Computer-Simulated Display to Advance the Understanding of Perceptual Motor Skills

Hirofumi IDA*

Department of Human System Science, Tokyo Institute of Technology, Tokyo, Japan

Abstract

Computer-simulated displays, known as computer graphics animation or virtual environments, provide biological psychologists and behavioral neuroscientists with novel inspiration in the study of human behavior. This review discusses current research trends in the use of computer simulated displays for the assessment of human perceptual motor skills, particularly sports related behaviors. Two research areas of longstanding interest to biological psychologists are reviewed, i.e., biological motion perception and perception-action coupling. A brief discussion of relevant case studies is also provided. The effective use of computer-generated visual stimuli can provide new insights into standard research questions and it is likely to provide new research insights in the future. However, these displays may have some disadvantages in terms of reality, or a sense of presence, when attempting to simulate a real-world situation. Researchers who study human perceptual motor skills should never neglect the ecological validity of their experiments, but computer-simulated displays will doubtless become a powerful tool for the investigation of motor behaviors.

Keywords: Computer graphics; Virtual environment; Biological motion perception; Perception-action coupling; Humanoid model; Sports-related behavior

Background

Computer graphics (CG) is no longer a mere visualization tool. Researchers in the fields of biological psychology and behavioral neuroscience use CG visual stimuli in a wide range of experimental scenarios. Complicated animations of human actions and physical phenomena are readily achieved in perceptual and cognitive tests as well as moving simple objects. Virtual environments (VE) developed with CG models, also known as virtual reality, provide users with a stereoscopic or pseudo three-dimensional visual perception experience. These computer-simulated displays now provide scientists with further tools for exploring the nature of visual perception and cognitive processing in more ecologically valid conditions than previously possible, while still allowing adequate control and simple manipulation.[1-3] This is mainly attributable to technical advances in CG rendering software and computer performance and the popularization of visualization tools such as auto-motion capture systems. A wide range of studies makes use of computer-simulated display, but this review will focus on visual perception and motor control in high-demand situations, i.e., sports-related skills. Athletes are required to acquire beneficial information, make an optimal judgment, and perform a fine-scale motion under spatially and temporally constrained conditions. One of the main integrative approaches is the assessment of perception-action coupled responses in simulated sports situations using VE.[4-13] There are high user expectations in these situations because simulated environments are employed for individual training purposes and for the coaching of trainees. Basic research on motion perception and motor responses is now presented to provide greater detail on sports-specific skills, beginning with biological motion perception.[14] Biological motion perception refers to the capacity of the visual system for recognizing complex human or animal movements, even when such actions are presented as a few moving dots (Figure 1). Previous brain imaging studies show that biological motion selectively elicits activity in the observer’s brain, particularly in the superior temporal sulcus.[15,16]. The second area is that of perception-action coupling, or the stimulus-response compatibility. This was recently developed further using common coding theory to suggest a shared representation of perception and action[17], and this should be taken into consideration when studying high-demand activities. In particular, the control schema of an interceptive activity such as ball catching is considered to require an interaction between perception and action[18,19]. This paradigm has been investigated using VE simulations. The first half of this article reviews the relevant studies on biological motion perception, from the early classical work to new computer-based techniques. At the end of this section, a case study on tennis motion perception will be introduced. The second half of this article reviews studies on perception-action coupling in VE, which have been growing in recent decades. This is followed by another case study that investigated virtual catching performance. This article has the goal of presenting milestone findings and promising approaches to facilitate the understanding of perceptual motor skills in the context of computer-simulated display.

Approach Based on Biological Motion Perception

From a point-light model to a polygon model

A point-light model is traditionally used to represent the joint positions of a gross motor pattern such as walking and it can provide the visual stimuli of biological motion (Figure 1). This approach was originally applied based on the rigid assumption that two or more moving dots are perceived as the end points (joints) or corners of a moving object (body segment) in an interpolative fashion. Several studies have demonstrated that the human visual system is capable of extracting the characteristic features of point-light motion, e.g., in
discerning the gender of a walker [20,21] and in identifying the walks of friends [22]. Biological motion also conveys mental states such as intention [23] and emotion [24,25]. The concept of biological motion is traced to its inception when starting with a point-light model. Thus, filmed, photographed, and CG-rendered real-world human motions are now frequently used as visual biological motion stimuli. In particular, auto-motion capture systems facilitate the parallel use of a point-light model and a polygon humanoid model to replicate motion data [26-28]. A previous study showed that the superior temporal sulcus was activated by a biological gait motion that was represented by both human-like and robot-like polygon models, whereas there was no response to the nonbiological motion such as disjointed objects and a grandfather clock [29,30].

Point-light display of competitive situations

For decades, sports psychologists have investigated the perceptual skills required to read a competitor’s intentions in a competitive face-to-face situation, i.e., the information pick-up strategy used by players while attempting to read their opponent’s next move [31,32]. Numerous studies with video displays and live situations have revealed the superior anticipatory performance of skilled players compared with less skilled counterparts, e.g., in tennis [33-35]. Studies have also detected expert-novice differences using a point-light display [36,37]. However, another study detected no performance differences between skilled and less skilled players using a point-light display, whereas the expert superiority was evident in live and video displays [34]. More recently, the use of a spatially occluded point-light model that presented selected body parts in isolation demonstrated that world-class badminton players acquire beneficial pre-contact information based on the opponent’s lower body and racket positions [38]. A point-light model using digitized coordinate data obtained from motion analysis facilitates the computational manipulation and modification of the biological motion that is displayed.

Use of other CG humanoid models

It is now easier than ever to produce detailed CG humanoid characters. In addition to simple models, e.g., stick-figure models [39-42] and ellipsoid models [10], finely textured polygon models are now available for researchers thanks to CG modeling software (Figure 2). For example, it was shown that the limb movements of a handball goalkeeper facing an avatar of the thrower in VE were almost equivalent to the real-world situation [4,6]. Furthermore, receivers in tennis can extract pre-contact information about the ball direction based on the motion of a polygon humanoid character, although the anticipatory performance in a CG display was less accurate than that in the corresponding video display [43]. The perceptual effects of information properties such as contour and texture that are embedded in a model were also evaluated using comparisons with different humanoid models. The perturbation of running motion was more accurately discriminated when the motion was displayed using a polygon model than with a stick-figure model [44]. Furthermore, when using a textured polygon model of a handball thrower as a reference, a point-light model and a silhouette model evoked different response patterns to a goalkeeper’s limb movement [11,13]. However, those studies also showed that other indices were not significantly different from humanoid model displays, i.e., time to respond, percentage of successful movements, and radial error (distance between goalkeeper’s limb and ball). The selection of a humanoid model is likely to affect the evoked perceptual response and it might potentially have a critical effect on observational learning and visual training in sports.

Manipulation of the test motion

A major advantage of using CG animation is the simplicity of computational manipulation and the availability of the manipulated environment. Many methods have been applied to motion displays when investigating the performer-observer paradigm in sports situation, i.e., the perturbation of motion using techniques such as dynamic simulation and noise addition [44]. Spatial exaggeration is also a perturbation technique and a previous study using this method in the study of tennis serve motion using a polygon humanoid model showed that a higher exaggeration rate increased the accuracy of serve type categorization [45]. Another study manipulated the tennis serve motion while maintaining the anatomical constraints on joint rotation and found that the anticipation of the ball direction was monotonically shifted based on the perturbation rate of the forearm pronation/supination and elbow extension/flexion of the polygon server [46,47]. Motion synthesis is useful for composing test motions. Using a normalization technique based on the decomposition of principal component analysis [21], tennis strokes were synthesized and displayed to tennis players using a stick-figure model [41]. The results suggested that tennis players extracted anticipatory cues of the ball direction, which was embedded in multiple low-dimensional dynamic modes of the performer’s motion. The technique was also utilized in subsequent studies that attempted to test the contribution of local (each body segment) kinematic information to the judgment strategy, by introducing motion perturbation techniques for occlusion and neutralization [46], and interchange [39,42]. Studies have generally provided evidence that skilled players rely on a global strategy for predicting the ball direction, where they effectively utilize visual

![Figure 1: Point-light model.](image1.png)

![Figure 2: Point-light (A), stick-figure (B), silhouette (C) and textured-polygon (D) models used for simulating tennis serve motion.](image2.png)
information from the opponent’s body areas as well as deterministic cues from the end-effector (racket).

Case study 1: Anticipatory judgments of complete novices

Although numerous studies have demonstrated that skilled players outperform less skilled players in terms of accuracy and speed of their responses, there is little understanding of the response characteristics of novices with no experience of playing with experts. Taking tennis as an example, depending on the experimental setting, the anticipation accuracy of early beginners could surpass a guessing level [34], but occasionally the accuracy might be significantly lower than chance [48]. A case study investigated the perceptual performance of complete novices when they viewed a captured tennis serve motion and made an anticipatory judgment of the ball’s direction, where an intervention with explicit feedback was given between the pre and post-tests. It has been stated that explicit instructions about the play should be presented to beginners during the first cognitive phase of learning [49].

Materials and methods: Two expert tennis players ranked in the Japan Tennis Association ranking, who were 22 and 25 years of age, acted as model players when creating test serve motions. The serve motions of the model players were aimed in three directions (center, body, and wide) and they were captured using an auto-motion analysis system with 48 retroreflective markers. Six test serve motions (2 model players × 3 directions) were reconstructed using a wire-frame model (see illustration in Figure 3). Fourteen novices who were naive to the test stimuli with a mean age of 23.9 years (SD 1.1) viewed the test serve motion and made an anticipatory judgment of the ball direction in pre- and post-test sessions. After finishing the pre-test the participants received feedback about the ball direction, before they moved to the post-test.

Results and discussion: The percentage of correct responses significantly increased in the post-test when compared with the pre-test, t(13) = 2.47, P = 0.028 (Figure 3). This indicated that feedback and familiarization had an immediate effect on improving the recognition performance of participants. The number of responses in the pre-test had a significant effect on ball direction F(2,26) = 6.89, P = 0.004, and post hoc multiple comparisons revealed a significant difference between the center (left side of the receiver) and wide (right side of the receiver) serve directions P = 0.001 (Figure 4). This response bias was not observed in the post-test. These results suggest that naive novice tennis players are likely to make a biased response to the left side when viewing a highly skilled tennis serve from the opponent receiver position. However, the response bias could be refined by familiarization along with explicit feedback that would work as the source of new knowledge [49].

Conclusions of biological motion perception approach: The first section discussed CG display techniques that advance the understanding of information pick-up strategy on the basis of existing biological motion perception approaches. The selection of CG models (e.g., point-light, stick-figure, and polygon) and the manipulation of the displayed motion (e.g., occlusion, perturbation, and synthesis) have considerable potential to provide further knowledge of skilled perceptual skills. A case study has demonstrated an example of a novice player’s biased response in tennis, which is one of the aspects that are different from the conventional expert-novice paradigm.

Approach Based on Perception-Action Coupling

Perception and action in simulator

VE is a computer-simulated environment that generally offers a stereoscopic visual experience to users, while some VE simulators are equipped with other displays such as stereophonic speakers and haptic devices. Motion tracking devices are also installed that capture user actions, e.g., camera-based optical systems and six-degrees-of-freedom magnetic systems. The vision-based perceptual effect is the main concern in sports-related behavior studies, where head-mounted displays and large projector screen systems are often used as simulation devices. One of the best known simulators is an immersive VE system known as CAVE (Cave Automatic Virtual Environment, (Figure 5) that provides room-sized multiple cubic walls allowing users to view objects three-dimensionally with stereoscopic glasses as they walk around in the space enclosed by the screen [50]. CAVE has been utilized in studies on automobile driving [51] and surgery training [52]. Researchers have investigated the paradigm of perception-action coupling in human motor control using immersive VE. Like CG, VE has many advantages in terms of control and the ease of manipulation in the environment, e.g., the viewpoint can readily be changed by moving the camera position (Figure 6). The immersive and interactive...
features of VE are advantageous for examining the pattern of online motor control where perception and action are mutually connected.

Interceptive activity in VE

Ball catching is a major experimental task used to investigate the interaction between the perception of an object and the coincident action of a performer during an interceptive activity. The Chapman strategy, also known as the outfielder problem, addresses how a fielder runs to catch a flying ball [53] and this problem has been investigated by cognitive psychologist and psychophysicists for decades [54,55]. As well as conventional real-world tests, several studies have used VE simulation to further understand the control strategy used in the fielder's running path. A study using the CAVE system showed that users could successfully intercept a virtual flying ball in a VE simulator [56]. This was also confirmed with a head-mounted display for soccer ball heading [57]. This study computationally modulated the ball trajectory in mid-flight from a normal quasi-parabolic motion to an artificial linear one. The results supported the optic acceleration cancellation theory that the angle of gaze elevation increases at a decelerating rate in an online control manner during the locomotion of the catcher. This agreed with a subsequent study where a similar technique for the modulation of the ball trajectory was applied to a baseball flying catch performed in a relatively wide (12 × 12 m) virtual environment [58].

VE simulation of competitive situations

Sports-related face-to-face situations have also been simulated in VE, although there is an overlap with studies on biological motion perception with regard to the experimental framework [4-6,10,11,13]. These situations generally require more temporally constrained control compared with simple flying ball catching. A pilot study examined motor behavior during baseball batting where the approaching ball was presented using a video monitor in monocular conditions and the bat motion was measured [59]. It was shown that baseball batters changed their swing pattern depending on the history of previous pitches, pitch count, and the rotational motion of the ball. Another pitcher-batter situation was subsequently reconstructed in a three-dimensional immersive VE [10]. The stepping motion pattern of batters was attuned to the motion of the opposing pitcher, while the swing time was related to the ball speed after release. A simulation of a soccer free kick in an immersive and interactive VE showed that the goalkeeper’s hand movement was biased to the initial direction of the ball’s flight [8]. In addition, the online movement pattern of goalkeepers was successfully explained using a mathematical model of the current position and velocity, whereas lateral acceleration was not used as a control variable.

Case study 2: Ball catching in two- and three-dimensional displays

Subects with normal stereoscopic acuity exhibit a reduced ball catching performance in monocular conditions when compared with binocular contexts, whereas an inherent weak stereoscopic acuity negates the effect of these viewing conditions on catching performance [60,61]. This suggests that stereoscopic information critically affects human perceptual motor performance. A case study investigated this stereoscopic effect from the aspect of VE display. To this end, the study compared the effect of two-dimensional (2D) and three-dimensional (3D) visual displays on virtual ball catching performance and this was compared with real-world catching.

Materials and methods: Twelve male collegiate students, with a mean age of 20.2 years (SD 1.4), performed a virtual ball catching task in CAVE under 2D (no visual disparity) and 3D (with default disparity) display conditions. They were also asked to catch a ball under quasi-randomized speed conditions in a real-world setting. The original coordinate data from the virtual ball in CAVE were acquired from the actual flight motion projected by a ball machine. This ball machine was also used for the real-world catching test.

Results and discussion: The percentage of correct catches was significantly higher in 3D conditions compared with the 2D conditions t(11) = 2.90, P = 0.014 (Figure 7). This supported the results of Mazyn et al. (2004). In the test of the catching performance under quasi-randomized speed condition, the Pearson’s correlation coefficient of 3D displays in the real-world (r = 0.490) was higher than that of the
2D display ($r = 0.255$). This indicates the superiority of the 3D display over the 2D display, in the sense that it evoked a more similar motor response to that found in real-world. Further studies using more elaborate motion analyses will be required to judge the potential applicability of stereo-displays.

**Conclusions of perception-action coupling approach:** The second section focused on perception-action coupled behavior in VE, particularly in regard to sports-simulated situations. Several studies have demonstrated that, to some extent, VE simulators can evoke natural responses in users, not only to displayed objects, but also to performing players. A case study has suggested the effectiveness of using a VE stereoscopic display for simulating ball catching behavior.

**Summary**

The use of computer-simulated displays facilitates the study of perceptual motor skills. Studies of skilled performance using CG and VE displays began in the 2000s and an increasing number of studies have made use of these display techniques. This review considered the traditional concepts of perception and action to illustrate the characteristic motor skills that can be investigated when computer-simulated visual stimuli are presented to a viewer, i.e., biological motion perception and perception-action coupling. Case studies were also provided to illustrate these two concepts. There are a number of intrinsic limitations when using computer-simulated displays. For example, the sense of depth in VE system is usually provided by projecting two slightly different images on the screen to generate a binocular disparity on the left and right retina. This is a simple depth cue, but the depth perception of the visual system also relies on additional monocular information that may be difficult to reconstruct in a two-dimensional image, e.g., ocular accommodation. Furthermore, the system processing performance of a complex VE system sometimes restricts the assessment of motor performance. In particular, the high-speed nature of sports-related behavior presents a problem in terms of computational power [56,62]. The settings of various visual effects should be carefully determined to match the task requirements, e.g., lighting and shading, texture mapping, and motion blur. The use of computer-generated visual stimuli can be used to address standard research problems related to human perceptual behavior and motor responses, but it might also have the potential to develop new research areas in biological psychology. CG animation and VE systems are unquestionably powerful tools in behavioral science, but an understanding of real-world human behavior is necessary to effectively manipulate the visual conditions used and practically interpret the data obtained.

**Acknowledgments**

The author would like to thank Dr. Kazunobu Fukuhara, Ms. Madoka Takahashi, Mr. Takahiyo Ogata, and Mr. Takahiro Noguchi for their valuable assistance when designing and conducting the experiments. The author also thanks Prof. Motonobu Ishii and Prof. Tetsuji Inoue for many helpful suggestions. This work was partially supported by a Grant-in-Aid from the Japan Society for the Promotion of Science (KAKENHI 22240070 and 23700715).

**References**