Accessibility, Assistive Technology and Digital Environments, Vol. 121, 2024, 1–8
https://doi.org/10.54941/ahfe1004608

Symbolism in Extended Reality

Yang Cai

University of California San Diego, La Jolla, CA, USA

ABSTRACT

For centuries, humans used symbols to describe scenes, objects, and events. Some of the symbols have evolved into pictorial languages such as petroglyphs and Chinese. The rapidly growing Extended Reality (XR) technologies enable users to experience immersive photorealistic views with head-mounted displays, 3D projection screens, or holographic display panels. The purpose of Extended Reality (XR) is to make invisible to be visible to the user. In this study, we explore how to superimpose symbolic feedback to the user in forms of geometric shapes, trajectory traces, spectra image, semantic network, and multimodal alerts, including audio and tactile signals. We zoom into the laparoscopic surgery training as a case study for applications. Our preliminary experiment results show that the symbolism approach does not significantly increase the distraction of attention or computational load. We also found that Extended Reality, especially, multimodal alert would be a bridge between manual operations and autonomous operations.

Keywords: Symbolism, Simplicity, Augmented reality, Virtual reality, Overlay, Visualization, Extended reality, Abstraction, Discovery

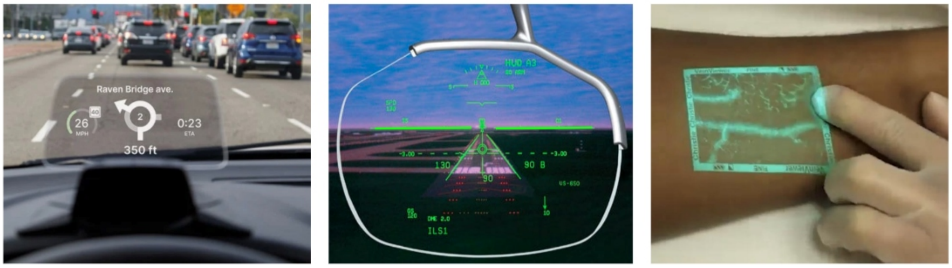
INTRODUCTION

After Pablo Picasso saw the famous ice age cave paintings in Lascaux, France, 1940, Picasso said, "Since Lascaux, we have invented nothing" (Jones, 2021). Those primitive cave paintings fit Picasso's metaphor of figurative abstract art. Two years later, Picasso created Bull's Head from the seat and handlebars of a bicycle and exhibited it at the Salon d'Automne in Paris. The art critic Eric Gibson described the sculpture as "*transparent*" for its *authenticity* and it is a moment of *wit* and *highly sophisticated* in its simplicity (Scott, 2019). Picasso transformed his art from realism to symbolism through his artistic life. He painted the very large mural Guernica with a limited palette of black, grays, and white to portrait the aftermath of the bombing within one contained space, where the forms of humans and animals were distorted and overlapped one another, creating more shapes and narratives within the anti-war symbolic language (Cole, 2021). Picasso emphasized the human factors in symbolism. He said, "it is not up to the painter to define the symbols. The public who look at the picture must interpret the symbols as they understand them."

For centuries, humans used symbols to describe scenes, objects, and events. Some of the symbols have evolved into pictorial languages such as petroglyphs and Chinese characters (Cai, 2016). The rapidly growing Extended

Reality (XR) technologies enable users to experience immersive photorealistic views with head-mounted displays, 3D projection screens, or holographic display panels. For example, some vehicles project the GPS navigation graph to the windshield in real-time. The augmented reality overlays align the map with the landmarks on the road. However, the holographic image may distract the driver's attention or block the frontal vision. Furthermore, mixed reality enables users to combine a physical model with realistic digital overlays. The border between these two worlds is getting blurry. The high-fidelity overlays certainly enhance the visual experience. But, it might add extra artifacts and distractions to the physical scene.

Extended Reality (XR) overlays real-time data to a screen, object, hand, or sound that can be perceived by the user. In our everyday life, there are working models. The ultimate goal of XR is to make the invisible, visible to the user. Figure 1 shows a symbolic navigation map projected to the windshield, the head-up-display (HUD) for the airplane pilot to land on the runway, and the near infrared image of the veins projected to the patient's arm (AccuVein, 2024). There are many applied human factors and ergonomic considerations in these devices: they are visual, symbolic, real-time, holographic overlay, multimodal, and deep in physics such as GPS mapping, aircraft landing angle, and spectrum imaging.



GPS Navigator **HUD for Pilots** **Projected Vein Finder**

Figure 1: Symbolic overlays in augmented reality systems.

The purpose of real-time interaction in the physical world such as driving and surgery, is that we want to explore a new way to apply Extended Reality (AR/VR design) with symbols rather than hyper-realistic 3D virtual objects. In this study, we articulate the laparoscopic surgery procedures and structures with symbols and polygons. For example, Calot Triangle for cholecystectomy surgery and training. This helps the surgeon locate the critical tissues and avoid risky moves within the critical area. We aim to show that the symbolic overlays reduce computing time and distraction, and improve procedural efficiency.

GEOMETRIC OVERLAY

Geometry is perhaps the oldest science that concerns the shape of individual objects, spatial relationships among various objects, and the properties of surrounding space. From ancient times, people have used line, shape, volume, distance, and angle to describe and measure objects. Geometric descriptions, in terms of a line and polygon, appear to be effective in the digital age. Instead of prevailing approaches that overlay high-fidelity 3D objects, which are computationally challenging and expensive, we explore simple 2D polygon overlays such as triangle and rectangular line drawings.

In laparoscopic surgery, especially cholecystectomy, surgeons need to identify the critical vessels around the gallbladder. The small area in the abdomen is called Calot's Triangle, formed by the border of the liver, Cystic duct, and common bile duct. Normally, the cystic artery and right hepatic artery would be inside the triangle. The surgeons have to explore Calot's Triangle to avoid cutting the arteries by mistake. Figure 2 shows the simulated physical model of the digestive cavity where the semantic description "neck" of the gallbladder is visible, and Calot's Triangle in yellow color.

We first developed a physical model of the digestive abdomen with realistic organs for haptic perception. We then simulate the laparoscopic camera motion and develop a machine vision algorithm to detect the Calot's Triangle area marked with a blurry bounding box. The real-time detection and tracking Calot's Triangle enable the trainee to learn the anatomic structure interactively with more realistic visual and haptic feedback. The symbolic overlay on the live laparoscopic surgery screen is consistent with the conventional surgical settings.

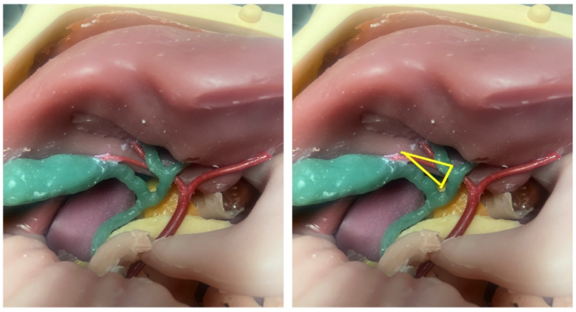


Figure 2: The simulated physical model of the digestive cavity where the gallbladder "neck" is visible (left) and the Calot's triangle in yellow color (right).

Figure 3: The machine vision algorithm detects and tracks the Calot's triangle and overlays a blue bounding box to the area on the live laparoscopic simulation video.

The method can potentially be transformed to in-vivo applications in surgeries. Figure 3 shows the machine vision algorithm detects and tracks the Calot's Triangle and overlays a bounding box to the area on the live laparoscopic simulation video.

TRAJECTORY OVERLAY

Surgical instrument tracking is important for trainees to learn how to manipulate and for instructors to assess the performance. In laparoscopic surgery, the camera control is necessary to follow the surgical instrument such as a grasper and adjust the camera to place the instrument at the center of the screen. Figure 4 shows the blue bounding box at the tip of the instrument. The score model calculates the distance from the center of the screen with the range from 0 (too far from the center) to 10 (perfect). This can be used to train the surgical coordination between the surgery assistant who normally holds the laparoscopic camera and the surgeon who performs the operation.

The trajectory tracking is dynamic. The tracking data is updated in real-time for monitoring and captured for post-procedure assessment. The trajectory can be overlaid to the screen with a bounding box along the target or plotted as a trace over time to show the trajectory patterns and efficiency of the manipulation. Figure 4 shows the real-time updated tracking status of the laparoscopic tip and the accumulated trace of the movements over time.

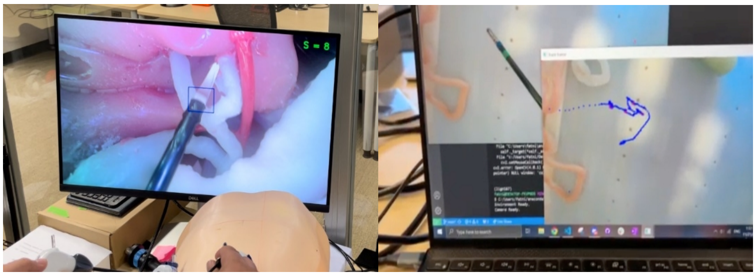


Figure 4: The blue bounding box indicates the tracked laparoscopic tip (left) and the trace of the movements over time (right).

SPECTRA OVERLAY

To make the invisible visible to the user is the ultimate goal of Extended Reality. Most autonomous driving vehicles use the near-infrared camera to detect objects on the road at night. Similarly, modern laparoscopic surgeries often use near-infrared (NIR) light and NIR camera to highlight critical vessels such as the cystic artery near the gallbladder and liver in cholecystectomy. We built a physical model to simulate the NIR view on the live laparoscopic surgery video. We then developed a machine vision algorithm to enhance the physical realism in terms of the color distribution, reflection, intensity, and the interaction between the normal visible light and the near-infrared light. This symbolism design in fact can be applied to many other fields, including but not limited to autonomous driving, medical diagnoses, structure inspection, environmental protection, and energy conservation. Figure 5 shows our simulated Near-Infrared Laparoscopic surgery with tangible digestive organ models in the abdomen cavity. The green color areas simulate the near-infrared reflections on the live video.

Figure 5: Simulated near-infrared laparoscopic surgery with physical organ models.

SEMANTIC NETWORK

A learning process is not just to build muscle memory, but also to develop a mental map through a sequence of decisions, actions, and consequences. A group of sequences of data can form a semantic structure that changes over time. Some knowledge links might disappear; some might stay; new ones could be added to the existing structure, or new ones could break out to a new paradigm.

Similar to other learning processes, laparoscopic surgeons need to build their mental model for complicated surgical procedures, especially for special cases such as abnormal anatomic structure or bleeding incidents. They sometimes have to switch procedures during the operation, e.g. switching from laparoscopic surgery to an open surgery.

We developed a semantic network visualization tool with 3D interactive sequential graph for representing the surgical scenarios in cholecystectomy, including some worst case scenarios such as cutting the wrong vessels twice, failing to switch to open surgery, bleeding to death, etc. The individual steps and decisions are coded with the medical teams as the graph nodes. The connections between the steps or decisions are coded as bidirectional links. We use Force-Directed Graph to spread the nodes evenly to avoid collisions and far apart. And we use arcs to connect the links instead of straight lines to accommodate the bi-directional actions and reasoning.

The interactive 3D sequential graph can be navigated with a keyboard shortcut or a game controller. We found that almost all medical students who experienced this prototype liked the game controller interface because they are familiar with it in video games. However, we want to expand deeper than a semantic network game. We encode related video clips to the specific surgical procedures, for example, insertion of a laparoscopic camera, or removal of the instrument, etc. Figure 6 shows the laparoscopic surgery worst case scenarios in the 3D interactive sequential graph and the close-up view of the steps and decisions. When the user zooms in to a specific step or decision, the user might click the node for more detailed information such as video clips.

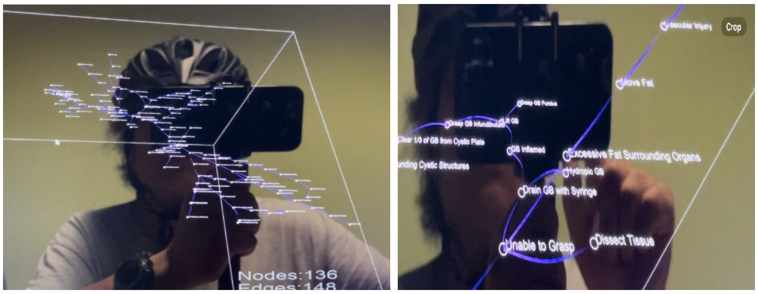


Figure 6: The laparoscopic surgery worst scenarios in 3D interactive sequences.

MULTIMODAL ALERT

Extended Reality uses all possible human perception channels to improve our learning and operation experience. It is a multimodal approach. Many existing systems already use the multimodal extended reality for safety alerts (Polaczyk, Croft, and Cai, 2021). For example, driving safety. The new Nisson vehicles has a symbolic alert sign on the side mirrors to indicate whether a vehicle is on another lane in the driver's blindspot before the driver switches lanes. The new Toyota car sends the audio and vibration signal to the driver when the vehicle is crossing the borderline (off the road). See Figure 7 for the two photos the author recorded from the two vehicles.

In laparoscopic surgery cases, we found multimodal alert is common. For example, when the surgeon heard the pulse rate from the ECG monitor speeds up, the surgeon would realize the patient might be in pain. So the surgeon would talk to the anaesthetist to adjust the dosages. Then the surgeon would listen to the ECG sound again to see any improvement.

Figure 7: The symbolic visual alert of the car in the blindspot (left), and the audio / haptic alert of lane-crossing (right).

CONCLUSION

The purpose of Extended Reality (XR) is to make the invisible visible to the user. In this study, we explore how to superimpose symbolic feedback to the user in forms of geometric shapes, trajectory traces, spectra image, semantic network, and multimodal alerts, including audio and tactile signals. We zoom into the laparoscopic surgery train as a case study for applications. Our preliminary experiment results show that the symbolism approach does not significantly increase the distraction of attention or computational load. We also found that Extended Reality, especially multimodal alerts, could be a bridge between manual operations and autonomous operations.

ACKNOWLEDGMENT

The author would like to thank the support of the Wellcome Leap SAVE Program and the NIST PSCR Program. The author would like to thank Alexandra Poltorak for the collaboration on Extended Reality and review of this work. The author appreciates Andy Cao, Anthony Wang, William Zhang, and Jackson Loughney for the implementation of related prototypes at the Visual Intelligence Studio.

REFERENCES

AccuVein (2024). https://www.accuvein.com/vein-visualization-system/

Cai Y., Arunachalam U. (2021) Interactive Floor Mapping with Depth Sensors. In: Ahram T. Z., Falcão C. S. (eds) Advances in Usability, User Experience, Wearable and Assistive Technology. AHFE 2021. Lecture Notes in Networks and Systems, vol. 275. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-80091-8_3

Cai Y., Kiefel J., Wu X. (2021) Gait Recognition from Drone Videos. In: Ahram T. Z., Falcão C. S. (eds) Advances in Usability, User Experience, Wearable and Assistive Technology. AHFE 2021. Lecture Notes in Networks and Systems, vol. 275. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-80091-8_46

Cai, Y. (2003). How Many Pixels Do We Need to See Things? Lecture Notes in Computer Science, Edited by Peter M. A. Sloot et al., LNCS 2657, Computational Science, ICCS 2003, Part III, Springer-Verlag.

Cai, Y. (2004). Minimalism Context-Aware Displays, Journal of CyberPsychology and Behavior, Vol. 7, No. 3, 2004, indexed by MedPub.

Cai, Y. (2011). Pattern Discovery from Eroded Rock Art, in the book (Ch. 16, pp. 429–449) Digital Imaging for Cultural Heritage Preservation – Analysis, Restoration, and Reconstruction of Ancient Artworks, edited by Filippo Stanco, S. Battiato, and G. Gallo, CRC Press, 2011 https://www.routledge.com/Digital-Imaging-for-Cultural-Heritage-Preservation-Analysis-Restoration/Stanco-Battiato-Gallo/p/book/9781138073791

Cai, Y. (2016). Instinctive Computing. Springer-London, 2016.

Cai, Y. (2020). Learning on the Fly, M. Zallio (Ed.): AHFE 2020, AISC 1210, pp. 1–8, 2020. https://cmu.box.com/s/cth4gslcimzglmxdpyw1s0703ky8ns3x

Cai, Y. (2021). Human-Machine Learning with Mental Map, in the volume of "Cognitive Computing and Internet of Things", in Proceedings of AHFE 2021, Springer- Verlag, July 2021. https://cmu.box.com/s/6akhu1q9ds0vqudjmanf8h7n4ll4ejhp

Cai, Y. (2022). Extreme Reality Interfaces, AHFE-2022.

Cai, Y. (2024). Episodic memory in 3D interactive sequential graph. Proceedings of AHFE 2024, Nice, France, July 2024.

Cai, Y. and Freeman, P. (2023). Real-time imaging processing for low-vision users, Electronic Imaging, January, 2023.

Cai, Y. and Klein-Seetharaman, J. (2005), The Gedanken Experiment of Human Dynamics. Annual Review of CyberTherapy and Telemedicine: A Decade of VR, Vol. 3, pp. 77–84, 2005.

Cai, Y. and Siegel, M. (2002) Texture Characterization of Visual Appearance of Paintings, IEEE IMTC'02, 2002.

Cai, Y. and Siegel, M. (2023). Mobile incident commanding dashboard (MIC-D), Electronic Imaging, January, 2023.

Cai, Y. and Vatura, T. and Cai, Y., Head-Mounted Display for Remote Breath Pattern Detection, AHFE, 2022.

Cai, Y. Bunn, A, Liang, P. and Yang, B. (2013). Adaptive feature annotation for large video sensor networks, Journal of Electronic Imaging 22(4), Oct.-Dec. 2013.

Cai, Y. et al. (2007). Visual Transformation for Spatiotemporal Data Mining, Journal of Knowledge and Information Systems (KAIS), Vol. 11, No. 5, 2007, Springer, https://www.cmu.edu/vis/media/pdf/paper.pdf

Cai, Y. et al. (2019). Spatiotemporal Analysis of "Jello Effect" of Drone Videos, AHFE Conference, July 29, 2019, published by Springer in 2019. https://cmu.box.com/s/z0ok7auscuatao36ovh6yoy92yojgvxk

Cai, Y. Franco, R. M. Herranz, M. G. (2010). Visual latency-based interactive visualization for digital forensics. Journal of Computational Science Volume 1 Issue 2 Pages 115–120. Elsevier. June 2010.

Cai, Y. Morales, J. A. and Wang, S. (2020). Cyber Attribution from Topological Patterns, In proceedings of International Conference of Computational Science (ICCS), Springer-Verlag, Amsterdam, 2020: https://cmu.box.com/s/9q7x4w1xxgkukvo5y8lsg9tf88vwecf1

Cai, Y. Snel, I. Bharathi, B. S. Klein, C. and Klein-Seetharaman, J. (2005). "BioSim: A Character-Based Biomedical Problem-Solving Environment", Journal of Future Generation Computer Systems, Volume 21, Issue 7, July 2005, Pages 1145–1156.

Cole, M. (2021) Picasso and 'Guernica': Exploring the Anti-War Symbolism of This Famous Painting. My Modern Met. https://mymodernmet.com/pablo-picasso-guernica/

Hackett, S. Alber, F. and Cai, Y. (2020). Haptic Helmet for Emergency Responses in Virtual and Live Environments, in LNCS, HCII 2020, Springer-Verlag. https://cmu.box.com/s/s31xvan45mukjytr2he0l9een5yuy0lm

Jones J. (2021). Did art peak 30, 000 years ago? How cave paintings became my lockdown obsession. The Guardian, 23 April, 2021. https://www.theguardian.com/artanddesign/2021/apr/23/cave-paintings-art-lockdown-obsession-30-000-years-lascaux

Polaczyk J., Croft K., Cai Y. (2021) Compositional Sonification of Cybersecurity Data in a Baroque Style. In: Ahram T. Z., Karwowski W., Kalra J. (eds) Advances in Artificial Intelligence, Software and Systems Engineering. AHFE 2021. Lecture Notes in Networks and Systems, vol. 271. Springer, Cham. https://link.springer.com/chapter/10.1007/978-3-030-80624-8_38

Scott, D. (2019). The Bull by Pablo Picasso – A Lesson in Abstraction. Draw Paint Academy. https://drawpaintacademy.com/the-bull/