Assessing Hazard Identification in Surface Stone Mines in a Virtual Environment

Jennica L. Bellanca, Timothy J. Orr, William Helfrich, Brendan Macdonald, Jason Navoyski and Brianna Eiter

Abstract Mine workers are expected to remain vigilant and successfully identify and mitigate hazards in both routine and non-routine locations. The goal of the current research project is to better understand how workers search and identify hazards. NIOSH researchers developed a data collection setup to measure a subject's gaze, head position, and reaction time while examining 360° 2D-panoramic images at a surface mine. The data is integrated in semi real-time to determine region of interest (ROI) hit accuracy for hazards within the images. The purpose of this paper is to discuss the development and implementation of the hardware and software. The following aspects of the setup will be explored in the paper: (1) environment selection, (2) image creation, (3) stimulus display, (4) synchronization, (5) gaze mapping, and (6) region of interest (ROI) hit calculation.

Keywords Eye tracking · Hazard recognition · Virtual reality

J.L. Bellanca (✉) · T.J. Orr · W. Helfrich · B. Macdonald · J. Navoyski · B. Eiter
Pittsburgh Mining Research Division, National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Pittsburgh, PA, USA
e-mail: JBellanca@cdc.gov

T.J. Orr
e-mail: TOrr@cdc.gov

W. Helfrich
e-mail: WHelfrich@cdc.gov

B. Macdonald
e-mail: BMacdonald@cdc.gov

J. Navoyski
e-mail: JNavoyski@cdc.gov

B. Eiter
e-mail: BEiter@cdc.gov

V.G. Duffy (ed.), *Advances in Applied Digital Human Modeling and Simulation*, Advances in Intelligent Systems and Computing 481,
DOI 10.1007/978-3-319-41627-4_20

1 Background

The metal/non-metal mining industry saw a spike in fatalities between October 2013 and January 2015. Thirty-seven mine workers were fatally injured in accidents, which is a significant increase relative to the record low number of fatalities in the two years prior—16 fatalities in both 2012 and 2011 [1]. In an effort to address this increase in fatalities, during the summer of 2015 the Mine Safety and Health Administration (MSHA) put forth new guidance on "working place" examinations. The goal of this guidance is to increase attention on mine workers' ability to identify hazards at a mine site [2]. Mining is a major undertaking that involves the use of complex heavy machinery, equipment, and processes, as well as numerous and diverse worker activities that take place in a dynamic, challenging environment [3]. In order to safely perform their duties, mine workers are expected to remain vigilant and successfully identify and mitigate hazards in both routine and non-routine locations. The goal of the current research project at the National Institute for Occupational Safety and Health (NIOSH) Pittsburgh Mining Research Division (PMRD) is to better understand the process of hazard recognition and to identify differences across various groups of participants (e.g., mine workers employed at stone, sand and gravel operations, safety professionals, and students). The long-term goal of this work is to develop training materials to help mine workers recognize and mitigate hazards.

PMRD researchers developed a virtual reality setup to examine participants' ability to search and identify hazards. Researchers developed a set of 360° 2D-panoramic images of locations at a surface stone mine that are displayed while collecting each subject's eye position, body position, and reaction time data. This data collection setup involves stimulus presentation and review in a single visit using two different virtual reality environments. The stimulus setup involves the integration of true-size panoramic images, millisecond accurate synchronization of image timing, motion capture data, scan path data, reaction time data, and high-speed video where subjects are allowed to move freely within the viewing area of the 360° cylindrical display. The gaze, motion capture, and reaction time data is integrated in semi real-time in order to determine region of interest (ROI) hit accuracy for researcher-identified hazards within the images. The review setup displays the fixation and hazard identification data to the subjects immediately following the panoramic data collection on a powerwall screen during a researcher-guided review session. The purpose of this paper is to discuss the development and implementation of the hardware and software. The following aspects of the setup will be explored in later sections of the paper: (1) environment selection, (2) image creation, (3) stimulus display, (4) synchronization, (5) gaze mapping, and (6) region of interest (ROI) hit calculation.

2 Simulation Environments

2.1 Stimulus Environment

Because of the impracticality of staging hazards and collecting gaze and reaction time data in the field, researchers were required to turn to the development of a virtual environment. It has been shown, especially for complex search tasks, that environments that are more realistic produce improved training transfer and performance results [4]. However, the cost of environment development and the end goal of developing distributable training materials made a full stereoscopic 3D environment impractical. This study was explicitly interested in hazard search and identification in surface stone facilitates. While these operations are very dynamic, the large scale of the locations and equipment generally requires the hazard identification to be completed at a distance to avoid entering into a dangerous location. Furthermore, a large room diameter would give the participant a more open experience that more closely mimics actual working place examinations.

From a technical perspective, stereoscopic stimulus presentation has been shown to provide significant improvements in only close task performance [5]. A static, monoscopic scene is sufficient because the large distance scale reduces the necessity of scene movement and perspective changes. Projection on larger screens offers a wider field of view that may increase peripheral awareness; this, while often linked to movement, allows us to understand scene context and is an important aspect of immersive environments that can play a role in search tasks [5, 6]. Furthermore, large screens have been shown to be just as effective as head-mounted displays and do not incur the technological restrictions [7]. Nevertheless, the importance of realism with respect to the task complexity and transfer still plays a key role [8]. Higher-resolution imagery has also been shown to increase search task performance; for this reason, high fidelity images were selected [9]. Finally, panoramic images are more deployable as they can be easily integrated into possible training products without extensive software development or hardware requirements. These factors make a large display of static 360° panorama images an appropriate choice.

The stimulus environment consists of a 360° panoramic theater roughly 10 m in diameter with a screen height of three meters. The screen is a polyvinyl material stretched over a steel tubular frame creating a parabolic toroid, where the mid-screen distortion is about 14 cm pulling inward toward the center of the space. The screen is coated with a silver, high-gain, retroreflective material to maintain polarization of incident light so that it can be used for passive stereoscopic viewing. Imagery is front-projected onto the surface from six 1920 × 1080 pixel projectors (Titian 1080p 3D, Digital Projection, Kennesaw, GA) mounted above the screen. The projected images are configured with a 10 % overlap. The setup also includes an array of 10 motion capture cameras (T20, Vicon, Oxford, UK) to provide real-time tracking and precise motion capture capabilities. Eye tracking is accomplished using SensoMotoric Instruments' (SMI) Eye Tracking Glasses (ETG) 2.0 (Teltow, Germany). Scenery for the system is rendered from a single workstation using in-house applications developed with Unity (Unity Technologies, San Francisco, CA).

2.2 Review Environment

The review environment is far less critical to the results of the study. However, maximizing participant engagement and image visibility is preferred. For this reason, a large powerwall setup was used that afforded a large high-resolution screen. The powerwall is a curved panoramic screen along one wall of a 10 × 10 meter room. The screen is approximately eight meters wide by three meters tall with a 50° curvature. The screen construction is similar to that of the stimulus environment where the screen is a coated silver, high-gain, retroreflective polyvinyl material stretched over a steel tubular frame. Imagery is front-projected by three Titan SX+ 3D projectors (Digital Projection, Kennesaw, GA) to provide a seamlessly blended 3182 × 1050 pixel image.

3 Panoramic Images

3.1 Image Acquisition

The panoramic images were captured at four locations at a typical surface stone operation: pit, plant, roadway, and shop. The images included both physically staged and digitally edited hazards. In order to accommodate the stimulus virtual environment and ensure as immersive an experience as possible, image composition and environmental conditions needed to be considered.

To ensure that images looked and felt realistic to the participants, the camera height, orientation, and position needed to be selected appropriately. First, the camera height was set to a reasonable standing height in order to give the participant a realistic perspective of the scene. The approximate eye height of the 50th percentile male—64.5 in.—was chosen in order to minimize the overall error of all the participants [10]. Since the stimulus environment had a level floor, it was also important to ensure that the camera was level on the tripod to maintain correct perspective. Similarly, in order to guarantee that near-field objects appeared to be the correct size in the image, they needed to be at least five meters away from the camera. Objects closer than that would be perceived as too large, and they would break the immersion; this is because the screen was five meters away from the participant in the stimulus environment. The 10-m diameter screen provides subjects with an open feeling, but limits the camera placement for more enclosed locations such as the shop. Given that this requirement was not possible to meet in all instances, the perspective distortion was minimized by positioning the camera close to only large plain surfaces such as a wall or cabinet while maintaining the desired distance from all other salient objects and hazards within a scene. To ensure that the hazards would be present in the final cropped image, the camera was positioned such that the objects required to be in the scene were within $\pm 15°$ of the camera vertically. Finally, to verify the accuracy of the projected panoramic images, the location and size of key objects were recorded. A laser rangefinder (Elite 1 Mile Arc, Bushnell, Overland Park, KS) was used to measure distance and inclination of objects. Inclination measures were used to ensure that objects were scaled properly in the vertical dimension and were located at the same inclination as represented in the actual setting. Object size was measured with a Digital Laser Distance Measurer (DLR165, Robert Bosch LLC, Farmington Hills, MI).

Changing environmental conditions and lighting were additional challenges to the realism and image quality of the panoramas. Flat, overcast lighting conditions would have been ideal to ensure even, well-lit scenes, but due to mine site availability, this was not always possible. The camera settings were optimized to minimize these effects as much as possible. The panoramic images were captured using a Nikon D3X (Tokyo, Japan) camera in conjunction with a GigaPan Epic Pro (GigaPan Systems, Portland, OR) mount. The GigaPan was configured to capture 180° vertically and 360° horizontally with an overlap of 30 %. With the fixed focal length of 14 mm, this amounted to 4×5 images, respectively, for each location. The aperture was set at f 5.6 to give sufficient depth of field for all images. The ISO was manually set for each location, where the ISO values ranged from 400 to 1600. The shutter speed was also manually adjusted depending on location and lighting, where inside shutter speeds were between 1/40 and 1/160 of a second and outdoor speeds ranged from 1/160 of a second to 1/3200. Shutter speeds in all conditions were fast enough to prevent workplace vibrations from interfering with the image quality. The camera was also configured for an exposure bracketing of three, where three exposures were taken for each image at ± one exposure value step.

3.2 Image Editing

To create the panoramic images, the corresponding 20 images for each shot were imported into PTGui (New House Internet Services B.V., Rotterdam, The Netherlands), a panoramic stitching software. PTGui automatically assigns control points to align adjacent images and blend them into a seamless panorama. If the software encounters problems in matching adjacent features, these require user intervention was necessary. Some images required manual alignment using the control points due to a lack of unique detail. Objects in motion (i.e. people, vehicles, moving equipment, clouds) caused most issues. In these instances, masking was used to delete the objects in motion so that they only appeared in one of the shots. The camera lens setting in the software was set at normal lens (rectilinear) and the final stitched panoramas were saved using equirectangular projection.

Adobe Photoshop (Mountain View, CA) was used for all the post-processing of the panoramas. Any additional lighting, contrast, and color issues were resolved, including the removal of image artifacts left by the camera such as lens flare. To keep the mine site anonymous, identifiers such as company signs and logos were removed from the panoramas. Although many hazards were staged and set up on the spot during the initial photo taking, several of the hazards could not be created at the site due to time, safety concerns, or other constraints. Therefore, hazards were digitally added into the panoramas during the editing process by additions or subtractions.

4 Stimulus Display

4.1 Calibration Setup

Spatial references were created for the display and motion tracking systems by creating a removable calibration plate at the center of the stimulus environment (see Fig. 4) and placing fixed reflective markers at various horizontal and vertical viewing angles with respect to the calibration plate origin. The calibration plate provided a repeatable, fixed origin and axes for calibrating the motion tracking system and aligning the projected images. The plate has three short pins that slip into holes drilled into the concrete floor. The plate was designed to catch the flanges on the standard Vicon active calibration wand to set the origin of the motion tracking system. A self-leveling laser level (HVL 100, Pacific Laser Systems, San Rafael, CA) with a 360° horizontal beam, two vertical beams at 90° angles, and a vertical plum beam was used to align the projection images with the calibration plate and fixed reflective markers. The fixed reflective markers consisted of a reflective sticker that was masked to the width of the laser beam to allow quick alignment of the laser level; these points were surveyed in using a Topcon QS-3R Robotic Total Station (Tokyo, Japan). The markers were placed above the screen and behind the screen door to provide absolute reference points every 15° for the vertical beams and every 5° for the horizontal, respectively. The laser level was set on a tripod with the plum beam centered on the calibration plate origin and aligned with the fixed markers. A texture map of a grid pattern with a 1° horizontal and vertical spacing was created to be used as a guide for aligning the projected images with the laser level beams during the image warping and blending calibration.

Image warping and blending was accomplished via a 3D Perception (3DP) display processor using Compact Designer (Orlando, FL). A 7 × 7 array of control points for each image channel was used to adjust the location of projected images on screen. Because the system does not have an array of alignment sensors, adjustment to the warp and blend calibration must be done manually. In order to ensure the best calibration was achieved in the central area of the screen, the laser was first set to eye height and moved out from there. The 3DP display processor controlled the blend regions of all six projectors as well. The pixel alignment of the overlap and gamma alpha adjustment was changed for each blend region in order to achieve a seamless blend between projection images. This process was not only manual, but subjective as well, since no good metric was available for blend control. The circular shape and retroreflective coating on the screens made blending the projectors especially difficult because of the directional dependence. One must be sure to manipulate the blend settings from the viewing area. Due to ambient vibration and system drift despite the projectors being mounted on an independent structure, the system required daily calibration prior to use. In the development of this routine, it was also discovered that the projectors suffer from thermal drift and needed to warm up for at least an hour prior to image calibration.

To ensure proper color balance between projectors, a spectroradiometer (SpectraScan 650, Photo Research Inc., Chatsworth, CA) was used to measure color values for each projector. Measured X and Y values of red, green, and blue were used to calculate target RGB values used to calculate target RGB reference values based on an algorithm supplied by the projector manufacturer, Digital Projection. The reference values were directly used by the Titan projector processor to better match the color values projected on screen. Additional color balance was also subjectively manipulated using color and gamma adjustment layers via 3DP's Compact Designer. Color manipulation via Compact Designer was preferred over direct projector settings because it provides real-time feedback to the user as adjustments are being made.

4.2 Screen Survey

Following the development of the calibration procedure, a topographical survey of the stimulus environment's screen was performed using the Topcon Robotic Total Station. A total of 648 survey points were collected on the screen using a grid pattern, measuring every five horizontal and vertical degrees visible on the screen and along the edge of the buffer regions above and below the panoramic images (Fig. 1). The survey allowed for a validation of the calibration routine, such that the pixel error was no greater than five pixels in any location and was lower than two pixels on average.

4.3 Stimulus Cropping and Alignment

In order to generate seamless stimuli that perfectly match the visual field that the participant would have seen if standing in the original panorama location, the

Fig. 1 Photo depicting the wrap point of the stimulus display. As shown on the image, the panorama coordinate space begins at the lower left of the panorama, with the x (*horizontal*) and y (*vertical*) axes running from (0,0) at the bottom left to (1,1) at the top right. Also shown is the buffer region above and below the image, which exists to ensure the panoramic image is uniformly level on the top and bottom, despite any small variations in screen height and projection region

projected image needed to precisely match both the relative bearing and inclination where the image horizon line matches the participant's eye height. Because the top and bottom framing of the images could not be reliably controlled when exported from PTGui, additional panoramic test shots were collected at each location including three marker staffs, as depicted in Fig. 2, placed around the scene. Eye height was determined by matching background features at the same y-pixel value with the red tape mark. The marker staffs also contained black tape marks every two feet that were similarly used to verify object scale and distortion.

Given the located eye height position, the equirectangular source panoramic images acquired from PTGui were then down-sampled to 12288 pixels horizontally, while maintaining the aspect ratio. The resolution was chosen to exceed the display system's resolution, preventing aliasing, and to keep the texture within typical hardware limits. The vertical crop was positioned to align the image eye

Fig. 2 A section of a test shot containing a marker staff with eye height marked in *red* and every two feet marked with *black tape*. Features on the same pixel level (*white dashed arrow*) allowed researchers to vertically align images to these landmarks in the final image

height 64.5 in. above the floor. Because the floor of the theater was not perfectly level, all screen height measures were referenced to the eye height horizon. A black mask was added to the top and bottom of the images near the limits of the vertical projectable area. These masks precisely control the top and bottom of the stimulus image relative to other screen measures, and conceal blend artifacts at the corners of the projector overlaps (Fig. 2). Given the surveyed grid pattern with eye height correctly located, it was determined that the top and bottom masks should be positioned at 17.5° and −16.7° from eye height, respectively. The resulting resolution of the panoramic image was then calculated to be 12288 × 1160, with a masked resolution of 12288 × 2048 to ensure full coverage.

4.4 Stimulus Rendering

Presenting stimuli within the stimulus environment requires more than simple image rendering, despite the panoramas being two-dimensional. The six-projector cylindrical display requires the image to be mapped onto an ideal cylinder and rendered for each projector. The rendered images then were passed to the 3DP display processor in order to map the image to the actual calibrated screen geometry. To create the image for each projector, in Unity, six cameras were generated and positioned inside the ideal cylinder to match the overlap of the projectors. The cropped and masked image was split into six 2048 × 2048 images that were then mapped on the ideal cylinder.

5 Synchronization

The data acquisition (DAQ) setup for the current study involved communication and synchronization between various systems. In this case, an in-house developed LabVIEW program was used to facilitate the study progression and synchronization of all data elements, including wired and wireless DAQ devices, optical motion tracking, video, eye tracking, and the stimulus display. The DAQ devices and motion tracking system had built in synchronization mechanisms that were handled via hardware triggers.

Because of the unpredictable drift of oscillator clocks used in computers, a simple one-time sync was not sufficient [11]. Clock drift was of particular importance in this study because oscillator clocks are notoriously sensitive to temperature changes such as that experienced by the eye tracking laptop used in this study, as it was semi-enclosed in a backpack [12]. Therefore, LabVIEW and the eye tracking laptop (X230, Lenovo, Morrisville, NC) were synchronized by calculating the clock offset before and after each two-minute stimulus trial. A simple two-way message paradigm was used that required 25 samples with 150 ms between each sample,

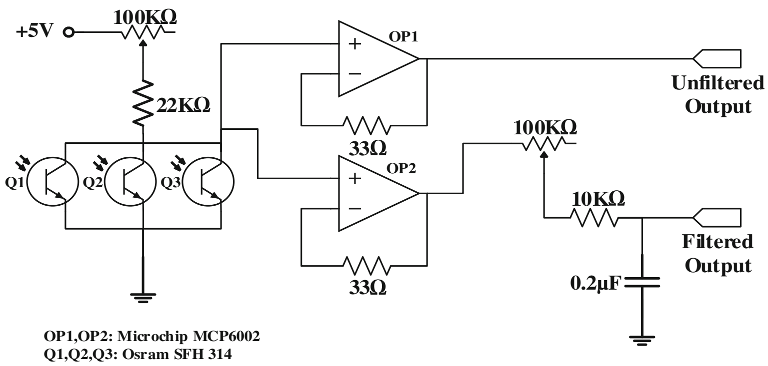

Fig. 3 This schematic drawing shows the circuit used in the photosensor. The circuit has an array of three phototransistors that, when exposed to light, pull the voltage at the collector towards ground. The signal is buffered by two op-amps, the output of one is then filtered through a low pass RC filter, and both outputs are recorded by an analog DAQ

discarding samples with a round trip time over seven milliseconds in order to eliminate any bad packets or network errors.

The display hardware, the rendering software, and the triggering mechanism used to display the panorama all have inherent latency. These fixed and variable latencies had to be accounted for to accurately determine when and how long a scene was displayed. To accomplish this, a photosensor device (Fig. 3) was mounted in the path of the light from the projector, and the output was wired to an analog DAQ (NI USB 6210, National Instruments, Austin, TX). The photosensor device consisted of several phototransistors biased with resistors to match the lighting conditions in the room and the light produced by the projector, such that when the stimulus was displayed, the voltage supplied to the DAQ would go low. Due to the digital light processing (DLP) projectors, the light produced was not continuous but rather pulsed, and two approaches were used to account for this. The output from the phototransistors was first buffered through two op-amps in a voltage-follower configuration, and one of the two outputs was then filtered through a resistor-capacitor low pass filter. The other output was sent directly to the DAQ, and both samples were recorded during the study. This allowed for the use of the analog filtered value directly, or analyzing the raw, photosensor data digitally to look for changes in peak value over a window.

To ensure that a good quality signal was supplied to the photosensor, it had to be mounted directly in the path of light from the projector to the screen, and the image projected on that portion of the projection needed to be of sufficient intensity to distinguish it from the background noise and ambient light between trials. To accomplish this, the photosensor device was mounted in the overprojection region of the projectors above the screen. Rather than fully mask this area as would normally take place, a white section of image was aligned with the position of the photosensor such that whenever time-critical stimuli were displayed, the photosensor would experience a strong transition from black to full white, thus encoding the stimuli start and stop.

6 Gaze Mapping

In order to combine the head position and eye tracking data, several systems needed to be synchronized and calibrated (Fig. 1 and Fig. 4). For future analysis, all systems had their clocks synchronized at multiple points during the study in post-processing by the methods described above. However, final gaze data in panorama space was needed for the debrief process immediately after stimulus data collection. Therefore, a simple real-time collection process was used for the first pass. Data from the eye tracking glasses was streamed over the network to the control computer that was recording the motion tracking and reaction time data, and everything was written out to a single file as it was received. While this did not account for differing latency between the eye and motion tracking systems, the fact that the latency was less than the fixation duration (75 ms) made it sufficient for debrief purposes.

In order to transform gaze data into panorama space, a static calibration was needed to create a reference between the two local coordinate systems. Functionally, this process was completed in two steps because the eye data and the motion tracking data were housed on separate hardware systems. Generally, the calibration accounted for both the orientation of the glasses relative to the participant's face ($\boldsymbol{q}_{etgCal}$ and $\boldsymbol{q}_{fwdCal}$) and the orientation of the glasses rigid body in motion tracking space ($\boldsymbol{q}_{mtCal}$). The static calibration was performed as the subject looked directly at a fixation + on the screen. Because the rotation from ETG space (physical coordinate system of the glasses) to local glasses space (local motion capture coordinate system) were separate during the static calibration, an arbitrary

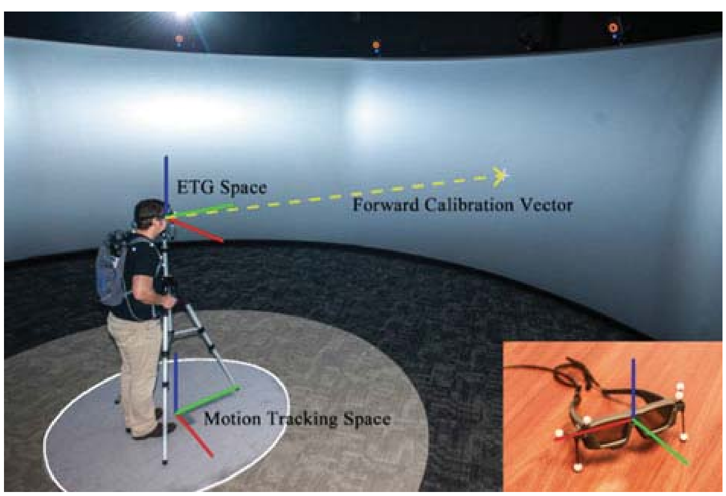

Fig. 4 Photo showing the calibration experimental setup. The participant stood in a fixed location at the center of the room with his or her head stabilized by a chin rest, while fixating on a + marker displayed on the screen. The forward calibration vector, used for calibration, is also shown (*dashed yellow line)*. A close-up of the eye tracking glasses is shown (*lower right*). The *light gray*, carpeted area (*highlighted in white*) shows the movement area in which the participant was allowed to move freely

unit-y vector was chosen as a temporary calculation point. The two rotations were then defined as $\boldsymbol{q}_{etgCal}$, the rotation of the averaged ETG gaze vector during calibration to unit-y, and $\boldsymbol{q}_{fwdCal}$, the rotation of the unit-y to the normalized forward calibration vector ($\hat{\boldsymbol{v}}_f$), defined in Eq. 1 below.

$$\hat{\boldsymbol{v}}_f = \left\| \boldsymbol{p}_{fix} - \bar{\boldsymbol{p}}_{mtCal} \right\|. \tag{1}$$

where $\boldsymbol{p}_{fix}$ is the static position of the fixation cross in global motion tracking space, and $\bar{\boldsymbol{p}}_{mtCal}$ is the average position of the origin of the eye tracking glasses in motion capture space during the static calibration. The rotation of the local glasses coordinate system in motion tracking space ($\boldsymbol{q}_{mtCal}$) was calculated as the average rotation of the rigid body during the static calibration. Figure 4 depicts these coordinate systems for clarity. The gaze vector in motion tracking space ($\boldsymbol{v}_{gaze}$) was computed for every ETG vector ($\boldsymbol{v}_{etg}$) during the trial as expressed in Eq. 2, where the correct vector was lastly rotated by its current orientation in motion tracking space ($\boldsymbol{q}_{mt}$):

$$\boldsymbol{v}_{gaze} = (\boldsymbol{q}_{mt}\boldsymbol{q}_{mtCal})\left(\boldsymbol{q}_{fwdCal}\left(\boldsymbol{q}_{etgCal}\boldsymbol{v}_{etg}\boldsymbol{q}_{etgCal}^{-1}\right)\boldsymbol{q}_{fwdCal}^{-1}\right)(\boldsymbol{q}_{mt}\boldsymbol{q}_{mtCal})^{-1} \tag{2}$$

The next step in the process was to compute where on the screen the subject was looking. This was made more complex by the fact that the screen on which the image was projected was not a flat surface or a perfect cylinder, but an irregular parabolic toroid that flattens out as it approaches the door. Due to the irregular nature of the screen shape, a simple equation did not prove sufficient in mapping the projection surface. Instead, a mesh was created using the topographic survey points (see Sect. 4.2 Screen Survey); this allowed for a raycast to be performed to compute the screen position. The mesh was extended vertically beyond the survey points so that data would still be available when the subject was looking above or below the screen. A simple ray/triangle intersection test was then done using the origin and ray from the previous step to compute, in motion tracking space, where on the screen the subject was looking.

Given the ray intersection position in motion tracking space on the screen where the subject was looking, one final computation needed to be done to determine the panorama position in the panorama space (Fig. 1). This was made somewhat simpler by the fact that the projection system was calibrated to produce a near-perfect cylindrical image from the center of the room at eye height. Given a point on the screen, the vector from the center of the room to that point could be used to determine what portion of the image was projected onto that point by treating the screen as if it was a perfect cylinder. For the horizontal axis, this was the rotation of the gaze vector around the center of the room offset by the location of the origin of the panorama space with respect to the motion tracking space. The rotation was then normalized to [0, 1]. For the vertical axis, only the vertical height of the physical point was needed, which linearly maps to the vertical position in the panorama normalized to [0, 1]. Values below zero and above one are still allowed, but these values indicate a position outside the panoramic image.

7 Region of Interest Hit Calculation

Immediately following the stimulus data collection, the first pass at ROI hits were calculated to facilitate discussion about the hazards within the stimuli images. The debrief software uses the panorama position to compute where the subject fixated during each trial. The fixations were computed according to SMI's dispersion algorithm [13]. Once the fixations were determined, button press times were checked and the last fixation before a button press was used to determine if a hazard was hit or missed. In the interest of both processing time and avoiding false negatives, the hit/miss detection was relatively generous. Three tests were used to determine a hit, if any of them passed, the button press was considered to have identified the hazard. If the button press fixation was within the hit area of two ROIs, both hazards were considered hits. The three tests used to check for a hit were a simple bounds check, distance to vertices, and distance to line segments. All of the hazard ROIs were specified by polygons with a minimum of three vertices. The bounding box of these vertices was first computed and if the button press was within the bounding box, it was considered a hit. The second test found the smallest distance to any of the vertices that made up the polygon. If the distance was below a threshold value (in pixels), it was considered a hit. Finally, the perpendicular distance to each of the line segments was computed and the minimum was compared to the same threshold as the vertex test. If it was below the threshold, the distance was considered a hit. A threshold distance of 150 pixels was used for all hit calculations.

Acknowledgments NIOSH would like to thank Holly Tonini for her help in taking and editing the panoramic images.

Disclaimer The findings and conclusions are those of the authors and do not necessarily represent the views of NIOSH.

References

1. Mine Safety and Health Administration. United States Department of Labor. http://arlweb.msha.gov/stats/charts/mnm-eoy-2015.asp
2. United States. Department of Labor: Mine Safety and Health Administration. Program Policy Letter No. P15-IV-01. By Neal H. Merrifield. (2015)
3. Scharf, T., Vaught, C., Kidd, P., Steiner, L., Kowalski, K., Wiehagen, B., Rethi, L., Cole, H.: Toward a typology of dynamic and hazardous work environments. Hum. Ecol. Risk. Assess. Int. J. **7**(7), 1827–1842 (2001)
4. Stinson, C., Kopper, R., Scerbo, B., Ragan, E., Bowman, D.: The effects of visual realism on training transfer in immersive virtual environments. In: Proceedings of Human Systems Integration Symposium (2011)

5. Bowman, D.A., McMahan, R.P.: Virtual reality: how much immersion is enough? Computer **40**(7), 36–43 (2007)
6. Gallop, S.: Peripheral visual awareness: the central issue. J. Behav. Optom. **7**(6), 151–155 (1996)
7. Patrick, E., Cosgrove, D., Slavkovic, A., Rode, J.A., Verratti, T., Chiselko, G.: Using a large projection screen as an alternative to head-mounted displays for virtual environments. In: SIGCHI Conference on Human Factors in Computing Systems, pp. 478–485. ACM (2000)
8. Lee, C., Rincon, G.A., Meyer, G., Hollerer, T., Bowman, D.A.: The effects of visual realism on search tasks in mixed reality simulation. IEEE. Trans. Visual Comput. Graph. **19**(4), 547–556 (2013)
9. Ni, T., Bowman, D., Chen, J.: Increased display size and resolution improve task performance in information-rich virtual environments. In: Proceedings of the 2006 Conference on Graphics Interface, pp. 139–146. Canadian Information Processing Society (2006)
10. Gordon, C.C., Blackwell, C.L., Bradtmiller, B., Parham, J.L., Hotzman, J., Paquette, S.P., Corner, B.D., Hodge, B.M.: Anthropometric Survey of US Marine Corps Personnel: Methods and Summary Statistics. No. TR-13/018. Army Natick Soldier Research Development and Engineering Center, Natick, MA (2013)
11. Sivrikaya, F., Yener, B.: Time synchronization in sensor networks: a survey. IEEE Netw. **18**(4), 45–50 (2004)
12. Marouani, H., Dagenais, M.R.: Internal clock drift estimation in computer clusters. J Comput. Syst. Netw. Commun. **2008**, 1–7 (2008)
13. SensoMotoric Instruments. BeGaze Manual. Version 3.5. (2015)