

Real Binoculars with Virtual Functions for Mixed Environments

Andrei Sherstyuk
Telehealth Research Institute
University of Hawaii
andreis@hawaii.edu

Kin Lik Wang
Telehealth Research Institute
University of Hawaii
kinwang@hawaii.edu

Anton Treskunov
Institute for Creative
Technologies
University of Southern
California
treskunov@ict.usc.edu

Jarrell Pair
myAWOL.com
jarrell@acm.org

ABSTRACT

Though often desirable, the integration of real and virtual elements in mixed reality environments can be difficult. We propose a number of techniques to facilitate scene exploration and object selection by giving users real instruments as props while implementing their functionality in a virtual part of the environment. Specifically, we present a family of tools built upon the idea of using real binoculars for viewing virtual content. This approach matches user expectations with the tool's capabilities enhancing the sense of presence and increasing the depth of interaction between the real and virtual components of the scene. We also discuss possible applications of these tools and the results of our user study.

This paper is an extended version of earlier work presented at the 4th International Workshop on the Tangible Space Initiative[5].

Categories and Subject Descriptors

H.5.1 [Information Systems]: Information Interfaces and Presentation—*Artificial, Augmented, and Virtual Realities*

1. INTRODUCTION

One of the trends in user interface (UI) design for mixed and augmented reality systems is the use of tangible interface metaphors where users interact with the environment with the help of real props [1]. Tangible interfaces blur the border between the real and the virtual by adding a tactile modality to interaction. Two examples illustrate the versatility of this approach. The *Tiles* system [2] provides a means for building general purpose interfaces by manually arranging physical tiles that represent operations and data. *MagicBook* [3] is a highly interactive tool in the shape of a

real book where fictional characters come to life as 3D objects when users turn the pages. In this case the interface is specifically designed to meet the goals of the application. It capitalizes on the perfect match between the appearance of the interface device (the book itself) and its expected function (a story-telling object).

We propose several techniques that, like the *Magic Book*, take advantage of the close match between the shape of the interface object and the effect of its application. Some of these techniques were mentioned briefly as extensions of the *optical sight* metaphor, originally developed for purely virtual environments [4]. In what follows, we continue exploring these ideas and describe how they can be applied in systems where virtuality and reality are both present. In section two, we outline our goals and place them in the perspective of related work. Sections three and four describe the proposed methods and extensions. Full details on the implementation and user evaluation are given in the last two sections.

2. MOTIVATION AND RELATED WORK

Our goal is to assist users in performing two basic interaction tasks: scene exploration and object selection. Both the complexity and the size of the scene (or *perceived* size) play significant role in UI design.

In scenes of low complexity and size, as in the table-top *AR²Hockey* game [6], all objects are in clear view and in close proximity which allows interaction to be direct and unassisted. A implementation of the virtual hand metaphor, in its basic [7] or extended form [8] could be appropriate in this case.

The *MagicLens* metaphor employs handheld interface devices shaped as a looking glass [9] which provides visual and semantic zoom into virtual scenes of high complexity. It works very well for physically compact scenes such as a 3D model of a house or small building for example. However, when users are placed in virtual settings with hundreds of life-size objects spread over hundreds of meters, as in the *FlatWorld* system [11], or *MR MOUT* training simulation [10], long range access enabling tools must be provided.

In one of the *FlatWorld* applications, a physical room with real furniture is augmented with a virtual urban combat scene displayed on rear-projection screens situated in window frames. Soldiers trainees are required to survey a vast city scene and react to circumstances. Lighting conditions and the level of hostilities vary from scenario to scenario. Long range vision enhancement tools are critical to their tasks and should be analogous to the binoculars, night vision goggles, and laser range finders they use in the real physical world.



Figure 1: A virtual desert scene projected on two screens is merged with the real room environment's prop window and doorframe. In this setting, a pair of binoculars could help to survey the scenes.

3. REAL BINOCULARS, VIRTUAL SCENE

In real life, vision enhancement can be achieved by any combination of the following methods: increasing magnification, choosing an alternative input signal range, and applying various filters. The corresponding devices are binoculars, night vision devices, and sunglasses with polarizing, ultraviolet, and other types of filters. In virtuality, these methods are replicated by using variable camera zoom, alternative shading models, and changing material parameters. For example, night vision may be implemented by storing and displaying surface temperature data as vertex colors. Reducing the value for Phong highlights on reflective surfaces mimics the effect of applying a polarized filter. Adding transparency can provide an X-ray imaging effect.

We propose using a pair of conventional binoculars as interface in spatially immersive display systems. The binoculars are used to control a virtual camera and to manipulate the rendering and shading of 3D objects in the scene. When placed next to eyes and directed towards the display screen, the binoculars can be used to control camera's orientation and zoom level. The lenses and prisms are removed from the binoculars, because their functionality is now performed in software. The remaining empty shell with rubber eyepieces serves only to block peripheral vision and provides tactile confirmation that the binoculars are in use. It is essential for the display screen to be large enough to provide a comfortable range of viewing angles.

By adding a wheel control and a few extra buttons, the baseline model of virtual binoculars can be significantly enhanced, as described in the next section.

4. EXTENSIONS AND APPLICATIONS

Because of its look and feel, a black box with two video channels, a pair of binoculars can be easily perceived as a general purpose view modifying instrument. Thus, it lends itself well to various modifications and extensions. Some of these are listed below.

- *Crosshair pointer.* By tracking the position and orientation of the binoculars, we can use it as a pointing device. A crosshair image sliding over the magnified view acts as a pointer. In conjunction with a simple form of ray casting, this option turns the binoculars into an optical sight which may be used for many purposes, including object selection and access. Variable zoom simplifies the selection of small, distant, and partially occluded objects, as shown in Figure 2.
- *X-ray vision.* Most users are familiar with the concept of X-ray vision, and it has already been introduced to AR and MR systems [12], both stationary [13] and mobile [14]. We implemented an X-ray vision mode by temporarily adding transparency to 3D objects in the virtual scene, as shown in Figures 2 and 6. One screen snapshot taken with the X-ray-enabled binoculars is shown in Figure 2, bottom-right; another example is shown in Figure 6.
- *Night-vision.* Night or thermal vision, displayed in gray or green monochrome or in pseudo-color, is another well known technique. It is widely used in video games, such as the *Splinter Cell* series [15]. We have not yet explored this option.
- *'Reality Freezer'.* 'Reality freezer' or snapshot mode allows one to pause virtual content by skipping all motion control code in the main graphics loop. Consequently all 3D virtual objects and characters freeze in place, tracking stops, and time does not advance. This mode is very convenient in searching for small objects within a single frame, counting objects, taking screen snapshots, and similar tasks which could be enhanced by a temporarily static virtual world. We implemented this feature very early in our development process.
- *LOD control.* When zooming in on distant objects, their polygonal nature becomes apparent, especially on object edges (Figure 5, left). This rendering artifact is undesirable in most applications. Virtual binoculars offer a convenient way of controlling the level of geometric complexity of the objects that fall into view. In a previous paper, we discussed potential applications of LOD-control for various VR and MR systems [4]. In this work, we tested it with a 3D model of a dolphin, as shown in Figure 5. By switching to a high-polygon version along with increasing camera zoom, the objects in view maintain their visual quality.

Note that all methods described above control the *virtual* component of a VR or MR system. That distinguishes our work from a family of vision-enhancing devices that augment *real* views with computer-generated content, such as coin-operated telescopes [16] and systems for astronomical observations [17].

5. PILOT TESTS: COCONUT SHOOTING EXERCISE

Before building a physical device, we conducted pilot tests on a desktop VR system, derived from Flatland [18], with a mouse-and-keyboard interface. In this exercise, the objective was set to find and shoot down large ripe coconuts on distant palm-trees, while avoiding hitting small ones, as illustrated in Figure 2. The coconuts were packed in tight bunches, which made precise pointing a very difficult task. In addition, large and small coconuts were occluding each other in most cases. No particular scoring system was involved, the purpose of this test was to get the ‘feel’ of the whole system.

For this exercise, we implemented the basic version of the optical sight metaphor [4], briefly outlined below. The view direction was controlled from the keyboard. Pointing at and selection of objects was implemented via ray-casting, with the probing ray originating from the camera and extended towards the location of a virtual hand. The position of the hand was controlled by a mouse and displayed as a small white circle. The system had a variable zoom, ‘reality freezer’, and X-ray vision mode, described above.

In this scene, the palm trees were modeled as articulated skeleton-based object clusters and animated with irregular swinging motions, imitating wind. That made unassisted aiming practically impossible. However, with all the enhancements in place, the system performed well. Zooming-in allowed unambiguous selection via ray casting. Reality freezer canceled the effect of simulated wind and made the trees stand still during aiming and shooting. X-ray vision helped to detect situations when small coconuts were occluded with larger targets. Figure 2 shows details on harvesting coconuts from one of the palm trees.

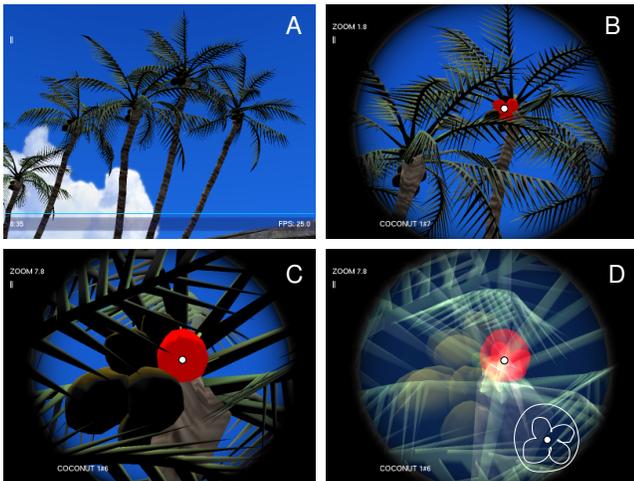


Figure 2: Coconut shooting exercise. Objects hit by a probing ray are highlighted in red. (A) original unassisted view; (B) low zoom value 1.8 results in multiple selections; (C) high zoom value allows reliable selection; (D) X-ray vision reveals hidden four small objects hit by the same ray (also contoured in white).

6. IMPLEMENTATION, HARDWARE AND SOFTWARE

To implement our virtual binoculars, a pair of Simmons 50x10 binoculars were fitted with two hollow plastic tubes that people looked through. For this particular model, it was nearly impossible to remove the lenses and use the original eyepieces for direct viewing. As a result, we had to use the additional tubes.

The binoculars were operated with an attached miniature mouse. The left mouse button switched the binoculars on and off, the middle button toggled an X-ray vision mode, and the right button toggled the ‘Reality Freezer’ and took screen snapshots. The mouse wheel changed zoom level between 1 and 40. The starting value of zoom was set to 1 which allowed seamless transition between views when the binoculars were turned on. The whole unit was tracked with an *InertiaCube 2* tracker for controlling the camera ‘look’ direction. The fully assembled binoculars are shown in Figure 3.

The 3D content was rendered at 25 FPS on a single PC with a 3.2 GHz CPU, 1GB RAM, and a nVidia Quadro NVS GPU, running customized *Flatland* engine [18]. Video and audio content was rendered with the OpenGL and OpenAL libraries. Dynamic sound localization was processed on the same PC.



Figure 3: A pair of real binoculars with additional equipment for controlling virtual views. Mini-mouse buttons and wheel operate zoom, X-ray vision, snapshot mode, and ‘reality freezer’. Intersense InertiaCube 2 tracker is attached to the top. At the bottom, an additional pair of viewing tubes can be seen.

7. USER EVALUATION

The system was tested with 20 volunteers, one at a time. Most volunteers were in their twenties or thirties, all had normal or corrected-to-normal vision. The experimental protocol was approved by the Committee on Human Studies of the University of Hawaii. Each participant was taken on a ten minute cruise around a virtual island as shown in Figure 4. The 3D content was displayed on an 8 x 6 foot

rear-projection screen. The participants were seated very close to the screen approximating the experience of being in a vehicle cockpit. The position and orientation of the virtual boat were updated automatically so the participants did not have to worry about navigation. While standing on the moving boat which was always oriented towards the center of the island, people could redirect their view approximately 90 degrees horizontally and 45 degrees vertically. For this particular setting, orientation-only tracking of the binoculars was quite sufficient, as the relative displacement of the camera position (a few inches) was negligible compared with the size of the virtual scene (several hundred feet).

The user's goal was freestyle exploration of the island which included a variety attractions to keep one interested for the duration of the cruise. Birds, butterflies, dolphins, crabs, and other tropical creatures were present. Participants could use the binoculars at will, with all its extensions as described in the previous sections. An informal competition to take the best pictures of the island's wildlife was announced. Incidentally, this component made the whole exercise somewhat similar to the once popular *Pokemon Snap* game for the Nintendo 64 video game console.

After returning from the trip, participants were asked to fill out a short survey about their background, such as gaming habits, previous VR/MR experiences, and how often they use real binoculars in everyday life. We also asked people to evaluate the virtual binoculars on a scale from 0 to 5 by answering the following questions:

For this particular scene, the virtual binoculars were:

- useful as a tool? (0-totally useless, 5-indispensable)
- easy to operate? (0-frustrating, 5-intuitive)
- enjoyable overall? (0-annoying, 5-very enjoyable)

A few words about what you didn't like...

And what you did like (if anything)...

To encourage constructive criticism, we explicitly asked people what they did not like first. Free form suggestions concluded the survey.

7.1 Quantitative analysis

Mean values of the ratings are summarized in Table 1. At a glance, the virtual binoculars were given fairly high scores in all questions by all groups of participants. Within 95% confidence intervals, people rated the virtual binoculars as

useful as a tool: 3.4 – 4.3

easy to operate: 3.4 – 4.2

enjoyable overall: 4.0 – 4.7

To determine if a participant's background had significant influence on their evaluation, we performed the Welsh two-sample t-test grouping people as gamers, non-gamers, VR-experienced, VR-novices, experienced with binoculars, and



Figure 4: On a bird-watching mission, most players never turned the binoculars off for the whole duration of the trip.

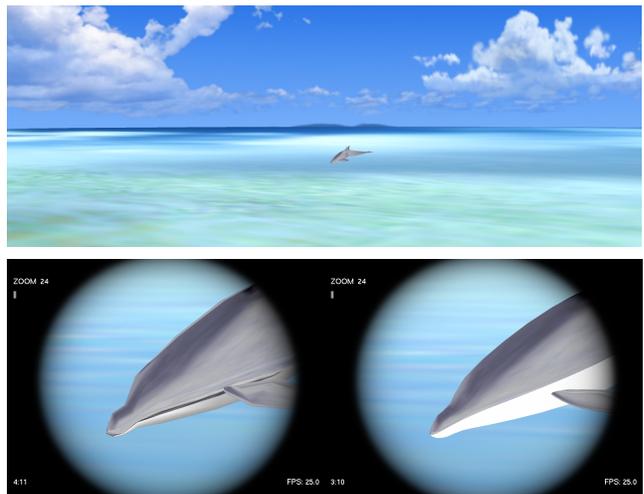


Figure 5: Zoom-controlled LOD. Top: panoramic view with a dolphin rendered in low-polygon form. Bottom: 24x magnification. The initial model (270 polygons, left) is replaced by a high-quality version (1440 polygons, right). The black round frame is normally turned off.

not experienced with binoculars. As the results show, there was no significant difference observed except for the question on usefulness, with $P = 0.03$, as answered by gamers and non-gamers. Another special case of $P = 0.98$ was observed in answers to the ease of use question given by people who had and had not previously experienced VR. To confirm these findings, we also performed a Wilcoxon rank sum test. The resulting P values together with observation medians are presented in Table 2. From both tests, we conclude that:

- Participants found the binoculars useful, easy to operate, and very enjoyable.
- Those participants who use real binoculars in everyday life found the device less useful than those who do not. We think that the reason for this result lies in the rather crude design of the device prototype. For users accustomed to the real thing, the resemblance breaks down.
- Previous VR experience or lack of it does not influence ratings on ease of use of the proposed device.
- There is no evidence that gaming habits have significant influence on the results of the evaluations among the participants.

7.2 Qualitative evaluation

Informal observations showed that people generally enjoyed playing with the virtual binoculars and did not show any signs of being bored or tired.

Surprisingly, of all 20 participants, only one person noticed that it is possible to ‘cheat’ with this device. Instead of holding the binoculars close to the eyes, one can just put it on the table and use it as a joystick to direct the view. All other people simply accepted the rules of the game and treated the binoculars as the real thing, diligently looking through the empty tube, as can be seen in Figure 4. We interpret this observation as complimenting the integration between the binoculars, the user interface, and the overall gameplay.

Another interesting observation is that most people practically never turned the binoculars off even when they used it at zoom level 1 (no magnification). This point supports the utility of the device and indicates that users were very comfortable with it. The free form written comments are summarized in Table 3.

8. CONCLUSION

In this paper we presented a number of techniques, based upon the idea of using real binoculars as an interface object to facilitate the viewing of virtual content. These techniques are simple and intuitive to use. Their implementation is straightforward and should be easy to integrate into most systems.

Enhancing users’ vision in mixed reality environments improves their ability to reliably point at and select small, distant, or partially occluded objects. Effectively, we increase the resolution of the user interface apparatus which enhances

Table 1: Mean values and Welsh two-sample t-test P values of evaluation ratings among different groups. Special cases are marked with an asterisk *.

Participants groups and sample size (N)	Useful as a tool	Easy to operate	Overall enjoyable
Proficiency in gaming:			
Regular or casual (N=12)	4.2	4.0	4.5
Not a gamer at all (N=8)	3.7	3.7	4.1
P	0.33	0.5	0.34
Had VR/MR experiences:			
Yes (N=11)	4.1	3.8	4.5
No (N=8)	3.7	3.8	4.1
P	0.48	0.98*	0.31
Use of real binoculars:			
Once in a while (N=13)	3.7	3.9	4.3
Almost never (N=6)	4.5	3.7	4.2
P	0.03*	0.61	0.64
All participants (N=20)	4.15	3.85	4.35

Table 2: Median values and Wilcoxon rank sum test.

Participants groups and sample size (N)	Useful as a tool	Easy to operate	Overall enjoyable
Proficiency in gaming:			
Regular or casual (N=12)	4	4	5
Not a gamer at all (N=8)	4	3.5	4
P	0.59	0.39	0.28
Had VR/MR experiences:			
Yes (N=11)	4	4	4.5
No (N=8)	4	4	4.25
P	0.85	1.0*	0.4
Use of real binoculars:			
Once in a while (N=13)	4	4	4.5
Almost never (N=6)	4.5	3.5	4
P	0.05*	0.55	0.61
All participants (N=20)	4	4	4.5

Table 3: Free form comments. Quoted text is reproduced verbatim.

Complaints	Compliments
binoculars too heavy	X-ray vision
drift in the tracker	interactivity, sense of adventure
shaky vision at high zoom levels	the novelty of the idea
slow zoom increment, must be progressive	“the comfort of the binoculars, shape and weight
when placed on the table, should switch itself off	“it’s nice to get caught up in the experience”

the level of user control over the scene. Direct and immediate tactile feedback from the interface objects adds to the overall sense of presence. Though not addressed in this paper, flashlights and headlamps are additional real world tools with potential as tangible interfaces.

Our user evaluation study shows that people easily grasp the idea of virtual binoculars, regardless of their proficiency in gaming interfaces and previous exposure to virtual and mixed environments. Favorable informal reports on usability and comfort are also very encouraging.

We believe that a large gamut of applications may benefit from our methods. These applications include professional skill training systems, educational tools, and interactive entertainment.



Figure 6: Top: approaching the Bird Island, panoramic view. Bottom: magnified view with a crosshair pointer for object selection (left); X-ray vision mode shows birds hidden behind the rock (right).

9. ACKNOWLEDGMENTS

We wish to thank Mike von Platen and Kristen Okahashi for their help with the system setup and testing and all volunteers, for their time and enthusiasm.

Portions of this paper were developed with funds of the Department of the Army under contract number DAAD 19-99-D-0046. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

10. REFERENCES

- [1] D. Bowman, E. Kruijff, J. LaViola, and I. Poupyrev. *3D user interfaces: Theory and practice*, Chapter 12, 2004: Addison-Wesley.
- [2] I. Poupyrev, D. Tan, M. Billinghurst, H. Kato, H. Regenbrecht, and N. Tetsutani. Developing a Generic Augmented-Reality Interface. *Computer* 35(3), March 2002.
- [3] M. Billinghurst, H. Kato, and I. Poupyrev. The MagicBook—Moving Seamlessly between Reality and Virtuality. *IEEE Computer Graphics and Applications*, 21(3), May 2001.
- [4] A. Sherstyuk, J. Pair, and A. Treskunov. Optical Sight Metaphor for Virtual Environments. *Proceedings of IEEE Symposium on 3D User Interfaces*, March 2007.
- [5] A. Sherstyuk, K. Wang, A. Treskunov, J. Pair. View Enhancement Techniques for Scene Exploration and Object Selection in Virtual and Mixed Environments. <http://www.ismar07.org/program/workshops/workshop-1-tangible-space-initiative>, 2007.
- [6] T. Ohshima, K. Satoh, H. Yamamoto, and H. Tamura. AR2 Hockey: A Case Study of Collaborative Augmented Reality. *Proceedings of the Virtual Reality Annual International Symposium*, March 1998.
- [7] L.D. Cutler, B. Frohlich, P. Hanrahan. Two-Handed Direct Manipulation on the Responsive Workbench, *Proceedings of ACM Symposium on Interactive 3D Graphics*, 1997.
- [8] I. Poupyrev, M. Billinghurst, S. Weghorst and T. Ichikawa. Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. *Proceedings of UIST'96* 79-80, 1996.
- [9] J. Looser, M. Billinghurst and A. Cockburn. Through the looking glass: the use of lenses as an interface tool for Augmented Reality interfaces. *Proceedings of Conference on Computer Graphics and Interactive Techniques (GRAPHITE)*, June, Singapore, 2004.
- [10] C.E. Hughes, C.B. Stapleton, D.E. Hughes and E.M. Smith. Mixed reality in education, entertainment, and training *IEEE Computer Graphics and Applications*, November/December 2005.
- [11] J. Pair, U. Neumann, D. Piepol and W. Swartout. FlatWorld: Combining Hollywood Set-Design Techniques with VR. *IEEE Computer Graphics and Applications*, January/February 2003.
- [12] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, 21(6), pp. 34-47, 2001.
- [13] N. Navab, A. Bani-Hashem, M. Mitschke. Merging Visible and Invisible: Two Camera-Augmented Mobile C-arm (CAMC) Applications, *Proceedings of Workshop On Augmented Reality (IWAR)*. San Francisco, October 1999.
- [14] R. Bane and T. Hollerer. Interactive Tools for Virtual X-Ray Vision in Mobile Augmented Reality. *Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, 2004.
- [15] Splinter Cell Double Agent, Official Site. <http://www.splintercell.com>
- [16] B. Lutz, M. Becker, D. Stricker, U. Bockholt. The Augmented Reality Ocular, *Proceedings of ACM SIGGRAPH VRCAI*, New York, 2004.
- [17] A. Lintu, M. Magnor. An Augmented Reality System for Astronomical Observations, *Proceedings of IEEE VR Conference*, March 2006.
- [18] Flatland Open Source Project, <http://www.hpc.unm.edu/homunculus/>