

# A Step Towards a Multimodal AR Interface: A New Handheld Device for 3D Interaction

Raphaël Grasset, Julian Looser, Mark Billinghurst  
HIT Lab NZ

University of Canterbury

Private Bag 4800, Christchurch, NZ

{Raphael.Grasset|Julian.Looser|Mark.Billinghurst}@hitlabnz.org

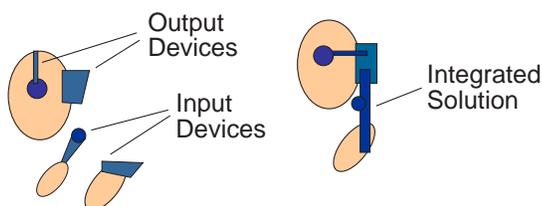
## Abstract

*In this paper, we describe the AR Mask, a novel handheld Augmented Reality (AR) input and display device that not only provides support for a full range of traditional interaction techniques, but also facilitates new metaphors. Unlike other AR display and input technologies, our device consolidates input and output within a single piece of hardware. We demonstrate how this design provides a closed control loop between the user's input modalities and natural sensory receptors.*

## 1 Introduction

With the MagicBook interface, Billinghurst et al. introduced a new kind of handheld AR display, providing the first transitional handheld device and an alternative to HMDs and their limitations [1]. Like Rekimoto's Transvision system[5], this device unified the visualization and the interaction space in the same device. However, interaction in the MagicBook was limited to navigation with a movement rate control and a switch with which to transition between the real and virtual worlds.

We are extending this work by adding the AR Mask; a fully functional device for multimodal interaction. Unlike current displays, the AR Mask will unify all required elements within a single hardware platform, in an attempt to increase interactivity and usability (see Figure 1).



**Figure 1. Multimodal HandHeld Approach: we aggregate components from the traditional approach (left) in one device (right).**

In the next sections, after reviewing related work, we

present our new approach. We then propose interaction techniques supported by the AR Mask.

## 2 Related Work

Hand-held interface research is now a sizeable field, largely due to the proliferation of powerful PDAs and cell-phones, where different work has proposed a spatially-aware approach [2] or background sensing technology [3].

The use of such hand-held interfaces for augmented reality has been demonstrated before. Transvision was a hand-held 3D tracked interface where a user could select and manipulate 3D objects with a ray-based approach [5]. The most similar work to our own is that of AR Pad - a video-see through handheld screen coupled with a Spaceball device for manipulating 3D objects [4]. Unfortunately, These devices generally require two-hand grasping, limiting the use of bimanual interaction techniques.

In contrast, the AR Mask integrates multiple modalities and does not require two-hand grasping. This leaves the user's second hand free for bimanual interaction or secondary tasks.

## 3 Our Approach

We identified three main requirements in the context of multimodal AR interfaces: a large choice of input/output modalities, a closed loop between input and output and awareness of the real world.

*Multiple Modalities:* We identify five main categories of modalities that need to be supported: visual, aural, speech, motor and haptic. The hardware also needs to provide efficient access to both real and virtual content.

*Closed Loop:* The cumbersome aspects of using multiple devices can be removed by integrating these elements into a single piece of hardware. The interface thus becomes more transparent, and provides a simple physical 'externalization' of the interface between the real (physical senses) and virtual worlds (task content). Grouping input and out-

put modalities in a more closed loop may also improve interaction efficiency.

*Real World Awareness:* In contrast to a VR device, the main purpose of a mixed-reality device is to preserve contact with the real world context.

With this in mind, our design requirements are:

- *proprioception awareness:* the device needs to provide an easy way to interact with real world content, and support natural proprioceptive vision of the real world.
- *collaborative presence awareness:* the device needs to support visual and audio cues of collaborators, and to not affect the speech of other users or the hearing of important external sounds.
- *collaborative interaction awareness:* the device needs to provide feedback of user state during interactions with another user. The device needs to provide feedback when the user is acting in the real world (head not behind the display) or when the user quits the application (and releases the display).

### 3.1 Implementation

Our prototype AR Mask (see Figure 2) is based on the first MagicBook interface. We use off-the-shelf components to enhance portability and simplify software development.

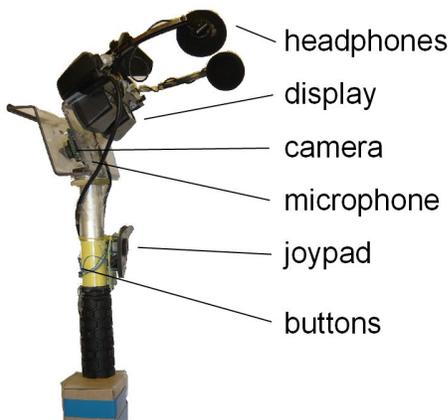


Figure 2. The AR Mask prototype.

A variety of sensors are embedded in the AR Mask. To detect handling by the user we use a pressure sensor mounted within the handle grip. A photosensor in the head-set display indicates when the user is actually holding the mask to their face (this approach has been used previously on Sony Glasstron HMDs).

A small joystick and buttons on the handle can be easily manipulated by the hand used to hold the device. A microphone is used to capture audio for speech recognition or communication, and headphones provide audio output.

## 4 Interaction Techniques

The AR Mask supports a number of different interaction techniques both on its own or with additional props such as a tracked paddle.

We have implemented some interaction techniques for the standard tasks of selection, manipulation, navigation and system control. Some possible selection strategies are shown in Figure 3.

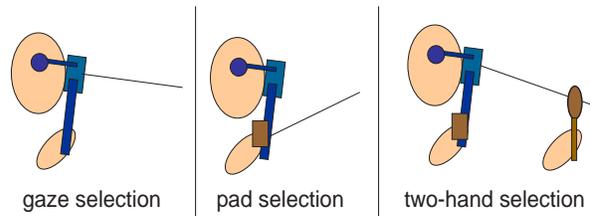


Figure 3. Selection techniques: gaze, joystick or aperture-based.

## 5 Discussion and Conclusions

We have developed the AR Mask, a new multimodal AR input and display device that supports two-handed interaction. Simple interaction techniques have been proposed and initial user feedback shows that the device could be interesting within specific application areas, such as in museum exhibits. User evaluations will be conducted soon to compare the device with an HMD, and the performance of motor skill coordination.

In the future, we hope to extend our device to support tactile or haptic feedback and ambient sensors for a richer experience with mixed-reality applications.

## References

- [1] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook: A transitional ar interface. *Elsevier Computers and Graphics*, pages 745–753, 2001.
- [2] G. W. Fitzmaurice. Situated information spaces and spatially aware palmtop computers. *Communications of the ACM*, 36(7):39–49, 1993.
- [3] K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing techniques for mobile interaction. In *Symposium on User interface software and technology (UIST '00)*, pages 91–100, 2000.
- [4] D. Mogilev, K. Kiyokawa, M. Billinghurst, and J. Pair. Ar pad: an interface for face-to-face ar collaboration. In *Extended abstracts on Human factors in computing systems (CHI '02)*, pages 654–655, 2002.
- [5] J. Rekimoto. Transvision: A hand-held augmented reality system for collaborative design. In *Virtual Systems and Multi-Media (VSMM'96)*.