing these effects is the use of traditional human-computer

interaction styles and input devices, such as a keyboard, mouse, or game pad that promote a sedentary lifestyle.

Our work is part of a wider project that aims to provide an immersive and physically engaging alternative to traditional computer games by making use of computer vision and hearing technology. Compared



to other games based on sensory devices, such as dance mats or skateboard controllers, our perceptive user interface is wireless and does not require any contact with the input devices during game play.

Due to the novelty of vision-based game interfaces, there are few established playing conventions, for example, for controlling a flying or swimming game character. This calls for careful consideration of the playability of game controls as well as the evaluations of the controls with the target group. The key requirements for computer vision-based games are robustness, responsiveness, intuitiveness, and physical appropriateness; particularly the last two since they play a central role in providing a pleasant playing experience and in making the learning phase shorter—an important factor in children’s interactive products [5, 6]. Current research, however, does not provide enough data on what gestures children find natural and intuitive in different game contexts.

In this article we discuss how children have helped us define and develop the control gestures for our game, *QuiQui’s Giant Bounce* (www.webcamgames.com), through traditional usability tests and Wizard of Oz (WOz) prototyping sessions. We focus on presenting the movement styles children preferred for the

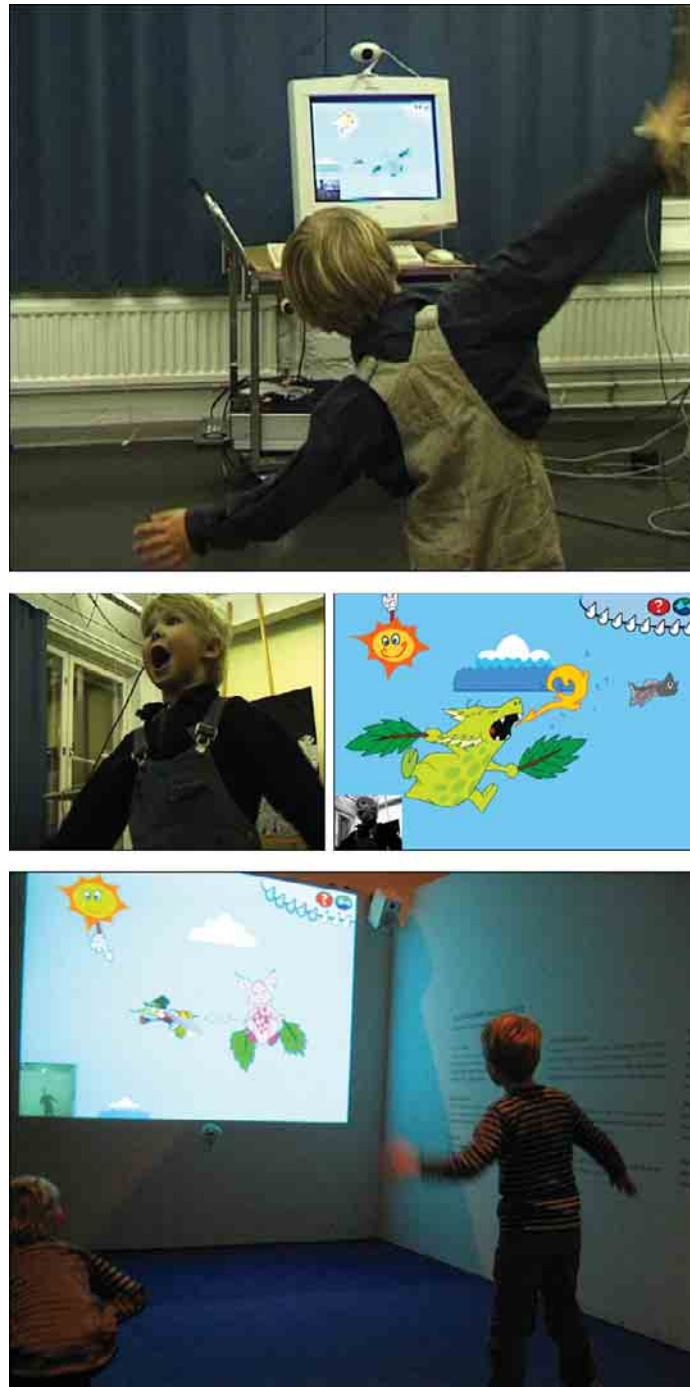


Figure 1. A child flaps his hands to make QuiQui fly (top). When the player shouts, QuiQui breathes fire (middle). The game has also been on display at various exhibitions (bottom).

different game contexts and how these findings have affected the game design.

The game is aimed at children ages 4–9 and uses simple sets of movements that vary according to the game task. It closely resembles KidsRoom [2] and other physically interactive story environments. However, our game has three unique characteristics: it uses a 2D computer-animated dragon as an avatar that mimics the player’s movements instead of a video avatar (video image of the user with or without background subtraction) as seen in many systems [5, 10, 11]. The user interface also adds voice as a second input modality: when the user shouts, QuiQui the dragon breathes fire to scare mean game characters away (as illustrated in Figure 1). And, in contrast to commercial games using custom-made cameras, QuiQui’s Giant Bounce works on a Windows computer equipped with a low-end USB camera and microphone. To help

children stay in the camera’s view, a Web cam image of the user appears at the bottom-left corner of the screen as seen Figure 1 (middle, right).

Studying Children’s Intuitive Gestures

Intuitive game controls are largely dependent on a child’s previous experience, the distinctiveness of the game world, and a child’s conception of movements that could be carried out in the game environment. Thus, the narrative context as well as game character’s animation play an important part in defining



Figure 2. The wizard observes player's actions and controls the games accordingly. The swimming and running games are presented on the left.

the gestures for controlling the game. For example, when QuiQui needs to fly and is depicted holding two leaves in his hands, the player most likely tries to flap his or her hands to control the character.

Intuitiveness also relates to having natural mappings between the avatar and player movements even in cases where the avatar does not act as a mirror image of the user, or the perspective or orientation of the virtual world differs from the real world. For example, when QuiQui is flying in the air, there is no obvious mapping for the user to move sideways. Since games should allow the player to explore unknown worlds, the natural interaction style between an avatar and a player must be found through iterative design and testing with children.

Usability testing with functional prototypes. We initially used fully functional prototypes in traditional usability tests with 28 children ages 5–9 to evaluate how they controlled an avatar [7, 9] flying in different directions. However, this approach required time-consuming reiteration of the vision algorithms due to incorrect assumptions of intuitive control gestures. In the first prototype, children could make QuiQui fly

upward by flapping both hands up and down. Steering QuiQui to the left or right was achieved by flapping only one hand.

Flying upward was an easy task for most children, but controlling sideways movement proved to be difficult and frustrating. When analyzing the videos we found the gesture most frequently attempted by the children was to lean to the side while flapping their hands as shown in Figure 1. The next version of the game implemented this manner to control QuiQui. In a subsequent test children gave less frustrated comments and spent on average 34% less time completing the game level, indicating a significant improvement.

Wizard of Oz prototyping. From our experiences, we found testing the intuitiveness of game controls does not necessarily require functional prototypes. We have used WOz methodology to first gather children's movements during simulated playing sessions and only then applied that data in refining the vision technology [8].

We carried out the WOz study with 34 children (14 boys and 20 girls) ages 7–9 at a local elementary school. The evaluated prototypes were single-player games where the player controls QuiQui, who swims, runs, jumps, and tries to escape from spiders. The study was conducted as a form of pair-testing to make the situation more relaxed for the children.

INTUITIVENESS RELATES TO HAVING NATURAL MAPPINGS BETWEEN THE AVATAR AND PLAYER MOVEMENTS EVEN IN CASES WHERE THE AVATAR DOES NOT ACT AS A MIRROR IMAGE OF THE USER, OR

the perspective or orientation of the virtual world differs from the real world.



Figure 3. The examples of the most common movement styles for children when QuiQui is swimming, diving, riding toward the sea surface, or running. The jumping column shows how one girl crouches (a) before taking off (b) and how she lands (c).

In these sessions, the children thought they were interacting directly with the games, but instead, an adult (the wizard) interpreted their gestures and controlled the action using a keyboard and mouse (as shown in Figure 2). To study what gesture patterns would emerge, the children were not given any hints other than the visual cues of the game. The video material was shot to design and test computer-vision algorithms, and to facilitate further game character design.

Analyzing children’s gestures. Although no single practice for analyzing and annotating human movements in HCI has yet established itself as the prevalent method (possibly due to the high variation in gestures used with different systems and input techniques), research on nonverbal communication and gesture recognition [1, 12] can provide valuable guidelines also in the study of children’s game controlling movements and their phases.

The analysis of children’s intuitive game-controlling gestures is twofold. First, we must determine what movements children prefer in a particular game context. Second, we must study the properties and individual differences in the children’s movements such as the components of the gesture that are repeated, range of motion, symmetry, pace, space used, and transitions from one movement to another.

The movements were analyzed using video annotation techniques. Here, we present the most common types of intuitive gestures gathered during the WOz prototyping sessions (see Figure 3 for reference). Detailed observations, movement descriptions, and children’s interviews are beyond the scope of this article.

Swimming and diving are movements containing various styles that could not be anticipated in advance, for example, due to QuiQui’s animations. In the swimming game, QuiQui was portrayed moving sideways and he had to collect pearls by swimming or diving on top of them. The study of children’s swimming movements showed that children most easily adopted the dog paddle (main style for 38% of chil-

dren), breast stroke (35%), and crawl (12%) as shown in Figure 3.

Surprisingly, the avatar animations (QuiQui's crawl) did not seem to restrict or direct the children's movements. Even though QuiQui moved from left to right, the children did not turn in the direction of the character until they were forced to change QuiQui's route in order to continue. This also means that QuiQui's animations had to be redesigned to better match the children's movements. Additionally, any earlier swimming training did not seem to fully correspond with their movements either, as could be seen in the popularity of the breast stroke, which Finnish children do not usually learn until they are 10 or 11 years old.

Although the children could easily imitate realistic swimming gestures with their hands, making a character dive up or down was more confusing for them. Here, the direction of the movement was key. To dive downward, the children continued swimming, but directed the movement toward the floor and accentuated that by bending their knees or upper body. To get QuiQui to rise to the surface, the children continued swimming, but with their arms above shoulder level. Some even jumped or rose on their toes.

In the running game we tested how children controlled QuiQui's running, passing trees, and collecting butterflies by jumping. As opposed to various swimming and diving styles, the children's jumping was most similar across all participants. The duration and height values in the different phases of the jumps are shown in the table here. The preparation phase (crouching before taking off) could be used as a pre-action cue to make vision algorithms react quicker when the actual jump begins.

Additionally, the game would behave more realisti-

Phase	Mean duration (in ms)	Post-phase height in relation to the children's head length.
Preparing	212.3	-0.65
Taking off	308.0	0.84
Landing	287.4	-0.41
Recovery	342.8	

The duration and height values of 130 jump samples. The height values represent how high or low the children moved in relation to the length of their heads at the end of each phase shown in Figure 3. The length of the head is one unit.

cally if QuiQui crouched as well. The recovery phase gives a cue of when the next movement can begin, which in turn affects the pacing of the game. In general, if the game does not tolerate a long enough recovery period after a movement, children can become frustrated if they are not able to succeed due to their physical performance. However, the avatar doesn't necessarily need to follow the player's exact jump height (as long as the player knows how high the avatar jumps), but it is important that QuiQui responds fast enough when a child takes off.

The study of the video material also revealed some design issues to be studied further. Improper system latency leads to redundant actions and oscillation [4], which breaks the illusion of tightly coupled interaction. This could also be seen as the successive jumps by the player if QuiQui did not respond fast enough.

When QuiQui was running in a field presented in a semi-3D perspective, the children were not sure when a butterfly was close enough to be caught. Thus, the avatar's range and distance from platforms or objects to be reached should be clear enough; otherwise children might accentuate their own movements unnecessarily or approach the objects repeatedly.

To make QuiQui run, the children used both subdued (minor vertical alteration of the head level due to small leg movements) and lively movements. Surprisingly, there was a gender difference in the preference of these movements. Nine of the 11 boys favored subdued styles such as marching, strutting, or walking as shown in Figure 3. However, girls seemed to prefer more dynamic styles; 18 of the 25 styles the girls used were either running or swinging the feet from side to side. The problem with the subdued styles is they are difficult for the vision algorithms to recognize if the camera's field of view is only capturing the player's upper body and most of the movement is appearing below the waist. Thus, in the process of building a running game, the designers must pay attention to clear instructions and the presentation of the running avatar.

Even though there are individual characteristics in children's movements, patterns do exist **AND THIS INFORMATION CAN BE USED TO FURTHER IMPLEMENT BOTH COMPUTER-VISION ALGORITHMS AND AVATAR ANIMATIONS.**

Changing the direction of a movement (for example, running left to right or altering hand levels when swimming), does not generally cause problems for children. They did, however, have trouble in some transitions. For example, a direct transition from running to jumping was challenging even for some seven year olds as it requires being able to enter the preparation phase of the jump while having only one foot on the floor at the end of the running cycle.

One-third of the transitions from jumping to running or halting consisted of a so-called secondary jump the children used to dampen the previous jump. This behavior, however, can cause problems if this type of jump is tracked like a real one. In addition, transitions that cause children to lose visual contact with QuiQui, such as diving by crouching, are cumbersome. Visual contact with the avatar must be regained before being able to continue.

Young children, in particular, should be given enough time between exiting one movement and entering another. Generally, halting between movements should be allowed but children should be encouraged to practice more fluent transitions, for example, by rewarding them with extra points.

Children can play action games for longer periods of time when they use traditional game-control devices as opposed to vision-based games that require gross motor skills, continuous broad movements, and rapid shifts from one movement to another. Extended physical play might result in a lowered game experience, less easily detected movements (for example, running turns into walking), or even injuries due to tiredness. According to our experiences, resting should be provided at approximately four-minute intervals for children 5–6 when they flap their hands to control a flying game character. Thus, effective game design takes into account resting and other non-play time, and also varies the movements to prevent children from straining themselves.

Conclusion

Controls for vision-based action games must be intuitive and physically appropriate. These qualities can be studied most reliably by involving children in an actual play situation. Although we have previously adopted traditional usability testing with functional prototypes, we believe that game concepts and controls can be evaluated and redesigned less laboriously using WOz simulations. We have found the system latency introduced by the wizard is tolerable even in action games. Our study also shows that even though there are individual characteristics in children's movements, patterns do exist and this information can be used to further implement both

computer-vision algorithms and avatar animations.

Current computer vision-based game development tends to focus on the upper body (that is, hand, head, and torso movements) due to technological constraints. Nonetheless, future game design should aim for more holistic movements such as jumping, running, and even richer combinations for the benefit of children's physical development. **C**

REFERENCES

1. Aggarwal, J.K., and Cai, Q. Human motion analysis: A review. *Computer Vision and Image Understanding* 73, 3 (Mar. 1999), 428–440.
2. Bobick, A., Intille, S., Davis, J., Baird, F., Pinhanez, C., Campbell, L., Ivanov, Y., Schutte, A., and Wilson, A. The KidsRoom: A perceptually based interactive and immersive story environment. *Presence, Teleoperators and Virtual Environments* 8, 4 (Aug. 1999), 367–391.
3. Cordes, C., and Miller, E., Eds. *Fool's Gold: A Critical Look at Computers in Childhood*. Alliance for Childhood, College Park, MD; www.allianceforchildhood.net/projects/computers/computers_reports.htm.
4. Crowley, J.L., Coutaz, J., and Berard, F. Things that see. *Commun. ACM* 43, 3 (Mar. 2000), 54–64.
5. D'Hooge, H., and Goldsmith, M. Game design principles for the Intel Play Me2Cam Virtual Game System. *Intel Technology J.* Q4 (2001).
6. Druin, A., Bederson, B., Boltman, A., Miura, A., Knotts-Callahan, D., and Blatt, M. Children as our technology design partners. A. Druin, Ed. *The Design of Children's Technology*. Morgan Kaufman, San Francisco, CA, 1999, 52–72.
7. Hämäläinen, P., and Höysniemi, J. A computer vision and hearing based user interface for a computer game for children. In *Proceedings of the 7th ERCIM Workshop: User Interfaces For All*. (Paris, Oct. 2002), 299–318.
8. Höysniemi, J., Hämäläinen, P. and Turkki, L. Wizard of Oz prototyping of computer vision based action games for children. In *Proceedings of International conference on Interaction Design and Children*. (College Park, MD, July 2004), 27–34.
9. Höysniemi, J., Hämäläinen, P., and Turkki, L. Using peer tutoring in evaluating the usability of a physically interactive computer game with children. *Interacting with Computers* 15, 2 (2004), 205–225.
10. Krueger, M., Gionfriddo, T., Hinrichsen, K. VIDEOPLACE—An artificial reality. In *Proceedings of CHI 85*. ACM Press, New York, NY 1985, 35–40.
11. Sony EyeToy; www.eyetoy.com.
12. Wilson, A., Bobick, A., and Cassell, J. Recovering the temporal structure of natural gesture. In *Proceedings of Second International Conference on Automatic Face and Gesture Recognition*. (Killington, VT, 1996).

JOHANNA HÖYSNIEMI (johanna@cs.uta.fi) is a researcher with the Tampere Unit for Computer-Human Interaction, Department of Computer Sciences at the University of Tampere, Finland.

PERTTU HÄMÄLÄINEN (pjhamala@tml.hut.fi) is a researcher with the Telecommunications Software and Multimedia Laboratory at Helsinki University of Technology, Finland.

LAURA TURKKI (laura.turkki@hut.fi) works as a usability professional at Helsinki University of Technology, Finland.

TEPPO ROUVI (teppo@rouvi.com) is a freelance animator and works as a usability professional with a Finnish telecommunication company.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.