International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2016)
D. Reiners, D. Iwai, and F. Steinicke (Editors)

Dynamic View Expansion for Improving Visual Search in Video See-through AR

Yuki Yano †2 Jason Orlosky ‡1 Kiyoshi Kiyokawa §1,2 Haruo Takemura ¶1,2

[1]Cybermedia Center, Osaka University [2]Graduate School of Information Science and Technology, Osaka University

Abstract

The extension or expansion of human vision is often accomplished with video see-through head mounted displays (HMDs) because of their clarity and ability to modulate background information. However, little is known about how we should control these augmentations, and continuous augmentation can have negative consequences such as distorted motion perception.

To address these problems, we propose a dynamic view expansion system that modulates vergence, translation, or scale of video see-through cameras to give users on-demand peripheral vision enhancement. Unlike other methods that modify a user's direct field of view, we take advantage of ultrawide fisheye lenses to provide access to peripheral information that would not otherwise be available. In a series of experiments testing our prototype in real world search, identification, and matching tasks, we test these expansion methods and evaluate both user performance and subjective measures such as fatigue and simulation sickness. Results show that less head movement is required with dynamic view expansion, but performance varies with application.

Categories and Subject Descriptors (according to ACM CCS): Multimedia Informaton Systems [H.5.1]: Artificial, augmented, and virtual realities—; User Interfaces [H.5.2]: Evaluation / methodology—

1. Introduction

Augmenting the human visual system has been the goal of vision and augmented reality (AR) researchers for years. With recent advancements in display technology and optics, we now have the ability to modulate background information to improve or enhance human performance in a variety of scenarios. One important type of enhancement is that of field of view (FoV) expansion. This is relevant in many areas, such as search and rescue, cycling, and for head mounted displays (HMDs) that occlude a portion of the user's natural field of view. Unlike optical see-through (OST) displays, video see-through (VST) displays are especially suited to this task since they can manipulate characteristics of the background image. Several models for "always on" vision expansion have already been proposed, such as FlyVIZ, SpiderVision, and Fisheye Vision [ALM*12][FHNI14][OWK*14].

However, one common characteristic of all these displays is that the modulated field of view is either always on, or activated in a binary on-off fashion. This can inhibit use of the device and degrade esoteric motion perception [OWK*14], especially in the periphery. To help overcome this problem, we have implemented several methods for dynamically increasing peripheral vision based on head movement. To accomplish this, we made use of wide field of view lenses coupled with a video see-through version of the Oculus Rift DK2 as shown in Figure 1, but we expand the view algorithmically rather than statically.

Methods tested in our experiments in addition to the standard one-to-one VST view include minification, diverging (a variant of one of the methods from [STG12]), and shifting, which are described in more detail later. The expansion factor in all of these methods is modulated by head rotation. To test these modes in a variety of situations, we conducted two experiments. The first experiment included a test of the expansion methods in a seated area where participants had to identify colors on a series of surrounding displays. Participants searched for objects of certain colors on the displays by rotating on a swivel chair, and marked their answers using a controller. The second experiment was a mobile task in which participants had to walk around a room, identify and match cards, and then physically sort matched cards in various locations about the room. These experiments allowed us to test a number of metrics for each individual technique, and showed that less head movement was required for dynamic view expansion in the seated task and that no significant differences in mental workload occurred in comparison with standard video see-through.

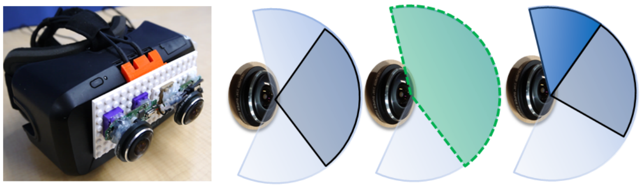

Figure 1: Image of our prototype display, which utilizes wide field of view (FoV) lenses to manipulate peripheral information (far left). Next, diagrams of peripheral FoV gains with respect to method: original fisheye (light blue), undistorted fisheye (gray with black border), gains for dynamic *minification* and *diverging* (dotted green), and *shifting* modification (dark blue).

---

† yano.yuki [at] lab.ntt.co.jp
‡ orlosky [at] lab.ime.cmc.osaka-u.ac.jp
§ kiyo [at] ime.cmc.osaka-u.ac.jp
¶ takemura [at] ime.cmc.osaka-u.ac.jp

DOI: 10.2312/egve.20161435

2. Related Work

Limitations on FoV in HMDs are known to affect a number of factors, including performance, presence, simulation sickness and safety [Art00]. Up to now, a number of works have studied these effects. For example, limiting the FoV will result in deterioration of motor skills [Ish88] but also diminish simulation sickness [FF16].

More recently, many techniques have been proposed that allow for optical or algorithmic (effective) expansion of FoV. For example, Ardouin et al. propose FlyVIZ [ALM*12] for view expansion that provides a 360 degree view of the user's real world surroundings using an catadioptric camera and HMD, though this type of compression significantly distorts field of view. Fan et al. propose SpiderVision [FHNI14], which dynamically overlays the user's rear view upon detection of scene changes, e.g. moving objects. These studies provide a modified or wider FoV than the normal human field of vision. However, such view expansion may distract the user because it is either always on, occludes the user's FoV, or is activated regardless of intention. Some display systems such as the dome screen-based CYLINDA [NYKY03] and the HMD-based Fisheye vision [OWK*14] retain the central FoV while displaying or expanding a view of the periphery using a combination of optics and non-linear compression. These displays support natural stereo vision in the central FoV, but spatial perception can be difficult due to constant peripheral distortion.

In a VR environment, Debarba et al. [DPHB15] explored the impact of non-planar projections and found that a wider, distorted FoV yields shorter search time with degraded quality of interaction. However, these effects would not necessarily be present if the view expansion method was activated dynamically. Several studies have explored view expansion in VR that focus on amplification or expansion techniques. For example, Jay et al. [JH03] found that head motion amplification can increase search performance and that amplified head motion is perceived as more natural. Yano et al. [YKS*14] compared the impact of a number of head based view expansion techniques in a VR environment, and found that some types yield better search performance than others.

In the case of AR, mismatch of characteristics such as viewing scale and orientation between the virtual backdrop and real environment are more problematic. Moreover, differences in latency must be taken into account in video see-through AR in comparison with VR, and the effects of view expansion will have different results on performance and simulation sickness. To study these characteristics and test the effectiveness of our system, we designed experiments to evaluate the impact of dynamic view expansion on real world visual search.

3. Framework for Dynamic View Expansion

To implement our view expansion methodology, we have developed algorithms designed for use in AR. Implementations in VR, such as that of [YKS*14], are relatively straightforward since it is possible to manipulate the 3D environment as desired. In contrast, the wide FoV lenses we utilize for AR must be undistorted and properly mapped to the user's field of view. The view expansion algorithms, two of which are based on [STG12] and [YKS*14], are implemented for AR and tested as described below.

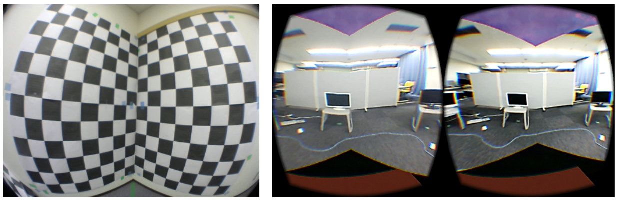

Figure 2: Image showing the wall-sized checkerboards used to calibrate the image planes (left), and the resulting one-to-one view in which video see-through images are undistorted (right).

3.1. Head Mounted Display Setup

The head mounted display we use is the Oculus Rift DK2. Video see-through functionality is achieved using Logicool C310 web cameras with 238° wide FoV lenses as can be seen in the prototype on the left of Figure 1, which is based off of the implementation in [OWK*14]. These camera images are rendered as textures in the display, though in contrast, our model uses two separate planes to render a single image. These planes can be seen on the left of Figure 2, which also shows the calibration boards used to undistort and align both planes. Whereas a one-to-one representation of the world would limit the user's vision to about 90°, we have access to an additional 148° degrees of FoV that we can dynamically manipulate, as described next.

3.2. Minification

The first method we implemented for AR was that of minification, which decreases virtual image size, but increases the amount of viewable content. In contrast with the VR case, where the scene camera is simply moved backwards, in AR, we decrease the size of the object showing the video background textures. The result is slightly different since the slightly distorted outer edges of the wide FoV lens become visible in the periphery, but the improvement to FoV is relatively similar.

3.3. Shifting

Motion amplification (referred to as shifting in this paper) is another technique in VR to show more information in the periphery in the direction the user is moving [JH03]. The users' virtual view rotates more to the direction of rotation than the rotation of the display when turning, thus combining body-head-eye rotation and allowing users to search a wider area more quickly. In our AR implemenation, we move the virtual eyes back to the real eye pose after 500ms to match decreases in head motion.

3.4. Diverging

The diverging algorithm implements a type of “diverging-prey" vision by increasing or decreasing the vergence of each camera image. Typically, vergence in almost all HMD models is fixed when in use. For our diverging method, we rotate the virtual eyes outwards, which will actually give the user a wider total field of view, though depth perception will be slightly distorted by the change in vergence. We hypothesize that despite changes in depth perception, search performance will show an increase due to the temporarily improved FoV.

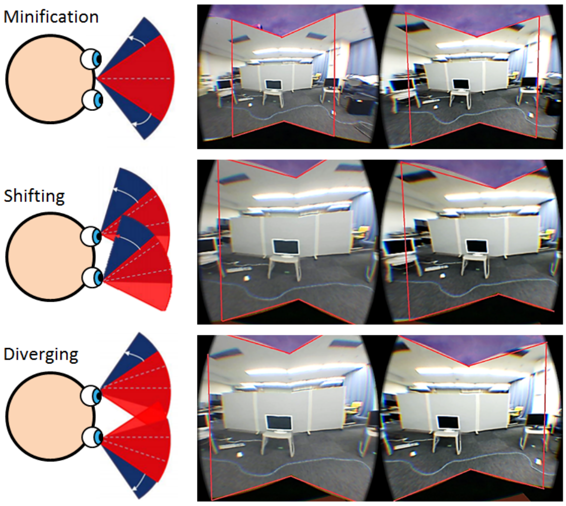

Figure 3: Logical representations of each of the three methods (left) and a view through the display showing actual visual expansion, including *minification* (top), *shifting* (middle), and *diverging* (bottom), techniques. Red represents the original FoV.

3.5. Parameter Optimization and Selection

Initially, we tested three parameters in informal experimentation to define the behavior of our expansion mechanism. The first parameter is the minimum head rotation (v_{start}) required to activate view expansion, which is used to prevent engagement during small or slow head movements. The second parameter is the upper limit (U_{Lim}) of the degree of view expansion, which increases linearly with head rotation speed. Based pilot tests with several users, high-speed expansion would likely distract users from tasks, so the degree of expansion was capped to the thresholds specified below. The view returns to the normal after these method-specific delays. The third parameter is return time (t_{return}) , which specifies how long before returning to normal after head movement returns below v_{start}. Parameters for v_{start}, U_{Lim} and t_{return} are as follows:

- minification: v_{start}: 40°/s, U_{Lim} 20°, t_{return}: 500*ms*.
- shifting: v_{start}: 40°/s, U_{Lim} 10°, t_{return}: 200*ms*.
- diverging: v_{start}: 40°/s, U_{Lim} 7°, t_{return}: 200*ms*.

4. Experiments

To test the effectiveness of our view expansion methods, we designed and conducted two AR experiments to test the three different modes, as described in the previous section. These modes were tested in two different environments, including a seated task where participants were surrounded by target monitors, and a standing/walking task where participants had to physically move and sort cards containing patterns.

4.1. Setup and Tasks

Our experiments were designed to evaluate how the different view expansion methods would perform in a 360 degree viewing environment. This resembles tasks common in a number of areas, such as vehicle navigation, surveillance, and environmental safety.

Figure 4: Image of the environment for the seated search task (left). Subjects rotated on the swivel chair to find the monitor containing a target color and recorded answers using a hand-held controller. Image of the setup for the mobile search task (right). Subjects traversed the room searching for target cards to place onto matching markers on separate tables.

4.1.1. Experiment 1: Seated Search and Identify

Experiment 1 was a seated task where participants had to find and identify colors on a series of displays that completely surrounded them in a 2.0m circle, as seen in the left image in Figure 4, and as represented on the top of Figure 5. Target objects were 160mm × 160mm squares that were blue, red, yellow, or green. These appeared on one of the eight displays among black squares of the same size shown on the remaining seven displays.

Participants used a game controller to record the color of the object with a corresponding button, and the next object was displayed. One of the displays was selected as a "home position," where the initial target object appeared at the beginning of each trial. The target objects were shown on one of the other seven displays in randomized order, which was repeated five times for each color, resulting total 70 trials per session. A total of 9 participants, 7 male and 2 female (average age 24.9, stdev 3.01) took part.

4.1.2. Experiment 2: Mobile Search and Match

Experiment 2 was a mobile task in which participants had to walk around a room, identify and match cards, and then physically place matched cards in various locations about the room, as shown on the right of Figure 4. Figure 5 illustrates a sample start configuration. Participants were first positioned in the middle of a 3.3m × 4.5m space surrounded by desks on which several A4-sized cards with different symbols and different colors were placed. They had to find matches between the sets of cards, stack them together, and repeat the process for all cards. They were forbidden from picking up more than one card at a time to ensure consistency. 8 subjects (7 male, 1 female, mean age 24.75, stdev 2.43) participated.

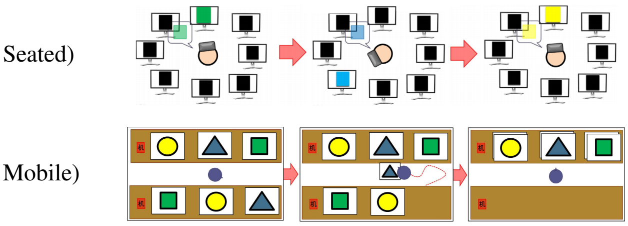

Figure 5: Block diagrams representing sample layouts for the seated search task (top) and matching task for the mobile setup (bottom). For seated tasks, one colored object was shown on a single monitor at a time. For mobile tasks, participants were required to match sets of cards between tables, one by one.

4.2. Test Conditions

We tested a total of four conditions (no expansion, minification, shifting, and diverging) for use in our real world tasks. The amounts of view expansion for each of the three view expansion techniques were determined by both pilot testing and logical selection so that search performance is increased, while level of discomfort is minimized. We use a within-subject design, i.e., each participant repeated the four conditions five times. The order of the conditions was randomized for each participant using a Latin square distribution and participants completed a NASA-TLX questionnaire [HS88] after each task. We analyzed task completion time and head rotation, as well as subjective measures of comfort. These were analyzed with one-way ANOVA and MSRB for post-hoc analysis.

4.3. Experiment 1 Results

On average, task completion time in minification and shifting conditions for the seated experiment were slightly smaller than the None condition. However, no statistically significant difference was found ($F_{(2,8)}$= 1.66, p= 0.206). Head rotation for shifting was smaller than for None and diverging conditions ($p < 0.01$), and minification was smaller than for None ($p < 0.05$), as shown in Figure 6. This suggests that users can search for targets more quickly with minification and shifting view expansion techniques than in a standard video see-through mode.

4.4. Experiment 2 Results

Surprisingly, dynamic expansion had little effect on mobile tasks. On average, task completion time for the shifting and diverging conditions were slightly smaller than that in None condition. However, no statistically significant difference was found ($F_{exp2(2,7)}$= 1.68, p= 0.201) Regarding head rotation, the None condition had the smallest total and shifting was the largest. However, no statistically significant difference was found either (F= 0.798, p= 0.509)

No statistically significant difference was found for subjective measures in the NASA-TLX mental task load in either experiment ($F_{exp1(2,8)}$= 0.817, p= 0.499), ($F_{exp2(2,7)}$= 1.15, p= 0.352), which suggests that our view expansion techniques should not cause more fatigue or simulation sickness than typical video see-through.

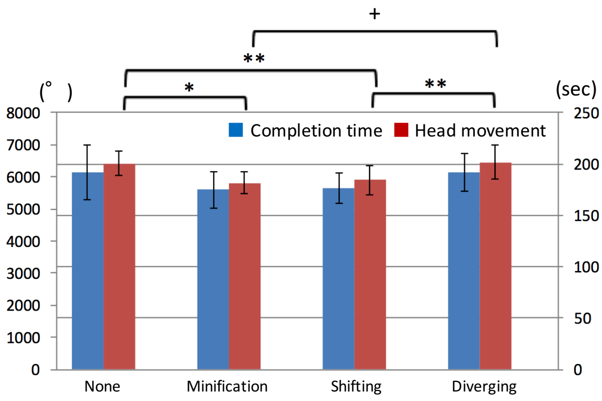

Figure 6: Performance results for the seated search and identify task. Red shows head movement in degrees (left axis), and blue shows completion time in seconds (right axis).

5. Discussion and Conclusion

For the seated experiment, the minification and shifting techniques showed an improvement in performance, but interestingly, no difference was found in the mobile experiment. This suggests that dynamic is heavily application dependent. It is likely that applications requiring low body rotation but high head rotation like driving, cycling, or ornithology will benefit more from on-demand expansion.

In conclusion, we presented a number of different techniques for expanding a user's natural field of view with respect to head movement. These take advantage of wide field of view lenses that allow us to access and manipulate peripheral information more effectively. We then set up a series of experiments including both seated and mobile tasks to test the view expansion techniques. Results showed that view expansion has a positive effect on performance that varies depending on task. We hope that these view expansion techniques will influence future methods for improving FoV.

References

[ALM*12] Ardouin J., Lécuyer A., Marchal M., Riant C., Marchand E.: Flyviz: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology* (2012), ACM, pp. 41–44.

[Art00] Arthur K. W.: *Effects of field of view on performance with head-mounted displays*. PhD thesis, UNC at Chapel Hill, 2000.

[DPHB15] Debarba H. G., Perrin S., Herbelin B., Boulic R.: Embodied interaction using non-planar projections in immersive virtual reality. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (2015), ACM, pp. 125–128.

[FF16] Fernandes A. S., Feiner S. K.: Combating vr sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)* (2016), IEEE, pp. 201–210.

[FHNI14] Fan K., Huber J., Nanayakkara S., Inami M.: Spidervision: extending the human field of view for augmented awareness. In *Proceedings of the 5th Augmented Human International Conference* (2014), ACM, p. 49.

[HS88] Hart S. G., Staveland L. E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology 52* (1988), 139–183.

[Ish88] Ishigaki H.: On the influence of the masking of the peripheral vision on the motor skill. *Bulletin of the Aichi Institute of Technology, Part A. 21* (1988), 268–276.

[JH03] Jay C., Hubbold R.: Amplifying head movements with head-mounted displays. *Presence: Teleoperators and Virtual Environments 12*, 3 (2003), 268–276.

[NYKY03] Nagahara H., Yagi Y., Kitamura H., Yachida M.: Super wide view tele-operation system. In *Multisensor Fusion and Integration for Intelligent Systems, MFI2003. Proceedings of IEEE International Conference on* (2003), IEEE, pp. 149–154.

[OWK*14] Orlosky J., Wu Q., Kiyokawa K., Takemura H., Nitschke C.: Fisheye vision: peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2nd ACM symposium on Spatial user interaction* (2014), pp. 54–61.

[STG12] Sherstyuk A., Treskunov A., Gavrilova M.: Predator-prey vision metaphor for multi-tasking virtual environments. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on* (2012), IEEE, pp. 81–84.

[YKS*14] Yano Y., Kiyokawa K., Sherstyuk A., Mashita T., Takemura H.: Investigation of dynamic view expansion for head-mounted displays with head tracking in virtual environments. In *Proc. of the 24rd International Conference on Artificial Reality and Telexistence (ICAT 2014)* (2014), Citeseer.