24-090

Using Augmented Reality Assistance in Mine Rescue

Doga Cagdas Demirkan
Mechanical Engineering, Colorado School of Mines, Golden, CO

Ava Segal
Mechanical Engineering, Colorado School of Mines, Golden, CO

Abhidipta Mallik
Mechanical Engineering, Colorado School of Mines, Golden, CO

Sebnem Duzgun
Mining Engineering, Colorado School of Mines, Golden, CO

Andrew J Petruska
Mechanical Engineering, Colorado School of Mines, Golden, CO

ABSTRACT

Search and rescue operations in mining heavily rely on humans. In the context of emergency situations within underground tunnels, the potential to utilize sensory technology for navigation assistance holds the promise of saving human lives. Our research introduces a device configuration designed to enhance the performance of search and rescue teams. This setup integrates LIDAR and thermal cameras with the Microsoft Hololens, facilitated by the NUC computer and ROS Bridge. Performance assessment involves quantifying the extent to which a user is willing to venture into a light-deprived section of the mine prior to necessitating supplementary illumination. The preliminary outcomes highlight the significant potential of augmented reality assistance in offering indispensable aid to first responders, thereby greatly augmenting the efficiency of search and rescue missions conducted during critical circumstances.

INTRODUCTION

=In the aftermath of natural or human-caused disasters, saving lives becomes an urgent and time- sensitive task[1]. Search and rescue teams face environmental obstacles while racing against time [2]. Underground spaces like tunnels, mines, or caves pose challenges due to no GPS, low visibility, and uneven terrain [3]. Advanced indoor navigation solutions are required due to the challenging nature of underground mines, where devastating disasters often occur [4]. In an emergency, mine rescue personnel face limited visibility due to dust and smoke, often reducing their visual perception to 1–3 feet [5].

Traditional emergency guidance in underground mines includes hand lines, pinwheels, and strobe lights [6]. Using green laser beams can provide situational awareness in smoky environments by penetrating the smoke [5]. Thermal imaging is another useful approach for smoke-occluded situations, such as those encountered by firefighters and military personnel. This technology can penetrate dust and smoke to provide clear images [7]. Nevertheless, restoring visual information to first responders in emergency situations will enable them to search the environment more safely and efficiently, as they will be able to scan an area for hazards and will provide personnel with better situational awareness [8].

Given the need for visual awareness in underground search and rescue missions, this work focuses on developing a hybrid human-machine system for solving situation awareness problems in pitch-black and smoke-filled underground mines.

In our previous study [9], we proposed a hardware setup and explained its integration with utilized software. The proposed methodology is to combine thermal imaging, LiDAR, and AR, which aims to penetrate the smoke by leveraging thermal imaging while having real-time world construction with LiDAR and visualizing it on an AR device. By having this hybrid system, we aimed to provide an AR image of the surrounding environment when pitch-black or visually occluded by smoke, dust, or other small particulates. This system was successfully built around an AR device, namely, Microsoft HoloLens, which can display a real-time image to the user that updates at 5 hz. However, the proposed hardware was never tested in an underground mine environment as a scenario-based solution, such as underground tunnel evacuation, wayfinding and recognition, or search and rescue missions.

As a continuation of our study that aims to fill the necessity of having a vision in pitch-black conditions, we took the prototype into the Edgar Mine at Idoha Springs, Colorado. After finding the optimum hardware and software combination and building a prototype, the last step is to find how to visualize the data stream, as it only comprises numbers and letters. To solve this problem, we ask, "What are the best ways and visual variables for first responders to visualize reconstructed data in augmented reality?" This question hypothesizes that AR interface with multi- colored reconstruction with depth data might help decrease response and assessment time, hence cognitive load. This will result in faster search and rescue operations.

Since emergency response is a time-critical mission, appropriate visualization needs to be employed. Our overarching goal is to investigate and benchmark visual variables for pitch-black augmented reality assistance.

RELATED RESEARCH

Visual variables are fundamental components of data visualization. In 1967, Bertin proposed the first categorization of visual parameters, including size, color, orientation, texture, shape value (brightness), and position (dimensions on the plane) [10][11]. Later, MacEachren suggested three more visual parameters: crispness, resolution, and transparency [12]. In 1994, Wolfe discussed the effects of complexity (referred to as cluttering) in visual search, which increases cognitive load [13]. In 2014, Zhang et al. examined the compatibility and limitations of 2D visual variables in 3D visuals [14][15].

Similar to studies on 2D and 3D compatibility of visual parameters, virtual and augmented reality compatibility also needs to be investigated. Computer Human Interaction (CHI) (one of the most prestigious [16] conferences in human-computer interaction) mentions the most significant gaps and challenges in the field [17]. Designing guidelines for visualization and understanding human senses and cognition in situated contexts are two challenges related to spatially situated data visualization. Improper utilization of visual elements can lead to the misinterpretation of visualization by the users [18], which might cost human lives in underground mining emergencies.

Laha et al. studied the effects of visual variables of head tracking, field of regard, and stereoscopic rendering for volumetric data interpretation. They found that the field of regard has a positive impact on tasks, but other parameters have mixed effects [19]. Adams et al. studied shadow vs. shadowless and position (floating vs. grounded) variables for the depth perception of users in augmented reality. They found that current devices make users underestimate the distance regarding the visual variable but the position of the object influences users' decisions [20]. Arjun et al. evaluated the effects of color, size, orientation, opacity, and shape for chart data understanding. They found that accuracy is affected the most by size and color. Further, cognitive load is affected the most by color, size, opacity and brightness [21].

APPROACH

The resulting system integrates a LiDAR, a thermal camera, and a Microsoft HoloLens onto a wearable platform (e.g., hardhat, belt, or backpack. The processing and power storage are integrated into a waist-belt mounted package with a cap-lamp style cable connecting the two. This enables the user to look around with minimal impact on their motion.

However, which visual variable to use to visualize the sensor data needs to be tested. Therefore, the research team internally tested the initial demo version of visualization varieties. Color is selected as a visual variable due to its pre-attentive attribute. However, failure to develop the correct color might result in increased cognitive load, distraction, and overlooking critical data [22][23][24][25].

The visual variable *color* is tested as in three different settings: i) single color mapping, ii) thermal sensor mapping, iii) depth mapping. These settings are tested for navigation in a tunnel and detecting a person together and seperately.

RESULTS AND DISCUSSION

Initial results yield possible visualization varieties. These visualization options are given in Figure 1 as a) single color mapping, b) thermal mapping, c) depth mapping.

Among these visualization options, depth mapping with a thermal camera yields the most robust results. Color schemes to provide depth perception information are added as a user option, where red schemes indicate closer objects and green schemes indicate farther objects.

The display capabilities of the system have been expanded to include several different shapes that have tradeoffs in terms of both point density and fidelity. Each of these display variants can be generated directly from the sensor data and evaluated as one of the following:

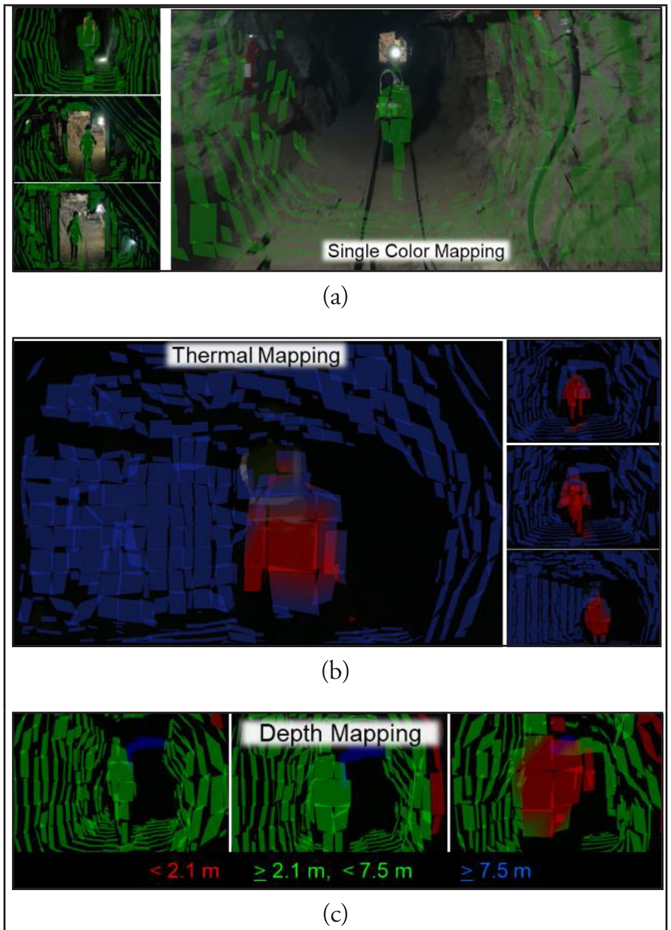


Figure 1. Visualization options of (a)single color mapping, (b) thermal mapping, (c) depth mapping

- Triangulated mesh, where the point-cloud is taken and the gd3 algorithm from the point- cloud library [26] is used to generate a tessellation,
- Square surface patch, where the point cloud is packed into an Octree that tracks a Gaussian representation of the points measured in each voxel. A square patch is then generated that represents the average surface for that voxel,
- Triangle surface patch, where the point cloud is packed into an Octree that tracks a Gaussian representation of the points measured in each voxel. A triangular patch is then generated that represents the average surface for that voxel,
- Cube occupancy grid, where the point cloud is packed into an Octree, and the voxels that contain points are extracted and displayed to the user,
- Cuboid surface representation, where the point cloud is packed into an Octree that tracks a 3D Gaussian representation which is converted into a rotated and scaled cuboidal shape,
- Ellipsoid surface representation, where the point cloud is packed into an Octree that tracks a 3D Gaussian representation which is converted into an ellipsoid according to the standard deviations in each direction for display.

For all images in Figure 1 the number of vertices and triangles with 2D performance of each display type is given in Table 1.

After choosing the best performing visualization method in terms of agility and refresh rate, the final run snapshots are given in Figure 2, which shows the user perspective. It should be noted that the user experience in a dark environment is more crisp than the figures can capture.

As a result, users can successfully walk through the dark tunnels, navigate and detect other persons with the given surface representation, depth coloring and thermal mapping.

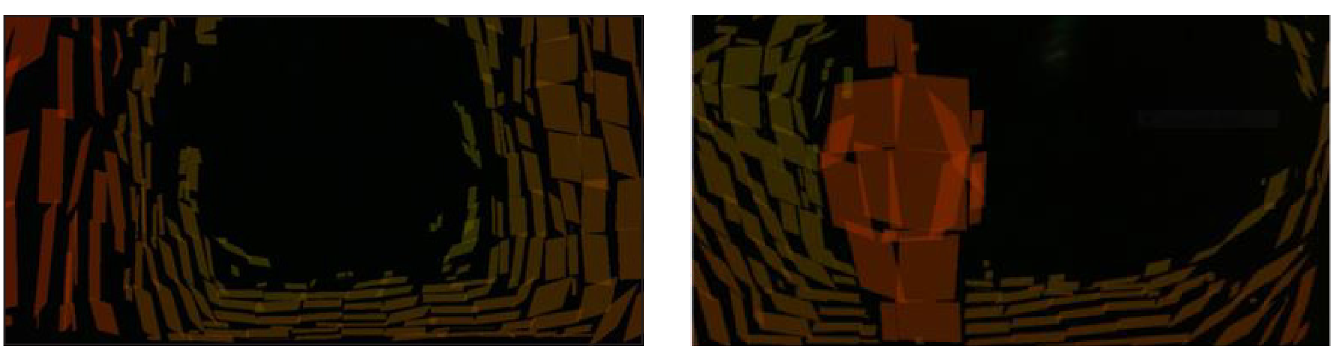
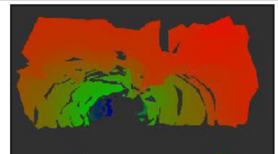
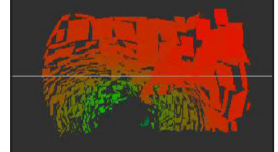
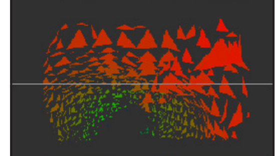
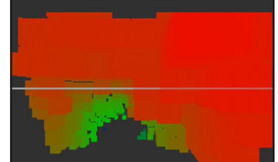
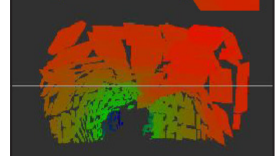
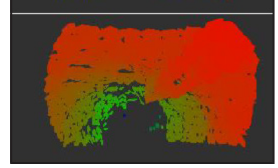


Figure 2. Example snapshot of the enhancement

Table 1: Display types with their respective number of vertices and triangles, and performances

Display Type	Number of Vertices	Number of Triangles	Qualitative 2D performance	Visualization
Triangulated mesh: Triangle tessellation using gd3. Provides a smooth surface but is unpredictable between scans.	1,253	1,671	Poor-Good	
Square Surface Patch Square surface representation requires minimal triangles and vertices and provides a reasonable level of fidelity.	996	1,992	Good-Excellent	
Triangle Surface Patch Triangle surfaces minimizes the number of triangles to send and voxels, but is disconnected and only provides moderate fidelity.	476	1,428	Good	
Cube Occupancy Grid: Cube occupancy. Very consistent but overly conservative and blocks much of the view.	5,712	3,808	Poor	
Cuboid Surface Representation: Cuboid provides a good view of the surroundings but requires a lot of vertices and triangles.	3,070	2,456	Good	
Ellipsoid Surface Representation: Ellipse representation has the best fidelity (note the hanging pipes clearly visible in the upper right) but requires a lot of vertices and triangles.	9,600	5,760	Excellent	

Although it had multiple successful runs, one of the challenges while testing was that the sensors and hard hat had fixed dimensions, and the users' head sizes and eye distances to the sensors were different. This made a slight offset for the visual enhancement interface. Since the search and rescue operations are time-critical, we added a wireless gaming controller for real-time alignment for the visual enhancement stream. The alignment helped users to fit the enhancement in three degrees of freedom.

CONCLUSIONS AND FUTURE WORK

In this work, we combined a thermal imaging camera with a LiDAR sensor and visualized a real- time world construction as an AR interface. The feasibility and progressive nature of the device were tested in an experimental underground mine located in Idaho Springs, Colorado. The results showed that combining the technologies we used enables faster, safer, and more effective disaster response for mine rescue operations. Not only does it allow the responders to search the environment more rapidly, but it also enables them to detect unexpected hazards before they become imminent threats. Moreover, the utility of the developed system is far-reaching, for example, for first responders searching smoke-filed burning structures. In the future, it might also enable autonomous systems to navigate these occluded environments effectively and enable disaster response to focus on the rescue in search-and-rescue.

We are currently designing a future human subject study to test the visual variable *color*. The design is already submitted to the institutional review board of the Colorado School of Mines. The application is approved as an exempt study under 45 CFR 46.104(d)(3) (July 19, 2018).

ACKNOWLEDGMENTS

This study was sponsored by the Alpha Foundation for the Improvement of Mine Safety and Health, Inc. (ALPHA FOUNDATION). The views, opinions and recommendations expressed herein are solely those of the authors and do not imply any endorsement by the ALPHA FOUNDATION, its Directors and staff.

REFERENCES

[1] Chen J, Li S, Liu D, Li X. Airobsim: Simulating a multisensor aerial robot for urban search and rescue operation and training. Sensors (Switzerland). 2020;20(18):1–20.

[2] Beerbower D, Energy P, Biggerstaff R, Coal A, Blackwell WK, Energy C, et al. Mine Rescue Handbook.

[3] Bertrand JWM, Fransoo JC. Modelling and simulation. Research Methods for Operations Management. 2016. 290–330 p.

[4] Zlot R, Bosse M. Efficient Large-Scale 3D Mobile Mapping and Surface Reconstruction of an Underground Mine. In: Yoshida K, Tadokoro S, editors. Field and Service Robotics: Results of the 8th International Conference [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p. 479–93. Available from: doi.org/10.1007/978-3-642-40686-7_32.

[5] Conti RS, Chasko LL, Wiehagen WJ, Lazzara CP. Fire Response Preparedness for Underground Mines. Inf Circ 9481 [Internet]. 2005;25. Available from: www.cdc.gov/niosh/mining/UserFiles/works/pdfs/2006-105.pdf.

[6] Zimroz P, Trybała P, Wróblewski A, Góralczyk M, Szrek J, Wójcik A, et al. Application of UAV in search and rescue actions in underground mine—A specific sound detection in noisy acoustic signal. Energies. 2021;14(13):1–21.

[7] Willette K. First Responder. NFPA J [Internet]. 2018; Available from: www.nfpa.org/News-and-Research/Publications-and-media/NFPA-Journal/2018/January-February-2018/Columns/First-Responder.

[8] Demirkan DC, Duzgun S. An Evaluation of AR-Assisted Navigation for Search and Rescue in Underground Spaces. In: 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR- Adjunct). 2020. p. 1–2.

[9] Demirkan DC, Segal A, Malik A, Duzgun HS, Petruska AJ. Real-time perception enhancement in obscured environments for underground mine search and rescue teams. [Manuscript Submitt Publ. 2023;.

[10] Bertin J. Semiologie graphique. Paris; 1967.

[11] Bertin J. Semiology of Graphics: Diagrams, Networks, Maps. UMI Research Press; 1983. 429 p.

[12] MacEachren AM. Some Truth With Maps: A Primer on Symbolization and Design. Association of American Geographers; 1994. 129 p.

[13] Wolfe JM. Guided Search 2.0 A revised model of visual search. Psychon Bull Rev. 1994;1(2):202–38.

[14] Zhang ZP, Liu J. Research on the symbol vision variable of the three-dimension virtual battle environment. Geomatics Spat Inf Technol. 2014;37(9):7–9.

[15] Hong S, Mao B, Li B. Preliminary Exploration of Three-Dimensional Visual Variables in Virtual Reality. In: 2018 International Conference on Virtual Reality and Visualization (ICVRV). IEEE; 2018. p. 28–34.

[16] Conferance Ranks. No Title [Internet]. Available from: www.conferenceranks.com/index.html? search all=Human+Factors+in+Computing+S ystems#data.

[17] Ens B, Bach B, Cordeil M, Engelke U, Serrano M, Willett W, et al. Grand Challenges in Immersive Analytics. 2021;17:1–17. Available from: hdl.handle.net/1880/112984.

[18] Dasgupta A, Poco J, Wei Y, Cook R, Bertini E, Silva CT. Bridging Theory with Practice: An Exploratory Study of Visualization Use and Design for Climate Model Comparison. IEEE Trans Vis Comput Graph. 2015;21(9):996–1014.

[19] Laha B, Sensharma K, Schiffbauer JD, Bowman DA. Effects of immersion on visual analysis of volume data. IEEE Trans Vis Comput Graph. 2012;18(4):597–606.

[20] Adams H, Stefanucci J, Creem-Regehr S, Bodenheimer B. Depth Perception in Augmented Reality: The Effects of Display, Shadow, and Position. Proc - 2022 IEEE Conf Virtual Real 3D User Interfaces, VR 2022. 2022;792–801.

[21] Arjun S, Reddy GSR, Mukhopadhyay A, Vinod S, Biswas P. Evaluating Visual Variables in a Virtual Reality Environment. 34th Br Hum Comput Interact Conf Interact Conf BCS HCI 2021. 2021;133–8.

[22] Hosseinkhani J, Joslin C. Significance of Bottom-Up Attributes in Video Saliency Detection without Cognitive Bias. In: 2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC). 2018. p. 606–13.

[23] Wang Y, Su H, Zhang B, Hu X. Learning Reliable Visual Saliency For Model Explanations. IEEE Trans Multimed. 2020;22(7):1796–807.

[24] Bruce NDB, Tsotsos JK. Saliency based on information maximization. Adv Neural Inf Process Syst. 2005;155–62.

[25] Duzgun S, Isleyen E, Demirkan D -C., Orsvuran R, Bozdag E, Pugmire D. Virtual and Augmented Reality for Visualization of Big Data: Examples from Deep Earth to Subsurface. In: AGU Fall Meeting Abstracts. 2019. p. IN21B-05.

[26] Rusu RB, Cousins S. 3D is here: Point Cloud Library (PCL). In: IEEE International Conference on Robotics and Automation (ICRA). Shanghai, China; 2011.

24-091

Using Event-Based Imaging and Deep Learning to Generate 3D Surface Maps for Autonomous Roof Bolting

Rik Banerjee and Andrew J. Petruska
M3 Robotics Lab
Colorado School of Mines Golden CO, USA

ABSTRACT

This study explores implementing a machine learning-based system to generate a 3D surface repre- sentation of the roof and support straps in the mine. Event cameras have been chosen for their performance in high-dynamic-range lighting conditions and for their low latency. To enable automated drilling and bolting, 3D vision using event-based cameras has been developed. A ground-truth set is created using two, time-synced event cameras and a LiDAR camera. These sensors are used to construct a ground-truth dataset of corresponding event- camera images and surface maps from the LiDAR. The network is tested with stereo-pairs of event images and produces a depth image with ±5 mm RMS error on average across 1000 test images.

INTRODUCTION

This paper looks to introduce the use of neuro- morphic sensors called event cameras, to perform the 3D perception task [1] of surface representation in the harsh environmental and lighting conditions of underground mines. These cameras offer several advantages over traditional cameras such as low latency, high temporal resolution, and very high dynamic range [2]. This allows for the capture of information in an active mining environment as it is resistant to motion blur due to vibrating sensor platforms, shadows due to single-point source lighting, and dust clouds. However, this sensor is novel enough to not have an associated data set that is both large and diverse [3]. This is an additional hurdle to using commercial vision products as there simply is not enough labeled data to design and train a generalized machine learning model. This problem is usually solved by using simulated data [4]. Nevertheless, simulators frequently lack the representation of visual features commonly present in mining environments, as can be seen in Figure 1. Additionally, the shift from simulation to real- world conditions is recognized to be challenging [5]. The pipeline suggested in this paper serves as an alternative, offering a quick and cost-effective means of generating a labeled data set.

The 3D perception task requires that image points (x, y) in pixels be re-projected into world coordinates (X, Y, Z) in physical distance units. This can be achieved by incorporating depth by using either a depth sensor [6] or a stereo camera pair [7], during operation. To mitigate power and dust concerns associ- ated with active sensors, stereo-vision has been chosen to compute the depth of each pixel.

Stereo vision takes two images taken by cameras that have a known distance offset and outputs a dis- parity map. Disparity is the lateral perspective shift between a pair of corresponding pixels in the left and right images. Since the baseline distance (distance between the cameras) is known, the per-pixel disparity can be projected into depth using epipolar geometry [8] as in Equation 1.

$$\mathrm{d} = \frac{\mathrm{b} \times \mathrm{f}}{\mathrm{z}}$$

where,

d = the per-pixel disparity value,
b =the distance between the stereo cameras,
f = the focal length of the cameras used,
z = the depth of the pixel.

This work is supported by National Institute for Occupational Safety & Health | NIOSH/Project 75D30121C12206.