DESIGNING ADAPTIVE AND CONTEXT-AWARE AR/VR/MR SYSTEMS

Anyone who has witnessed the adoption of the Internet remembers the static and non-context-aware websites. Compare that with the powerful engines behind the websites of today. These context-aware engines are equipped with machine learning and optimization algorithms that allow them to adapt and cater to their user's behavior and environment. This deep understanding of the user and their needs created a more personalized and efficient experience. Current AR/VR/MR systems are not quite as static as the websites of yesteryear but are still a long way to go from becoming as powerful and context-aware as browsing the web is today. **In my research, I aspire to facilitate the road to achieving context-aware AR/VR/MR systems that elegantly adapt their interactions to their user's behavior and environment.**

I typically approach building these systems using the **Computational Interaction Design** paradigm [17] which leverages Machine Learning and Optimization to guide the interaction between the user and system. For example, in the past I built a Virtual Reality (VR) navigation tool that caters to users' needs without undermining the immersiveness of the virtual experience. I also developed a tool capable of optimizing the arrangement of elements (e.g. artwork, advertisements) in a virtual space according to the amount of visual attention they receive. Furthermore, I helped design a virtual pet trained with real pet behavioral data for Mixed Reality. The virtual pet was also equipped with scene awareness and is capable of detecting various objects in the scene (e.g. chairs, beds, scratchers) and behaving on them accordingly (e.g. eating from a bowl and not from a bed). In my latest work as an intern with Adobe, I explored using augmented reality (AR) sketching in the design process by developing a tool for drawing 3D curves in AR using only a tablet and a stylus.

My future work will continue to explore context-aware AR/VR/MR systems at the intersection of **Human-Computer Interaction**, **Computer Graphics**, and **Machine Learning**. I will investigate **human-centered** techniques to apply scene understanding and/or behavioral analysis for AR/VR/MR for **(1) content anchoring**, **(2) interactive storytelling**, and **(3) generating content**.

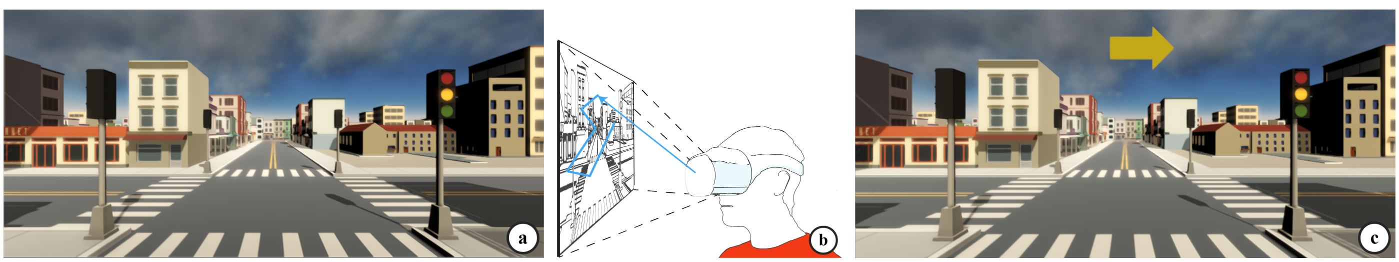

Figure 1: An illustration of our navigation tool. (a) While navigating a scene in virtual reality, (b) the user's gaze sequence can indicate his/her need for navigation help and (c) an aid is adaptively displayed when needed.

GAZE-DRIVEN ADAPTIVE VR NAVIGATION AID

In "Lost in Style" (CHI'19 [3]), I created a Long Short-term Memory (LSTM) network that can predict when users feel the need for navigation aid in VEs using their gaze patterns. I trained the LSTM using data collected from users navigating a VE while wearing a gaze-tracking VR headset. I used this LSTM to devise an adaptive aid that is shown to users only when they need navigation help (Figure 1).

Users often face difficulties exploring and navigating a virtual environment which is frustrating and weakens the immersive factor that makes virtual experiences so enticing. These feelings of "needing navigation aid" often go unnoticed by developers until the user issues a complaint about the navigation difficulty. In response, developers are left to make difficult design choices to incorporate permanent navigational guides (e.g. mini-maps, arrows, overlaid paths). These navigational guides can take up considerable screen space, be visually unappealing, or be distracting to embed in the VR interface.

Adding immersion to the virtual experience is not the only benefit of our adaptive tool. Our tool could be used by decision-makers to unintrusively collect data from users about an environment. This can allow architects to analyze the navigation difficulty of spaces under development. We can even use these cues to automate the design of real-world spaces.

SPATIAL DESIGN GUIDED BY VISUAL PERCEPTION

2D Spatial Design. I devised a system that can optimize the placement of 2D visual elements (e.g. advertisement, artwork) in a virtual environment (VE) according to their visual attention (IEEE VR'19 [4]). I achieved this by training a Support Vector Machine (SVM) regressor that can estimate the amount of visual attention a visual element will receive with data collected from users navigating a virtual museum and subway station. I used Markov Chain Monte Carlo (MCMC) to find the right arrangement of visual elements to accomplish a target amount of attention (Figure 2a).

This is particularly important for VR designers because it can streamline the cost/benefit analysis of advertising at a particular location in a VE, as well as allow them to maximize the amount of attention advertisements in their VE can receive, and as a result increase financial gain. Our method can also aid curators in finding the optimal arrangement of artwork in museums by equalizing and maximizing the opportunity for artists to showcase their work. We involved curators and design experts in the creation of our system.

Figure 2: I created systems that optimize the arrangement of elements in VEs (e.g. (a) artwork in museums, (b) clothing in stores) according to their visual attention.

3D Spatial Design. The rise in popularity of consumer-grade VR headsets opens the door for researchers to explore creating interactive e-commerce experiences in the form of "virtual stores". Virtual stores can be designed to have the cohesion of real-world stores with an added bonus of the interactivity that VR can offer.

We extended our spatial design system [4] to optimize the placement of 3D items (e.g. toys, clothing) in a virtual store by maximizing their visual attention (IEEE VR'23 [12]). After consulting with experts and referencing store design books [8, 20], we identified a set of spatial constraints (e.g. avoid over-crowding items on shelves, place items within equally-spaced intervals) that should also be considered for an ideal placement of items in a store (Figure 2b).

MULTI-VIEW CURVE DRAWING IN AUGMENTED REALITY

The introduction of AR glasses like the HoloLens paved the way for in-situ sketch-based design. Though these devices are fun to work with, they are not yet widely accessible to designers unlike tablets (e.g. iPad). In my internship with Adobe, I utilized the ubiquity of tablets to introduce in-situ sketch-based interactions into designers' workflow (IEEE VR'23 [2]). A user aiming to prototype a chair with scale and style similar to a table existing in the real environment could do so by sketching 2D strokes on a tablet. The 2D strokes are overlayed onto the real world via the AR device while considering the user's intention. Projecting the sketch into AR can be accomplished using a 2D sketch to 3D curve projection algorithm [10]. I integrated this into a system that enables users to scan the environment geometry and to sketch curves that wrap around and interact with the geometry (e.g. curve that wraps around a tree to decorate it with ornaments in Figure 3).

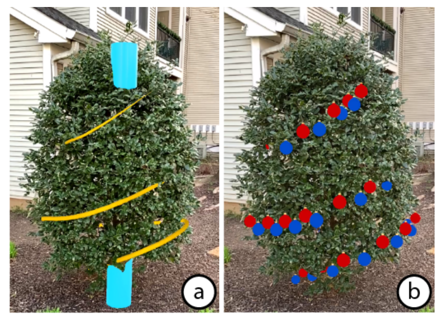

Figure 3: (a) A curve drawn with our environment-aware mobile AR drawing system. (b) This curve can decorate the real world with virtual ornaments.

Unlike previous AR drawing approaches [6, 11], drawing curves around geometry enables designing in in-accessible locations (e.g. using the sketched curve to place flowers on the high arc). Moreover, our multi-view drawing system affords users the ability to create large-scale animation motion paths (e.g. a flight path for an animated plane that flies around in a large room), by allowing them to draw smaller curves around the geometry that are joined to create a larger curve that considers the underlying geometry.

SCENE-AWARE MIXED REALITY SYSTEMS

Interactive Pets. Virtual pets can be a great alternative for those who aren't able to physically or financially care for a real animal. In addition to gameplay; virtual pets have been used in therapy [5], learning [9] and elderly care. I helped to create an interactive Mixed Reality (MR) pet (Figure 4) with a behavioral model capable of understanding scene semantics and mirroring behavior patterns obtained from an indoor pet (CHI'21 [13]). We achieved this by training an LSTM with behavioral data collected from a cat to synthesize the virtual pet's behavior in Mixed Reality. To create a more realistic experience, we also incorporated scene understanding into the virtual pet. We elevated the interactivity of the Mixed Reality experience by utilizing HoloLens' voice recognition feature to allow the user to interact with the pet by giving it commands (e.g. sit, fetch).

Figure 4: Our scene-aware virtual pet.

Assistive Technology. Augmentative and alternative communication (AAC) is a communication mechanism for those with complex communication needs, and existing AAC devices are forms of assistive technology comprising hardware and software that can support or replace natural speech entirely.

I also helped develop a novel context-aware mixed-reality AAC system in collaboration with Assistive and Special Education Technology experts at George Mason University [16] (UIST'23, to be submitted). This system aids users with communicating in grocery shopping scenarios by recognizing objects in the scene and generating the proper spoken message.

RESEARCH AGENDA

My future work will build on my past work in scene understanding, spatial design, behavioral analysis, and AR/VR. I will focus on the problems related to integrating AR content seamlessly into the real-world environment and devising intuitive interactions for this content.

Content Anchoring. Lindlbauer et al. [14]'s system adapts AR content according to the user's cognitive load and field-of-view. Content is manually anchored to the scene by the user. In Cheng et al. [7]'s system, content is placed in the scene with an optimization. However, the system relies on the user to specify an ideal placement and its parameters in an initial environment that can then be transferred and optimized to another environment. I hope to create a standalone system that's able to optimize the placement of content automatically without the need for a reference placement.

Figure 5: AR content anchored into the scene.

I'm interested in exploring AR systems that integrate scene understanding to seamlessly and automatically anchor and adapt content into our world (Figure 5). To place content in a real-world scene we must consider the scene's semantics and user behavior. For example, contents of a presentation should be anchored on the meeting room table instead of on the snack table or a chair. The placement should also consider the content's visibility and avoid placing it in occluded locations. We could also encode multi-view visibility and interactivity if the content is meant to be viewed by multiple users in the scene. The content's placement could also be optimized to maximize visual attention similar to my previous work in spatial design [4]. Furthermore, if the content is interactive we must consider interactivity and accessibility in placing the content.

Interactive Storytelling. In my previous work, I explored integrating animation motion paths into the environment geometry [2]. Inspired by other works in gesture-driven storytelling and animation [19] (Figure 6), I would like to explore gesture-driven animation of 3D content that has been integrated into the environment geometry.

I'm also interested in exploring incorporating interactions into Spatial AR. These interactions could aid the viewer in obtaining more information, by for example, manipulating or expanding the content. Or they may be used to furnish the content creator with the ability to author these storytelling AR experiences in a standalone system, providing them with the ability to specify the effects, key frames, motion paths, timing, and/or triggers (e.g. gesture, user position, speech).

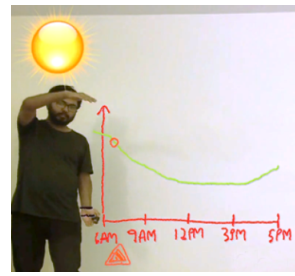
Figure 6: Gesture-driven AR animation [19].

Generative Content. In my work with Adobe, we trained a Generative Adversarial Network (GAN) to automatically create a colored cartoon character from a rough sketch (GI'21 [1]). Using the GAN, we developed a tool that allows artists to visualize their character designs by automatically coloring them from a selected color scheme as they sketch. This system enabled artists to design characters faster than traditional methods—a useful feature in the early stages of character design—when artists need to iteratively re-design their characters according to clients' feedback and experiment with multiple color schemes and styles.

There's great potential in using generative models for AR . For example, researchers like Liu et al. [15] explored using GANs to light AR scenes. While Peebles et al. [18] used a GAN to generate and align AR content (Figure 7). I would like to continue utilizing generative models, especially for generating AR content and populating/decorating scenes.

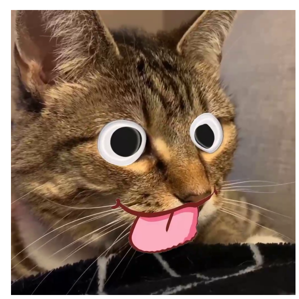
Figure 7: GAN-aligned AR content. [18].

COLLABORATIONS

Dr. Marc Pomplun from the University of Massachusetts-Boston consulted me on visual attention and employing eye-tracking data for behaviour classification. Dr. Cuong Nguyen, Research Scientist at Adobe, consulted me about the latest advances in AR sketch-based interactions. Dr. Vojtěch Krs, Dr. Nathan Carr and Dr. Radomír Měch, Research Scientists at Adobe, consulted me about 2D sketch to 3D curve projection. Dr. Vivian Genaro Motti and Dr. Yoosun Chung from George Mason University consulted me on assistive technology.

POTENTIAL FUNDING SOURCES

I identified several awards and grants that I intend to apply to for funding:

- NSF Future of Work at the Human-Technology Frontier
- NSF CISE IIS Human-Centered Computing (HCC)
- NSF Education and Human Resources (EHR) Core Research
- National Endowment for the Humanities (grants related to AR/VR/XR in digital humanities)

REFERENCES

[1] Rawan Alghofaili, Matthew Fisher, Richard Zhang, Michal Lukáč, and Lap-Fai Yu. "Exploring Sketch-based Character Design Guided by Automatic Colorization". In: *Graphics Interface 2021*. 2021.

[2] Rawan Alghofaili, Cuong Nguyen, Vojtěch Krs, Nathan Carr, Radomír Měch, and Lap-Fai Yu. "Warpy: Multi-view Curve Sketching in Augmented Reality". In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2023.

[3] Rawan Alghofaili, Yasuhito Sawahata, Haikun Huang, Hsueh-Cheng Wang, Takaaki Shiratori, and Lap-Fai Yu. "Lost in Style: Gaze-driven Adaptive Aid for VR Navigation". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: ACM, 2019, 348:1–348:12. ISBN: 978-1-4503-5970-2.

[4] Rawan Alghofaili, Michael Solah, Haikun Huang, Yasuhito Sawahata, Marc Pomplun, and Lap-Fai Yu. "Optimizing Visual Element Placement via Visual Attention Analysis". In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 464–473.

[5] Eric Lewin Altschuler. "Play with online virtual pets as a method to improve mirror neuron and real world functioning in autistic children". en. In: *Med. Hypotheses* 70.4 (2008), pp. 748–749.

[6] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. "SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–15. ISBN: 9781450356206.

[7] Yifei Cheng, Yukang Yan, Xin Yi, Yuanchun Shi, and David Lindlbauer. "SemanticAdapt: Optimization-Based Adaptation of Mixed Reality Layouts Leveraging Virtual-Physical Semantic Connections". In: *The 34th Annual ACM Symposium on User Interface Software and Technology*. UIST '21. Virtual Event, USA: Association for Computing Machinery, 2021, pp. 282–297. ISBN: 9781450386357.

[8] Claus Ebster. *Store Design and Visual Merchandising: Creating Store Space That Encourages Buying*. Business Expert Press digital library. Business Expert Press, 2011. ISBN: 9781606490952. URL: `https://books.google.com/books?id=1kg8S3LuzCQC`.

[9] Andreas Holzinger and Hermann Maurer. "Incidental learning, motivation and the Tamagotchi Effect: VR-Friends, chances for new ways of learning with computers". In: *Computer Assisted Learning, CAL 99*. Elsevier BV, 1999, pp. 70–70.

[10] Vojtěch Krs, Ersin Yumer, Nathan Carr, Bedrich Benes, and Radomír Měch. "Skippy: Single View 3D Curve Interactive Modeling". In: *ACM Trans. Graph.* 36.4 (July 2017). ISSN: 0730-0301.

[11] Kin Chung Kwan and Hongbo Fu. "Mobi3DSketch: 3D Sketching in Mobile AR". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–11. ISBN: 9781450359702.

[12] Wei Liang, Luhui Wang, Xinzhe Yu, Changyang Li, Rawan Alghofaili, Yining Lang, and Lap-Fai Yu. "Optimizing Product Placement for Virtual Stores". In: *2023 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2023.

[13] Wei Liang, Xinzhe Yu, Rawan Alghofaili, Yining Lang, and Lap-Fai Yu. "Scene-Aware Behavior Synthesis for Virtual Pets in Mixed Reality". In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2021. ISBN: 9781450380966.

[14] David Lindlbauer, Anna Maria Feit, and Otmar Hilliges. "Context-Aware Online Adaptation of Mixed Reality Interfaces". In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. New Orleans, LA, USA: ACM, 2019. ISBN: 978-1-4503-6816-2/19/10.

[15] Daquan Liu, Chengjiang Long, Hongpan Zhang, Hanning Yu, Xinzhi Dong, and Chunxia Xiao. "ARShadowGAN: Shadow Generative Adversarial Network for Augmented Reality in Single Light Scenes". In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 8136–8145. DOI: `10.1109/CVPR42600.2020.00816`.

[16] Yu Liuchuan, Huining Feng, Rawan Alghofaili, Boyoung Byun, Swati Rampalli, Yoosun Chung, Vivian Genaro Motti, and Lap-Fai Yu. "An AI-driven AAC system on Hololens". 2023. *(to be submitted)*.

[17] Antti Oulasvirta, Per Ola Kristensson, Xiaoju Bi, and Andrew Howes. *Computational interaction*. Oxford University Press, 2018.

[18] William Peebles, Jun-Yan Zhu, Richard Zhang, Antonio Torralba, Alexei A. Efros, and Eli Shechtman. "GAN-Supervised Dense Visual Alignment". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022, pp. 13470–13481.

[19] Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, and Wilmot Li. "Interactive Body-Driven Graphics for Augmented Video Performance". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, pp. 1–12. ISBN: 9781450359702.

[20] L. W. Turley and Ronald E. Milliman. "Atmospheric Effects on Shopping Behavior: A Review of the Experimental Evidence". In: *Journal of Business Research* 49 (2000), pp. 193–211.