Published in final edited form as:

Int Conf Virtual Rehabil. 2017 ; 2017: .

Development of a novel visuomotor integration paradigm by integrating a virtual environment with mobile eye-tracking and motion-capture systems

Haylie L. Miller, Ph.D.,

Departments of Physical Therapy, University of North Texas Health Science Center, Fort Worth, Texas, USA

Nicoleta Bugnariu, PT, Ph.D.,

Departments of Physical Therapy, University of North Texas Health Science Center, Fort Worth, Texas, USA

Rita M. Patterson, Ph.D.,

Family Medicine, University of North Texas Health Science Center, Fort Worth, Texas, USA

Indika Wijayasinghe, Ph.D., and

Department of Electrical and Computer Engineering, University of Louisville, Louisville, Kentucky, USA

Dan O. Popa, Ph.D.

Department of Electrical and Computer Engineering, University of Louisville, Louisville, Kentucky, USA

Abstract

Visuomotor integration (VMI), the use of visual information to guide motor planning, execution, and modification, is necessary for a wide range of functional tasks. To comprehensively, quantitatively assess VMI, we developed a paradigm integrating virtual environments, motion-capture, and mobile eye-tracking. Virtual environments enable tasks to be repeatable, naturalistic, and varied in complexity. Mobile eye-tracking and minimally-restricted movement enable observation of natural strategies for interacting with the environment. This paradigm yields a rich dataset that may inform our understanding of VMI in typical and atypical development.

Keywords

virtual environment; motion-capture; mobile eye-tracking; visuomotor integration

I. Introduction

Virtual environments have demonstrated validity for application in both research and rehabilitation/intervention [1, 2, 3]. Researchers have used virtual environments in both small-(e.g., joystick/mouse/keyboard controlled) and large-scale (e.g., surround screens, head-mounted devices [HMDs], CAVE systems) in combination with kinematic data to assess motor control and motor rehabilitation [4] or in combination with mobile eye-tracking to assess attention and gaze in semi-naturalistic tasks [5]. These technologies can provide

quantifiable data used to increase understanding of mechanisms that give rise to human motor behaviors and motor task performance. However, few facilities are instrumented for the use of immersive virtual environments to deliver user-controlled semi-naturalistic tasks in combination with both kinematic and eye-movement data.

Considered in isolation, kinematic and eye-movement data are independently useful in understanding the motor or visual system in typical and atypical development. But when studied in combination, they open new avenues for the study of *visuomotor integration*—the use of visual information to guide motor planning, execution, and modification. Visuomotor integration is a critical building block in real-world tasks such as postural control [6] and reaching [7].

With the aim of studying visuomotor integration in a comprehensive and multi-faceted way, we have integrated an immersive virtual environment, an instrumented treadmill, a full-body motion capture system, and mobile eye-tracking.

This combination of instrumentation presents technical challenges, but allows examination of important relationships between visual and motor systems that are linked in their structural and functional neural underpinnings, as well as in their behavioral consequences. In service to that goal, we developed a paradigm of eight novel tasks designed to target specific aspects of visuomotor integration and capitalize on the rich and complex dataset generated by the integration of these three systems.

II. Method

A. Apparatus

Data in our paradigm are integrated from three systems: a virtual environment and instrumented dual-belt treadmill with embedded force plates (V-Gait CAREN system and DFlow software suite; Motek ForceLink, Amsterdam, Netherlands), a 12-camera infrared motion-capture system (Eagle-4 cameras and Cortex software suite; Motion Analysis Corporation, Santa Rosa, CA, USA), and a mobile eye-tracking system (ETG 2.0 and BeGaze software suite; SensoMotoric Instruments, Boston, MA, USA). The instrumented lab space is depicted in Fig. 1.

a) Virtual environment—The V-Gait CAREN system (Fig. 2) includes a dual-belt instrumented treadmill that can deliver pitch (± 10 degrees variable incline) and sway (± 10 cm sideways) perturbations. Two force plates are embedded under each belt. The DFlow software suite collects force plate and behavioral data from the V-Gait CAREN system at a variable sampling rate averaging 120 Hz. DFlow sampling rate ranges from approximately 100–300 Hz based on processing demands related to rendering the 3D environment and writing data to file. Data from the DFlow system are transformed to temporally align with the sampling rates of the motion-capture and eye-tracking systems.

The virtual environment is projected on a 180-degree wrap-around screen (3 m tall $\times$ 5 m diameter) by 3 ceiling-mounted projectors and a custom driver to blend images. The origin

of the virtual environment ([0,0,0] in the XYZ coordinate space) is at the center of the treadmill, 1.5 m from the screen.

The force plates are calibrated at 3 points along the center of each force plate (for a total of 6 points) prior to each data collection, to ensure minimal artifact (± 10 N).

b) Motion capture—The 3D position of each reflective marker is captured by a real-time motion capture system consisting of 12 Eagle-4 infrared cameras designed for kinematic testing of normal and pathological motion. Two cameras are mounted on tripods near the base of the treadmill, and 10 cameras are mounted on a truss above the system at ceiling height (see Fig. 1 for camera array). Data are streamed from the infrared cameras at 120 Hz and recorded by Cortex, a software package designed for processing kinematic data. Data streams between Cortex and DFlow are linked with a pulse used to sync the recordings in time. Data output includes the position, velocity, and acceleration of individual markers as well as joint angles and body segments.

The motion capture system is calibrated prior to each data collection, to ensure that marker XYZ positions are accurate within 0.5 mm and that the XYZ position of the origin of the motion-capture space (0,0,0) matches the origin of the virtual environment (0,0,0).

Markers are placed on 28 anatomical landmarks of the participant's body in a set optimized to capture the head, torso, pelvis, and feet. Marker sets for a variety of child and adult body sizes were developed in Cortex and tested for robustness, with stretch parameters built in to accommodate for changes in distance between markers due to natural body movement. A dynamic capture was also collected for each participant to further optimize the model to accommodate their natural range of motion during a movement similar to the demands of the experimental tasks. This process of individualized adjustment resulted in stable marker sets that were able to accommodate participants of all sizes and motor abilities.

c) Mobile eye-tracking glasses—We used a mobile eye-tracking system (Fig. 3, ETG 2.0; SensoMotoric Instruments, Boston, MA) which consists of a recording device (Samsung Galaxy with iView ETG software) and glasses with three 60 Hz digital cameras. One camera records the scene with a 60° horizontal and 46° vertical field of view (FOV), and two cameras record binocular pupil position, acceleration, and velocity. The gaze tracking range is approximately 80° horizontal and 60° vertical. The glasses weigh 47 gm, and are adjustable to fit the shape and size of a participant's nose.

Prior to data collection for each participant, the glasses are adjusted for proper fit, and calibrated using a 3-point triangulation procedure based on objects displayed the virtual environment. Calibration verification is conducted by verbally instructing participants to fixate each of 13 objects displayed in a fixed array adjusted to the participant's eye level (Fig. 4).

A recording of calibration verification is made for each participant, so that any necessary adjustments can be made in post-processing. This is rarely necessary in typically-developing participants, but proves useful when participants in clinical populations present with

oculomotor complications such as monocular nystagmus, strabismus, or amblyopia that can impact the quality of calibration.

All objects used in calibration and verification are located in the positions of objects that participants will encounter in the experimental tasks. The Y coordinates of all stimulus objects are set to the Y position of a marker located on the upper right corner of the eye-tracking glasses, effectively adjusting all stimuli to the participant's eye height. The position of this marker is captured prior to each task, while a participant stands upright in a natural, comfortable stance with feet approximately shoulder-width apart.

The BeGaze software suite is used for basic processing of saccadic and fixation eye movements, including offset correction, marking and removal of artifact from blinks, marking saccades, marking fixations of > 200 ms, and calculating the proportion of time a participant spent looking at particular areas of interest (AOIs).

d) Tethers and harness—Head-eye integration strategies during pursuit eye movement used to track visual motion are notably different when employing head-free versus fixed-head methodology [8]. In this paradigm, the eye-tracking glasses are tethered only to the recording device worn in a pack on the participant's back. This tether does not restrict or limit head motion, and allows more natural head-eye integration strategies when using pursuit eye movement to track visual motion. Participants wear a lightweight harness for safety, which is sized and adjusted to fit each body type. They are tethered to support beams in the walls using side ropes attached to the harness. The ropes are adjusted for each participant to allow them to lean, step, or otherwise move within their comfort and the lateral limits of the platform.

Thus, participants are predominantly unrestricted in their use of strategies when tasks require a whole-body movement in response to a target. This, in turn, produces a more naturalistic set of motor behaviors than would be allowed by the use of fixed-head or body-restricted paradigms.

B. System Integration

We use software developed by the research team to register the coordinates of the eye data to the coordinates of the body markers and virtual environment, which enables calculation of time-on-target, positional accuracy of gaze (without requiring manual marking of each trial), and other related variables that require integration of the eye, motion capture, and virtual object data streams. Coordinates from the motion-capture system served as the base coordinate system, and coordinates from the eye-tracking and virtual environment systems were transformed to this coordinate space.

The eye-tracking coordinates are referenced to an origin approximately at the center of the right eye. Coordinate transformation was achieved by calculating the rotation and translation of the eye-tracking coordinate system relative to the motion-capture system, in part by using markers attached to fixed locations on the glasses. The specific algorithms used for data integration are described in greater detail elsewhere [9].

Since the virtual environment is calibrated to the motion-capture coordinate space prior to each data collection, these two coordinate frames are virtually identical. However, since 3-dimensional virtual objects are projected onto a 2-dimensional screen, a slight transformation was required to account for distortions that arise due to projection.

Four models were evaluated as potential transformations:

1. Translation and scaling;
2. Rotation, translation, scaling;
3. Perspective transformation, translation, scaling;
4. Perspetive transformation, rotation, translation, scaling.

Model fit was evaluated using six samples during which the point of visual fixation was at the approximate center of the target object. Performance of the model was measured by the root mean squared error (RMSE) of the angle between the object vector from the eye and the eye gaze vector in degrees:

$$RMSE = \frac{180}{\pi} \sqrt{\sum_{i=1}^6 \arccos \left(\frac{(p^b - p_{g, origin}^b) \cdot v^b}{\|p^b - p_{g, origin}^b\|_2 \|v^b\|_2} \right)^2}$$

where p^b is the object coordinate, $p_{g, origin}^b$ is the vector from the origin to the center of the eye, and v^b is the gaze vector after transformation to the base coordinate system. The second model, employing rotation, translation, and scaling, was ultimately the best fit for the data (Fig. 5).

The software developed and evaluated through this process results in data integration across the three systems. It enables visualization of body markers, eye gaze vectors, and virtual object locations simultaneously (Fig. 6), and data are written to an output file used in quantitative analysis.

C. Visuomotor Integration Tasks

We use a combination of commercially-available and custom-built virtual environment applications to deliver highly-controlled, semi-naturalistic tasks used to evaluate motor and task performance. These applications range from adaptations of traditional eye-movement tasks (e.g., fixation or pursuit of a static or moving object) to complex tasks that more closely approximate the attentional demands of the real world (e.g., a virtual cartoon shooting gallery). Here, we present only the custom application used to test visuomotor integration, which includes eight tasks of varying difficulty, stimulus complexity, and degree of motor response required.

Our paradigm of visuomotor tasks requires oculomotor (*watch-only*) and/or gross motor (*watch-and-move*) responses, and the difficulty and complexity of the tasks increases across a data collection session. Thus, early tasks serve as a baseline assessment of the fundamental oculomotor or gross motor skills necessary to complete later tasks. Watch-and-move tasks

require participants to make a physical movement in response to visual appraisal of a moving target in the virtual environment (e.g., lean or step laterally).

For example, basic oculomotor control skills were tested using a watch-only task, during which participants used saccadic and smooth pursuit eye movements to visually track a single oncoming target (Angled Pursuit). Basic postural control skills were tested using a watch-and-move task, during which participants needed to lean or step to match the position of a user-controlled object to a static target (Safe Zones). These “building block” skills were necessary for a more complex watch-and-move task that participants encountered later (Intercept). This task required participants to use information about visual motion to plan and execute leaning or stepping movements to match the position of their user-controlled object to the oncoming target (Intercept).

Standardized, pre-recorded instruction sets were presented on screen and via speakers, and a proctor was available for clarification if needed. Between each trial of all tasks, participants were instructed to fixate a cross in the center of the visual field, visible for 150 ms; a blank screen appeared immediately before and after each instance of a visual target (i.e., ball or fixation cross) for a total inter-stimulus interval (ISI) of 300–400 ms. This allowed the participant to start from approximately the same point for all trials.

a) Watch-only tasks—We developed a series of five watch-only tasks, which require participants to stand and watch static or moving stimuli. Watch-only tasks are presented in the following sequence:

- Quiet Standing: This task serves as a baseline measure of postural stability with and without the influence of visual stimulation and context in three conditions:
  1. 30 sec of eyes-open, with no projected image. Participants are instructed to, “Look straight ahead,” so that they do not attempt to obtain a visual point of reference for their position by looking down at the CAREN platform or the floor. In this condition, the virtual environment is black, removing a participant’s ability to judge postural sway based on visual cues. However, because the participant’s eyes are open and light is allowed to stimulate the retina, this condition induces visual stimulation without providing visual context or cues about depth or position.
  2. 30 sec of eyes-closed, removing the influence of both visual stimulation and visual context.
  3. 30 sec of eyes-open while fixating a cross in the center of the visual field. This condition provides visual stimulation along with visual context as feedback for maintaining postural stability.
- Fixation: This task serves as a control task used to acclimate participants to the stimuli, instructions, and system, as well as to confirm their ability to complete basic voluntary eye movements. Participants are instructed to fixate each of 13 static red balls (projected image 0.5 m in diameter) in a fixed-randomized

sequence. Each target appears for 500 ms, in locations where virtual objects are presented in subsequent tasks.

- Horizontal Pursuit: This task mirrors traditional pursuit tasks found in the fixed-head eye-tracking literature. Participants are instructed to track a red ball that appears in the middle of the screen for a random interval of 200–500 ms, then rolls across the visual field to the left or right limit of the field of view (FOV) of the eye-tracking system at a constant speed of 30°/sec, the optimal speed for human visual pursuit of motion [10].
- Peripheral Pursuit: While the horizontal pursuit task can be accomplished with either eye-only or a combination of head and eye movement, this task necessitates the use of both head and eye movement in visual pursuit of a moving target. The red ball appears at the right or left edge of the screen and rolls across the visual field at a constant speed of 30°/sec to the opposite edge.
- Angled Pursuit: This task requires integration of vertical and horizontal motion, given that targets roll down an angled grid. The grid has depth and height in the virtual environment, such that it is pitched toward the participant, creating the illusion of oncoming motion (Fig. 7). The red ball appears at the top left, right, or center of the grid, and rolls at 30°/sec to the bottom left, right, or center of the grid, for a total of nine trajectories.

b) Watch-and-move tasks—We developed a series of three watch-and-move tasks, during which participants make a body movement—lean or step—in response to static or dynamic visual stimuli. Tasks are presented in the following sequence:

- Safe Zones: This task serves to acclimate participants to how the user-controlled object (blue ball; Fig. 8) responds to changes in their body position when they lean or step laterally. The ball is controlled in the medial-lateral direction by the X position of a marker on the C7 vertebrae, but participants are blind to this information. This task is similar to Fixation in that it confirms that participants have the ability to perform basic voluntary body movements or postural shifts to match the location of a static target. Participants are instructed to move their blue ball into the green safe zone. Safe zones are displayed in 5 positions corresponding to where virtual objects would be presented in subsequent tasks.
- Intercept: In this task, the stimuli are identical to that of Angled Pursuit. Participants are instructed to move their blue ball to “block” the red ball at the bottom of the grid, and not let it go past them (Fig. 9). The addition of a motor response necessitates accurate visuomotor integration, namely, motor planning and modification based on on-line updating of information about the speed and direction of visual motion.
- Avoid: In this task, the stimuli are identical to that of Angled Pursuit and Intercept. Participants are instructed to not let the red ball touch their blue ball.

III. Data Processing and Variables of Interest

Integration of the three data streams in a custom software package enables faster, more automated post-processing, in particular with respect to eye-movement data. This integrated approach to analysis eliminates the need for hand-marking individual saccades, fixations, and pursuit trajectories for each trial of each task based on video data, as is standard format of most commercially-available analysis software. Given the volume and complexity of output generated by the virtual environment, motion-capture, and eye-tracking systems used in the visuomotor paradigm described here, efficient processing and analysis is critical to generate the highest yield from the resultant dataset.

Variables of interest resulting from integrated approach to analysis are those that measure the accuracy of participants' performance relative to their eye movement and body movement, and those that measure the coordination of eye and body movement (e.g., use of head-eye integration to pursue visual targets). These include, but are not limited to:

- Time-to-target: the time elapsed from an initial saccade to accurate fixation of a target
- Initial saccade gain: the percent gain achieved in the primary saccadic eye movement to a static or moving target (calculated based on the first saccadic movement in the correct direction; for trials that begin with anticipatory saccades in the wrong direction, this distinction is meaningful).
- Average deviation-from-target: the average difference between the position of a participant's point of gaze and the target location achieved in either saccadic (for fixed targets) or saccadic and pursuit eye movement (for moving targets) across a trial (calculated as deviation-from-target at each sample averaged across all usable trials in a task, such that lower numbers represent more accurate performance).
- Pursuit trajectory accuracy: the difference between the slope of the participant's pursuit eye movement trajectory and the target trajectory.
- Head-eye ratio: the ratio of head-to-eye movement used to achieve fixation or pursuit of a target.

In addition to visuomotor variables available through integration of these three systems, standard oculomotor, kinematic, and performance-based data are available from each individual system. As a result, the yield from projects employing this paradigm is rich and diverse, lending itself to both fine-grained analysis of the relationship between specific mechanisms and task performance (e.g., postural stability and number of "hits" in the Safe Zone task) and comprehensive examination of relationships between larger systems (e.g., head-eye integration strategies and postural sway during pursuit of peripheral targets).

IV. Conclusions

Recent advances in technology have made it possible to obtain high-quality data with reasonable temporal and spatial resolution from mobile eye-tracking devices. Improvements

to virtual reality technology have produced more lifelike environments, and the ability to incorporate user-controlled objects when combined with motion capture. These advances have opened new avenues for investigation of naturalistic strategies for locating, tracking, and interacting with objects. Although prior work has been done using two of the three technologies simultaneously, few groups have used all three in an immersive setting allowing for full-body movement and head-free eye-tracking.

Successful integration of virtual environments, motion-capture, and mobile eye-tracking proved technically challenging but feasible. In order to maximize output from the resulting system, we developed a series of eight tasks to assess visuomotor integration in a comprehensive and quantitative way, with each task increasing in complexity and difficulty. This paradigm yields a rich and flexible dataset, and requires minimal instruction for the participant, making it ideal for use in both typically- and atypically-developing populations.

Future directions include online, real-time integration of data streams through a wireless network to enable users to control the virtual environment with eye movements as well as body movements. By developing new ways to assess visuomotor integration, we aim to better understand the relative influence of visual, motor, and attentional factors on functional movement and behavioral performance.

Acknowledgments

We thank the staff of the Human Movement Performance Laboratory and MotekForcelink, who were instrumental in developing this paradigm and the virtual environments. We also thank staff and students whose continued involvement in data collection and post-processing supports several ongoing projects using the paradigm described in this manuscript. Special thanks to the many individuals who volunteered their time as participants to aid in development of the visuomotor tasks and systems integration.

National Science Foundation (SMA-1514495;NRI-1208623), National Institutes of Health/National Center for Advancing Translational Sciences (KL2-TR001103), UNTHSC seed grant (RI-6114).

References

1. Schultheis MT, Rizzo AA. The application of virtual reality in rehabilitation. *Rehabilitation Psychology*. 2001; 46(3):296–311.
2. Holden MK. Virtual environments for motor rehabilitation: Review. *CyberPsychology and Behavior*. 2005; 8(3):187–211. [PubMed: 15971970]
3. Miller HL, Bugnariu N. Level of immersion impacts the effectiveness of virtual environments used to assess or teach social skills in Autism Spectrum Disorder. *Cyberpsychology, Behavior, and Social Networking*. 2016; 19(4):246–256.
4. Sveistrup H. Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation*. 2004; 1(10)
5. Hayhoe M, Ballard D. Eye movements in natural behavior. *Trends in Cognitive Sciences*. 2005; 9(4):188–194. [PubMed: 15808501]
6. Ajrezo L, Wiener-Vacher S, Bucci MP. Saccades improve postural control: A developmental study in normal children. *PLoS One*. 2013; i(11):e81066.doi: 10.1371/journal.pone.0081066
7. Saunders JA, Knill DC. Humans use continuous visual feedback from the hand to control reaching movements. *Experimental Brain Research*. 2003; 152(3):341–352. [PubMed: 12904935]
8. Lisberger SG, Morris EJ, Tychsen L. Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Reviews of Neuroscience*. 1987; 10:97–129.
9. Wijayasinghe IB, Miller HL, Das SK, Bugnariu NL, Popa DO. Human-like object tracking and gaze estimation with PKD android. *Proceedings of SPIE*. 2015; 9859:98906.

10. Treue S, Maunsell JH. Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience*. 1999; 19(17):7591–7602. [PubMed: 10460265]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.

Diagram of the instrumented lab space, including placement of 12 infrared cameras (red), the instrumented treadmill (black), the screen (light gray), a participant facing the screen, and the control console.

Fig. 2. Child performing a task in the V-Gait CAREN system during recording of kinematic and oculomotor data from force plates, markers, and eye-tracking glasses.

Fig. 3. ETG mobile eye-tracking glasses instrumented with reflective markers; two eye-recording cameras embedded in the bottom of the frames and one scene-recording camera embedded in the front of the

Fig. 4.

Array of objects in the virtual environment used for calibration and calibration verification. Calibration is performed in a triangulated fixation procedure using the top left/right and bottom center objects.

Fig. 5.

Alignment of object vector from eye (red line) and gaze vector (blue line; large spikes indicate off-task looking away from target) in the X (a), Y (b), and Z (c) directions for the Model 2 transformation.

Fig. 6.
Array of marker data and gaze vector data integrated and visualized in custom analysis software.

Fig. 7. Screen capture of a sample trial trajectory of the Angled Pursuit task at the start and end point (top center to bottom center).

Fig. 8.

Screen capture of a sample trial of the Safe Zones task at the start point; blue ball is user-controlled object and moves in the medial-lateral (X) direction based on the position of the C7 marker, green box is static.

Fig. 9. Screen capture of a sample trial trajectory of the Intercept task at the start and end point (top left to bottom center); the blue ball is the participant's user-controlled object, and the red ball is the target object.