Active Appearance and Spatial Variation Can Improve Visibility in Area Labels for Augmented Reality

Hojung Kwon Yuanbo Li Xiaohan Ye Praccho Muna-McQuay Liuren Yin[‡] James Tompkin
Brown University [‡]Duke University

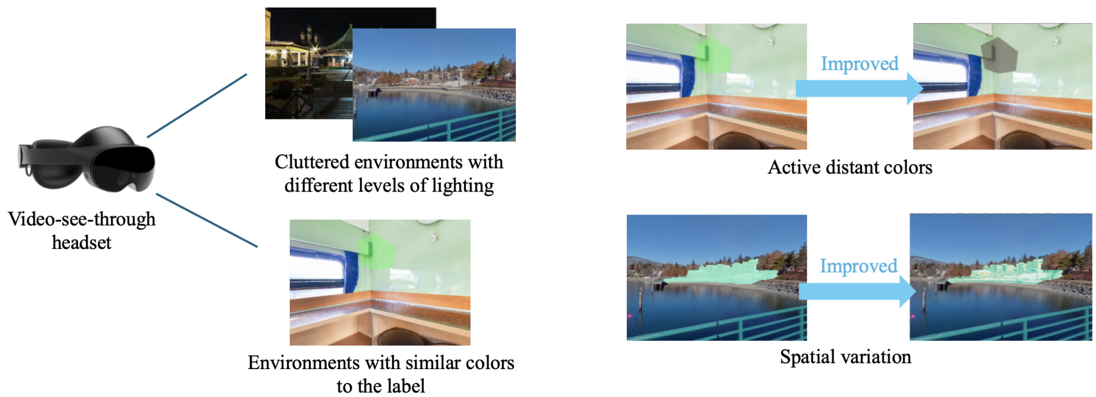

Figure 1: *Overview.* To address area label visibility and environment occlusion in AR, our system uses active distant colors to enhance perceived contrast between label and environment and uses spatial variation to reduce environment occlusion.

Abstract

Augmented reality (AR) area labels can visualize real world regions with arbitrary boundaries and show invisible objects or features. But environment conditions such as lighting and clutter can decrease fixed or passive label visibility, and labels that have high opacity levels can occlude crucial details in the environment. We design and evaluate *active* AR area label visualization modes to enhance visibility across real-life environments, while still retaining environment details within the label. For this, we define a distant characteristic color from the environment in perceptual CIELAB space, then introduce spatial variations among label pixel colors based on the underlying environment variation. In a user study with 18 participants, we found that our active label visualization modes can be comparable in visibility to a fixed green baseline by Gabbard et al., and can outperform it with added spatial variation in cluttered environments, across varying levels of lighting (e.g., nighttime), and in environments with colors similar to the fixed baseline color.

Index Terms: Augmented reality, active labels, environment-adaptive.

1 Introduction

Augmented reality (AR) labels visualize information that is invisible in the real world [10, 24]. These labels can reveal hidden details on historical artifacts [21], annotate defects on building materials [11, 13], or help people navigate to a destination [16]. AR designers must ensure visibility across environments with varying lighting and background complexity or clutter. This is difficult as the environment is often unknown. To overcome this, AR label generation systems must flexibly tailor label appearance to the environment. For instance, to enhance visibility, AR systems might place labels in less cluttered areas [12] or might add semi-transparent billboards [7, 22]. Some content design guidelines have adopted such approaches for AR text labels [2, 4].

We design a system for AR area labels, which we define as AR shapes that highlight visible scene features or visualize invisible features. We choose area labels because, for area labels that must be situated upon the real-world, existing solutions to enhance visibility are less applicable. For instance, an AR label might show the location of a tumor in a patient's brain [17]—this label cannot be moved without losing spatial correspondence with the object. A label outline may be appropriate but the outline does not allow details within the label to be highlighted. Further, enhancing the visibility within area labels is challenging in cluttered environments with many varying details, as labels are typically uniform in appearance. We investigate an approach for *active* AR labels [9] that can dynamically adjust to the surrounding environment both in overall appearance and in local appearance variation as in Fig. 2. This local variation considers the spatial content within and around an area label at different granularity levels, e.g., appearance variation per pixel within the label vs. whole label variation. We consider a video-see-through (VST) AR setting as this allows us to experimentally analyze the environment easily.

In a user study, we evaluated these area label visualization modes against a baseline constant label (semitransparent green as suggested by Gabbard et al. [8] and Hombeck et al. [14]). We evaluated overall label visibility in terms of interior detail preservation and in terms of exterior shape. We found that our visualizations modes outperformed the baseline in certain backgrounds. These findings suggest that, for VST AR, adaptive area labels are a promising alternative to constant labels when the environment is unknown, and that area label visibility can benefit from spatial variation to enhance detail in cluttered environments.

2 Related Work

AR label visibility enhancement. Maintaining the visibility of AR contents under different lighting conditions and background complexity is difficult [6]. Guidelines for optical-see-through (OST) AR devices suggest use only in ideal conditions such as indoor environments with lighting intensity at 500–1000 lux [18]. Some research suggests that there are colors that are suitable for AR labels regardless of background conditions. For text labels, Gabbard et al. [9] found that green labels and blue labels superimposed on white billboards had the best legibility against six real-life textures, visualized as real world posters, compared to red labels and active labels that changed color based on average background colors. Jankowski et al. [15] found that white text on a semitransparent black billboard (or vice versa) had the best visibility and legibility compared to plain texts and texts with shadow or outline. For area labels, Hombeck et al. [14] identified ranges of opacity levels at which AR area labels with different patterns and outlines would remain

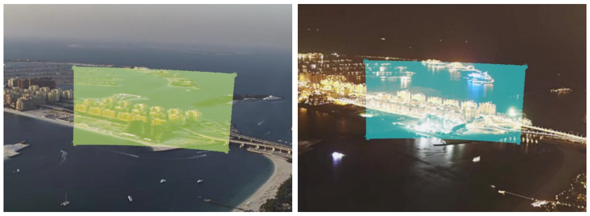

Figure 2: Active labels can dynamically enhance label visibility based on backgrounds. We visualized these area labels (40% opacity) at an identical location on different frames from a 360° time lapse video.

Figure 3: As different environment pixels can have the same most perceptually distant color, computing a per-pixel most distant color results in "clumping", the loss of the background details as in the image on the right.

visible, while not occluding background details. For OST, Gabbard et al. [8] again found that green, orange, and blue AR label colors kept contents more visible and legible without being unintentionally modified due to being mixed with background colors (perceptual shift).

Other studies, in contrast, showed that there is a potential benefit to labels that change their appearance based on environment conditions to maintain visibility. Ahn et al. [1] introduced a system that enhances AR content visibility by changing CIELAB luminance levels based on contrast saliency with the background. Riegler et al. [22] and Grasset et al. [12] introduced semitransparent billboards and leader lines that change their colors depending on environment lighting conditions.

Adopting such differing results into AR label design guidelines is complicated. Static labels, whose appearances remain consistent regardless of the background, have decreased visibility in backgrounds with similar colors as the labels. Therefore, we present an active way to modify label colors, spatial variation, and opacity level to enhance visibility.

AR area labels. These labels highlight scenes and objects to provide information that is otherwise unavailable to users. Ridel et al. [21] developed an AR overlay to highlight concavities and convexities on historical artifacts, aiding a human in the distinction of details from eroded areas. Chang et al. [5] evaluated gesture-based methods for virtually annotating features on a real-life object and introduced techniques for beautifying these annotations to increase their precision.

Area labels are also used to visualize features that are invisible in real life. Garcia-Pereira et al. [11] introduced an AR system for showing 3D models of building components at a real-life location where those components would be placed. Manual controls for the components' opacity level enabled users to choose an opacity value that would visualize the components without occluding background details. Liao et al. [17] designed a system that projects 3D images of brain tumor and ventricle onto a patient's head to help neurosurgeons locate them during surgeries and realistically visualize the tumor's shape.

These existing systems show that when highlighting real-life objects, AR overlays should clarify what objects they are highlighting, while not occluding details in the background under the overlay. These systems also reveal the need for an AR overlay whose shape remains identifiable regardless of its surrounding environment. With an active label approach for area labels and internal spatial variations within a label, we assess whether users can identify details from the background under the label, while also being able to recognize the shape of the label.

3 Methods

To create an environment-adaptive active area label system that can vary its internal label spatial variations, we begin with the video-see-through AR setting. This setting makes it simple to access the appearance of the surrounding environment at each display pixel. We implemented our approach in Unity on a Meta Quest Pro. As read access on the passthrough video feed is not available to third-party developers, we used 360° photographs to simulate real scenes.

Next, we consider four visualization features to define appearance to an area label: a *color space and distance measure* within that space, a method to pick a *characteristic environment color* from which to find a distant color for a label, a method to introduce *spatial variation* within a label, and label *opacity*. By combining these features, we aim to improve visibility so that people can perceive detail within an area label regardless of the scene lighting, environment color, and level of clutter. We also aim to enhance the visual contrast between the environment and the label so that the label is easy to locate and its shape is clearly recognizable.

3.1 Color space and distance measure

Contrasting colors are often used to draw attention to features in visualization [20, 23]. Therefore, when trying to assign colors that will be visible against a dynamic environment, one core task is to quickly find a color that is distant (or most distant) from another. We use the standard CIELAB human perceptual color space for all computations. As a distance measure within CIELAB, we use CIEDE2000 (or CIE ΔE^*_{00}) instead of earlier ΔE^* measures (CIE76 or CIE94) as it better measures a perceptually uniform distance, especially in saturated regions, blue regions, and for neutral colors. Unlike the Euclidean distance of ΔE^*_{76}, ΔE^*_{00} is not simple to compute: it is neither linear nor differentiable. This makes it hard to use within solvers and more expensive to reason with. Further, CIELAB does not have a simple boundary, making geometric queries more complex.

Precomputing Distant Colors. To ease this process, we precompute a 3D look-up table that stores the farthest color by ΔE^*_{00} from all input sRGB colors that might appear in the environment. Computed by brute force over 8-bit three-channel input, this is expensive but only has to happen once: each of the 256^3 inputs requires converting to CIELAB and computing ΔE^*_{00} to a large sample of the continuous outputs in CIELAB space. To speed up this computation, we sample every 4th RGB value in the 8-bit three-channel RGB space (e.g., (0,0,4), (0,0,8), ...), convert each to CIELAB, and then consider this our set of possible input and output points. Doing so leaves still 68.7 billion comparisons by ΔE^*_{00}, which took 41 minutes on 4 CPUs. After finding the farthest for each input point, we trilinearly interpolate the remaining points. We use the resulting table as a 3D lookup texture in Unity.

3.2 Characteristic environment colors

As area labels cover regions of an environment that extend across multiple pixels in the VST display, we must decide how to color the label so that it contrasts with the set of environment pixels. For instance, naively computing a mean distant color over the region will produce dull mid-gray labels if that environment region itself contains differing colors. Moreover, naively computing a distant color for each pixel in the environment region for spatial variation tends to remove environment details in the resulting label as in Fig. 3.

Instead, we first derive a characteristic environment color (CEC) from which to compute a distant color for our overall label appearance. This is based upon the binning method of Nedrich [19]. After dividing the RGB color space into 27 equally-sized bins, we bin each environment region pixel. We select the bin with the most assigned pixels, then use the average color of all assigned pixels as the characteristic environment color.

3.3 Spatial variation

Given the characteristic environment color $\mathbf{c}$ converted to a CIELAB value, assigning to our area label an LAB value $\mathbf{c}'$ that is most distant from $\mathbf{c}$ may still reduce contrast with some details of the environment that we would like to observe. As such, we design a method to restore

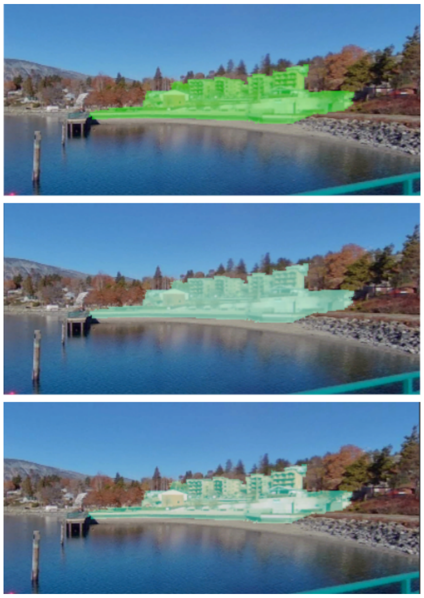

Figure 4: **Assignment modes.** Scene *outdoor daytime.* All shown at 40% opacity. *Top:* Mode A—baseline green. *Center:* Mode C—no spatial variation. *Bottom:* Mode E—spatial variation.

some of this contrast while maintaining an overall label appearance that is far from the environment. As a secondary effect, this method also compensates somewhat for the complex geometry of the valid CIELAB color space, such as corners, where different input colors may be most distant from the same output color—this causes "clumping" in appearance if a naive per-pixel distant color approach is used.

To resolve this, given the characteristic environment color $\mathbf{c}$, we compute the vector difference in CIELAB space $\mathbf{d_i} \in \mathcal{D}$ to the color of every pixel in the environment within the region of the area label. Then, we compute the most distant color $\mathbf{c}'$ from $\mathbf{c}$. Finally, we add back the vector difference $\mathbf{c}' + \mathbf{d_i}$ to produce a color for each output pixel in the area label.

3.4 Opacity level

Opacity level variations can help AR systems to highlight objects while reducing background occlusion [8, 14, 22]. Hombeck et al. [14] studied opacity for solid area labels without patterns. They found that the ideal opacity range at which 1) the label outline was visible, and 2) background details interior to the area label were visible, was between 20% and 70%. In our specific VST system with active labels, we as authors judged that 20% was too translucent; instead, we raised this floor to 40%.

4 User Study

We invited 18 participants (8 male, 10 female) to our study, with ages between 19–36. 11 participants had used AR or VR applications before, and 7 participants had developed AR or VR software. All participants had normal vision. Each user study session took 20–30 minutes and we gave participants a $15 gift card as a reward.

4.1 Color assignment modes

Given the features in the previous section, we define a set of visualization modes including baseline passive (not active) label modes.

Baseline. Given human visual system's high green sensitivity, research in both area and text labels found that green labels had the best legibility [8, 9], with green being among the best colors for semi-transparent area labels to 1) minimize perceived color changes of background details, and 2) maintain label shape visibility [8]. As such, we test semi-transparent green (0,255,0) labels (fixed; not active).

Our approach. We include our visualization modes (Fig. 4). Without spatial variation, our mode assigns the same color across the area label, so it may enhance overall visibility in comparison to the environment. Adding spatial variation tries to improve details within the area. We evaluate all modes at both 40% and 70% opacity. Alphabetizing all modes, we have:

A. Green color. Opacity: 40%.
B. Green color. Opacity: 70%.
C. Active distant color. Opacity: 40%. Spatial variation: off.
D. Active distant color. Opacity: 70%. Spatial variation: off.
E. Active distant color. Opacity: 40%. Spatial variation: on.
F. Active distant color. Opacity: 70%. Spatial variation: on.

4.2 Scenes

We define a set of area labels across three scenes, chosen to emphasize different environment characteristics. Scene *outdoor nighttime* is a park with bright lights, creating a high contrast environment that is more difficult for active labels to maintain interior detail due to the potential for saturation. Scene *outdoor daytime* is a lake scene with apartments in a forest; the apartments act as an object to be highlighted within an area of visual detail or clutter, but where the object itself also has interior visual details to be highlighted. Scene *indoor green* is the interior of a train with mint green walls; this evaluates the need for active labels when a fixed label appearance is similar to the environment.

4.3 Protocol

At the beginning of the study, we showed participants a short presentation defining ideal and non-ideal labels upon an otherwise unused scene of the Milky Way galaxy in the night sky. A label was said to be ideal if the exterior shape of the area label was visible against the environment, and if the interior details of the environment within the area label were visible. We showed example ideal and non-ideal conditions for each case.

Next, in each of the three test scenes, we presented labels and visualization modes in a random order to participants within a two-alternate force choice setting. Participants used the Meta controller to toggle between choices and make selections. Participants were shown a sequence of multiple comparison pairs across the three scenes. Upon making a choice of preferred mode for each comparison pair, participants had to select a rationale for the choice: "Background under the label was more visible", "Label outline was more visible", "All of the above", and "None of the above". The first three rationales let us identify which modes clearly visualize label shapes while not occluding details in the background. If the participant chose "None of the above", we asked the participant to elaborate using free speech to the experiment conductor.

We use the Spring Rank algorithm [3] to rank modes, with lower ranks indicating higher preference among participants. From sets of pairwise preferences, Spring Rank determines an overall ranking and assigns a distance score for each mode. A higher score for a mode indicates a higher likelihood for that mode to be preferred (and have a lower rank) compared to other modes.

5 Findings

Across all three scenes, mode E (active distant color, 40% opacity, spatial variation on) was most preferred and had the highest rank (Fig. 5; ordered vertically from most to least preferred across all scenes; Spring Rank mode scores: E=0.85, A=0.73, F=0.62, C=0.56, D=0.22, B=0). When looking at rankings of the three curated test scenes individually, we see the attributes of the methods producing more varied preferences.

Outdoor nighttime—cluttered high contrast. Mode E had the lowest rank in *outdoor nighttime* (Fig. 7). As rationale, most participants selected "Background under the label was more visible" (79.22%) as in Fig. 6. From these results, we see that in a cluttered environment with varying levels of lighting across different regions of the environment, our spatially variant active color assignment mode can be better than

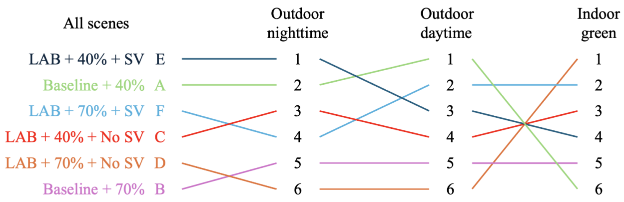

Figure 5: **Preference ranks.** Mode E is overall preferred, but each scene reveals different aspects of performance: Baseline green performs poorly in *indoor green*, whereas 70% opacity labels tend to perform well on *indoor green* as the area label region does not have complex interior detail.

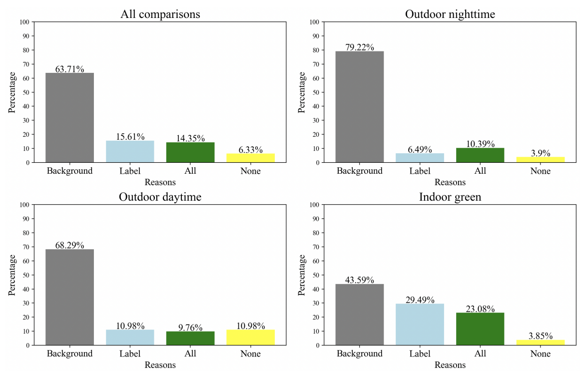

Figure 6: Participants predominantly chose "Background under the label was more visible" as their rationale. However, in *indoor green*, more participants chose "Label outline was more visible" than in other scenes.

the baseline in terms of minimizing occlusion, while also preventing saturation caused by a bright light source.

Outdoor daytime—cluttered bright. Mode A (baseline, 40% opacity) had the highest rank, as in Fig. 8, with "Background under the label was more visible" being the predominant reason (68.29%). Many participants who chose "None of the above" explained that visualization modes with spatial variations unrealistically altered details from the area under the label. Compared to the other outdoor scene, the overall brightness for this scene was higher. Thus, it may have been easier for participants to distinguish colors of objects from the environment, making the effects of color assignment with spatial variations appear more drastic.

Indoor green—similar colors to the baseline. In the *indoor green* scene, mode D (active distant color, 70% opacity, spatial variation off) had the lowest rank as in Fig. 9. We can also observe that the two baseline modes had the highest ranks in this scene. In this scene, the area under and directly surrounding the label had similar colors to the baseline. Therefore, the active distant color modes may have been better at capturing the characteristic color of the environment near the baseline.

The fact that more participants chose "Label outline was more visible" in this scene than in others (29.49% in this scene vs. 6.49% and 10.98% in other scenes) also supports that when the environment has similar colors to the baseline visualization modes, our active visualization modes with a higher perceptual contrast to the environment are more preferred.

6 Conclusion

We consider the under-studied problem of active area labels within VST AR systems. In an unknown environment, these must highlight an area while also showing details within it, requiring a label that can dynamically adjust appearances based on environment colors and human perception. Our visualization modes outperformed the baseline when minimizing environment occlusion due to clutter across two levels of lighting, and when enhancing the visibility of label shape in an environment with similar colors to a fixed green baseline label—*it's not easy being green*.

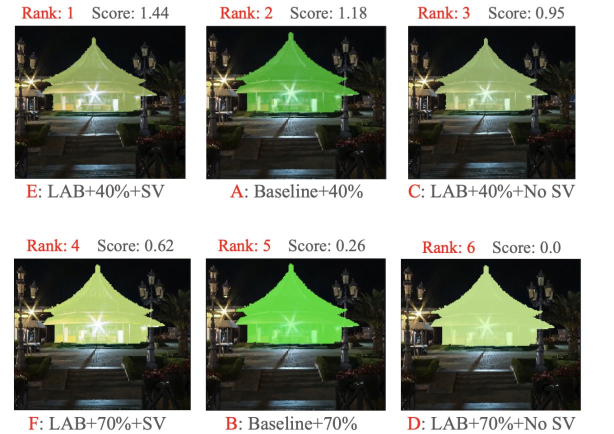

Figure 7: In *outdoor nighttime* with a high level of lighting contrast, our active distant color mode with spatial variation outperformed other modes.

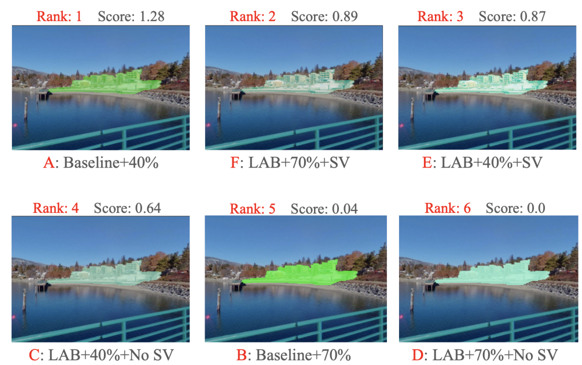

Figure 8: One of the baseline modes outperformed our modes in *outdoor daytime*. Many participants noted in this scene that some modes with spatial variations unrealistically altered environment details.

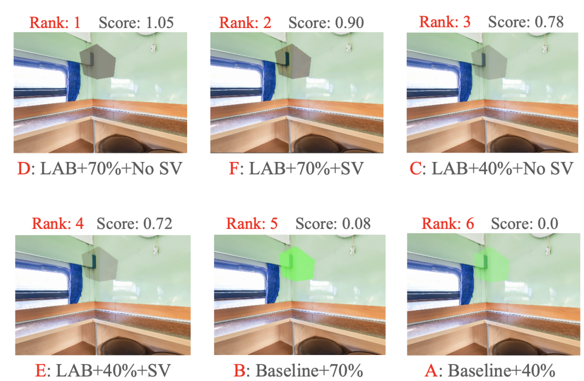

Figure 9: In *indoor green*, all of our active modes outperformed the fixed baseline. More participants chose "Label outline was more visible" to explain their decision in this scene than in the other two scenes.

Limitations. Some environments still cause the baseline green to be preferred. This may be because in scenes with a high level of ambient lighting, the contrast enhancement effect of spatial variation within the area label appears more noticeable. Future studies may find new approaches to appearance assignment to balance specific environment characteristics. Moreover, we could not identify one mode that consistently outperformed others across all test scenes. Except for mode B (baseline, 70% opacity), which consistently ranked 5, the rankings of other visualization modes varied across the three test scenes (Fig. 5). Designing additional users studies to measure participant performance such as the amount of time participants took to locate and identify label shapes may help us to more clearly identify modes that outperform others.

Acknowledgments

This work was funded by US NSF IIS-2107409. Thank you to Eva Schiller and Brayden Goldstein-Gelb for helping to develop the augmented reality system.

References

1. E. Ahn, S. Lee, and G. Kim. Real-time adjustment of contrast saliency for improved information visibility in mobile augmented reality. *Virtual Reality*, 22:245–262, 2018. doi: 10/gd2qd6 2
2. Apple Developer. Design for spatial input. `https://developer.apple.com/videos/play/wwdc2023/10073/`, Accessed: Apr. 2024. 1
3. C. D. Bacco, D. B. Larremore, and C. Moore. A physical model for efficient ranking in networks. *Science Advances*, 4(7), 2018. doi: 10/gdxkmn 3
4. S. Benham and R. Budiu. The usability of augmented reality. `https://www.nngroup.com/articles/ar-ux-guidelines/`, Accessed: Apr. 2024. 1
5. Y. S. Chang, B. Nuernberger, B. Luan, and T. Höllerer. Evaluating gesture-based augmented reality annotation. In *Proc. 3DUI*, pp. 182–185. IEEE, 2017. doi: 10/nb52 2
6. E. Cooper. The perceptual science of augmented reality. *Annual Review of Vision Science*, 9:455–478, 2023. doi: 10/nb5x 1
7. B. Dao and J. L. Gabbard. Early steps towards understanding text legibility in handheld augmented reality. In *Proc. IEEEVR*, pp. 159–160. IEEE Computer Society, 2013. doi: 10/nb5s 1
8. J. L. Gabbard, M. Smith, C. Merenda, G. Burnett, and D. R. Large. A perceptual color-matching method for examining color blending in augmented reality head-up display graphics. *IEEE Transactions on Visualization and Computer Graphics*, 28(8):2834–2851, 2022. doi: 10/nb5w 1, 2, 3
9. J. L. Gabbard, J. E. Swan, and D. Hix. The effects of text drawing styles, background textures, and natural lighting on text legibility in outdoor augmented reality. *Presence*, 15(1):16–32, 2006. doi: 10/dgjrhj 1, 3
10. I. García-Pereira, J. Gimeno, P. Morillo, and P. Casanova-Salas. A taxonomy of augmented reality annotations. In *Proc. VISIGRAPP*, vol. 1, pp. 412–419. SciTePress, 2020. doi: 10/nb5m 1
11. I. García-Pereira, C. Portalés, J. Gimeno, and S. Casas. A collaborative augmented reality annotation tool for the inspection of prefabricated buildings. *Multimedia Tools and Applications*, 79:6483–6501, 2019. doi: 10/ghq8hh 1, 2
12. R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-driven view management for augmented reality browsers. In *Proc. ISMAR*, pp. 177–186. IEEE, 2012. doi: 10/f3szz6 1, 2
13. S. Henderson and S. Feiner. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE Transactions on Visualization and Computer Graphics*, 17(10):1355–1368, 2011. doi: 10/d3ttbb 1
14. J. Hombeck, L. Ji, K. Lawonn, and C. Perin. A study of opacity ranges for transparent overlays in 3d landscapes. In *Proc. VIS*, pp. 66–70. IEEE, 2020. doi: 10/nb5z 1, 3
15. J. Jankowski, K. Samp, I. Irzynska, M. Jozwowicz, and S. Decker. Integrating text with video and 3d graphics: The effects of text drawing styles on text readability. In *Proc. CHI*, pp. 1321–1330. ACM, 2010. doi: 10/dh4rv6 1
16. T. Köppel, M. Eduard Gröller, and H. Wu. Context-responsive labeling in augmented reality. In *Proc. PacificVis*, pp. 91–100. IEEE, 2021. doi: 10/gnkpfm 1
17. H. Liao, T. Inomata, I. Sakuma, and T. Dohi. 3-D augmented reality for mri-guided surgery using integral videography autostereoscopic image overlay. *IEEE Transactions on Biomedical Engineering*, 57(6):1476–1486, 2010. doi: 10/c3f3fm 1, 2
18. Microsoft. Hololens environment considerations. `https://learn.microsoft.com/en-us/hololens/hololens-environment-considerations`, Accessed: Apr. 2024. 1
19. M. Nedrich. Palette maker. `https://tinyurl.com/577fs62m`, Accessed: Apr. 2024. 2
20. K. Reda and M. E. Papka. Evaluating gradient perception in color-coded scalar fields. In *Proc. VIS*, pp. 271–275. IEEE, 2019. doi: 10/nb56 2
21. B. Ridel, P. Reuter, J. Laviole, N. Mellado, N. Couture, and X. Granier. The revealing flashlight: Interactive spatial augmented reality for detail exploration of cultural heritage artifacts. *Journal on Computing and Cultural Heritage*, 7(2):1–18, 2014. doi: 10/gndpcp 1, 2
22. A. Riegler, A. Riener, and C. Holzmann. Adaptive dark mode: Investigating text and transparency of windshield display content for automated driving. In *Proc. MuC*, pp. 421–428. Gesellschaft für Informatik e.V., 2019. doi: 10/gq9krh 1, 2, 3
23. L. Wang, J. Giesen, K. T. McDonnell, P. Zolliker, and K. Mueller. Color design for illustrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1739–1754, 2008. doi: 10/bhdxjj 2
24. J. Wither, S. DiVerdi, and T. Höllerer. Annotation in outdoor augmented reality. *Computers & Graphics*, 33(6):679–689, 2009. doi: 10/ds74cw 1