Virtual Vouchers: Prototyping a Mobile Augmented Reality User Interface for Botanical Species Identification

Sean White[1]

Columbia University

Steven Feiner[1]

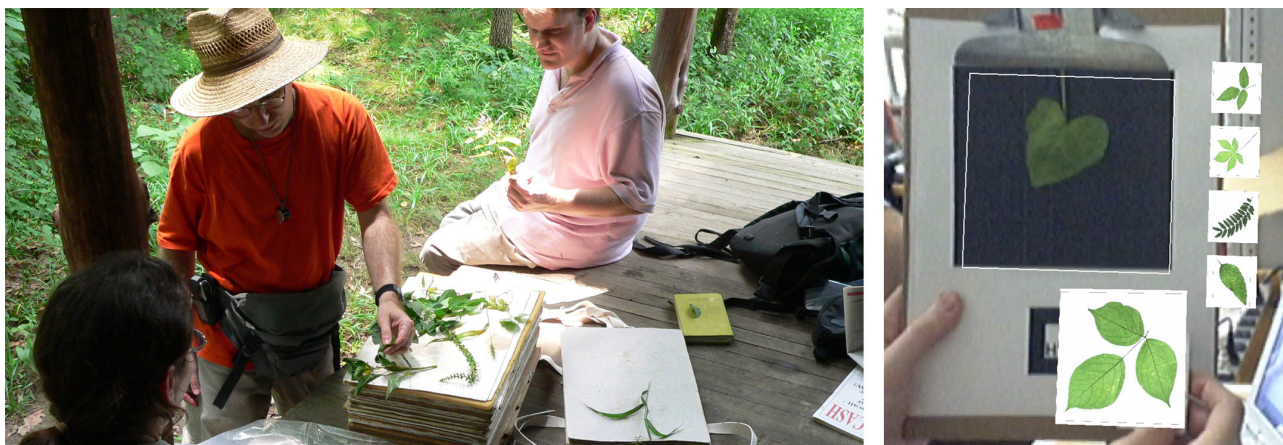Columbia University

Jason Kopylec[2]

Columbia University

Figure 1: (a) Botanists gathering samples in the field. (b) View through a video see-though display of first prototype of the tangible augmented reality user interface.

ABSTRACT

As biodiversity research increases in importance and complexity, the tools that botanists require for field-work must evolve and take on new forms. Of particular importance is the ability to identify existing and new species in the field. Mobile augmented reality systems can make it possible to access, view, and inspect a large database of virtual species examples side-by-side with physical specimens. In this paper, we present prototypes of a mobile augmented reality electronic field guide and techniques for displaying and inspecting computer vision-based visual search results in the form of virtual vouchers. Our work addresses head-movement controlled augmented reality for hands-free interaction and tangible augmented reality. We describe results from our design and investigation process and discuss observations and feedback from lab trials by botanists.

CR Categories: H.5.1 Multimedia Information Systems; H.5.2. User Interfaces; J.3 Life and Medical Sciences

Keywords: electronic field guide, augmented reality, mobile computing, wearable computing, tangible user interface, head-movement control

---

1 email: {swhite,feiner}@cs.columbia.edu
2 email: jkk2106@columbia.edu

1 INTRODUCTION

As the understanding of bioinformatics and biodiversity grows, so must the tools that botanists require for field-work and research. Of particular importance is the ability to identify existing and new species in the field [27]. When botanists in the field need to identify a collected specimen or verify the existence of a new species, they initially consult their own personal knowledge and a paper field guide (Figure 1a). However, paper field guides are difficult to use, do not represent the full corpus of the specimen collection, and do not provide access to species samples (called *vouchers*). Thus, botanists must eventually borrow physical vouchers, with which to compare the specimens they collect, from museums and herbaria, such as the Smithsonian type specimen collection. Because of this, and the fact that vouchers are unique and fragile, the process of obtaining them is time-consuming. In short, botanical research is constrained by availability and access to necessary data.

The opportunity exists to develop a new type of field guide that supports the work flow of a field botanist, combining immediate and holistic access to specimens in the form of *virtual vouchers* and providing even greater information than is found in the current physical voucher. We have been working with colleagues in computer science and botany to develop an electronic field guide (EFG) that addresses this opportunity and acts as a test-bed for exploring new user interface techniques for mobile augmented reality systems (MARS) in an iterative design process (Figure 1b).

The remainder of this paper follows the structure of our user interface design and investigation process. We first discuss related research. Then, we describe the tools, process, and techniques that we observed our botanist colleagues using in the field, which motivated our task analysis, requirements, and design decisions. Next, we present the concept of virtual vouchers and a proposed

workflow of the botanists in the context of an EFG incorporating a computer vision algorithm for identifying plant species. We then introduce three alternative user interfaces developed to address the workflow and explore ways of interacting with the computer vision algorithm and virtual vouchers: a tangible augmented reality (AR) interface, a head-movement controlled AR user interface, and a hand-held user interface with orientation control. Following that, we discuss early experiences and feedback from botanists who have tried the system. Finally, we present our conclusions and plans for future work.

2 RELATED WORK

A number of research projects have investigated EFGs to aid in species identification or fieldwork. The FieldNote system [35] focused on context-aware data collection and Minimal Attention User Interfaces—these were extended by the same team and used by ecologists in Kenya observing giraffes [31]. While not specifically intended for identifying species, the system supported data collection in species observation. Similarly, CyberTracker [1] is a PDA-based system that has been used in a number of fieldwork projects for tracking animals. These systems aid in recording observations from the user. While not designed for fieldwork, Cyberguide [3] addresses a similar goal of providing mobile context-aware information in the form of a tour guide. More recently, ButterflyNet [40] has been developed by Yeh and Klemmer as a mobile capture and access system for biologists to share notes and photos with colleagues. Online resources for species identification have also been explored. Student-built identification keys [38] have been used in the context of educational practices, and an online EFG with XML structuring and visual keys has been developed by Stevenson and colleagues as the enabling tool for "citizen science" models of biodiversity monitoring [37]. In contrast, our EFG investigates computer vision species identification, data collection, and browsing in the field.

3 BOTANISTS AND PLANT SPECIES IDENTIFICATION

As part of our investigation, we joined four botanist colleagues on two collection trips to observe their tools, process, and techniques for gathering plants and data. In preparation for a trip, appropriate tools are gathered and packed in a backpack. The individuals brought their own paper field guides, notebooks, pencils, and cameras. They shared pruning shears, a GPS receiver, collection bags, plant presses, and a list of species that need to be collected.

The botanists then travel to the collection site. When a plant on the collection list is observed, they first identify it using knowledge of the local plants and a paper field guide. Species identification involves inspection of multiple characteristics and comparison of these characteristics with field guide content. Characteristics for comparison include leaf outlines and venation (vein patterns), as well as plant structure, bark (if present), and roots. Leaves can be sufficient for identifying a plant, but closer examination is often required of both the leaf and the plant. In some cases, the plant cannot be identified using the field guide and must be compared with physical plant vouchers in an herbarium.

Once a positive identification has been made, the contextual information from the plant sample is recorded, including location, date, time, and descriptions of the plant and its local environment. The information is gathered into a paper notebook and saved for later use in labeling the specimen. The sample is then pressed in a plant press and brought back to the herbarium. Once it is correctly identified, it will be added to the herbarium specimen collection and serve as a voucher for that particular species.

A voucher herbarium specimen (Figure 2) is a plant sample that has been collected and preserved in a herbarium or specimen collection. The voucher is used for identification of species, and literally acts as the voucher for a species. It can be used for de-

Figure 2: An image of a voucher herbarium specimen. (Photo by Ingrid Lin, courtesy Smithsonian Institution.)

tailed comparison and may even be used for DNA sampling or similar physical testing. The voucher is labeled with information associated with its origins, such as the location and time of collection, name of the person who collected the specimen, and name of the person who made the positive identification. Multiple samples are often collected to create a set of vouchers that represent the diversity of a given species. A given trip may secure a few samples or as many as 25 different collected species. In cases where the location is unfamiliar, the botanist may collect even more.

We observed that the inspection and comparison tasks were particularly critical in finding the appropriate matching species. The sample leaf was often held in the hand and inspected from multiple angles and distances. Discussion was common among botanists trying to identify the species and, in some cases, the species was left unidentified and taken back to the lab for inspection by other experts in the field and, ultimately, for comparison with a voucher.

The inspection and comparison process involved constantly moving and manipulating objects. At the same time, we observed that the botanists often used their hands to move through terrain and inspect plants near the path. Our observations of tools, process and techniques for gathering plants and data along with contrasting requirements for physical manipulation and hands-free interaction motivated our initial design choices.

4 CONCEPTUAL MODEL AND REQUIREMENTS

In developing our conceptual model and requirements for prototype design, we first considered the task of species identification. We deconstructed the identification and collection process into the following subtasks:
- acquiring the unidentified specimen,
- identifying possible species matches,
- iterative comparison with the potential matches and inspection of details and characteristics,
- selection of the matching species, and
- collection and pressing of the specimen and associated contextual data.

In addition to task analysis, we considered the existing set of tools that were familiar to field botanists. In particular, we focused on voucher herbarium specimens and field guides.

4.1 Virtual Vouchers

The Smithsonian Institution maintains a large collection of botanical reference vouchers. Over the past few years, they have been creating a digital database of high-quality, scanned images based on the specimen voucher collection, currently containing over 85,000 images.

We use the term *virtual voucher* to describe a digital representation of the botanical reference specimen in conjunction with its contextual and characteristic data. This data includes additional imagery of the whole plant and root systems, location and date of acquisition, name of collector and of identifier, regional information, articles about the specimen, and links to related specimens. In contrast, previous uses of the term focus primarily on the image of the voucher specimen [2, 30]. The virtual voucher, more generally, acts as a holistic virtual representation for any object that necessarily exists in the physical world.

4.2 Electronic Field Guide Design Criteria

Field guides and flora are an important tool in the botanical research process. In our team's vision of the EFG, a botanist outfitted with a worn or held camera could view a plant to sample it non-invasively, identify the specimen, and see relevant information about the species (including geotemporally situated visualizations of prior collection). The botanist could also gather data for further research and share that data with colleagues in real time.

Although our team's broader aim is to support multiple aspects of the field-work process, our initial goal is to assist botanists in identifying the species of plants found in the field. To this end, the EFG must provide access to the full specimen collection, tools for helping find the closest match, and a means of accessing and inspecting each virtual voucher.

A computer vision algorithm has been developed by our colleagues to aid this process by performing visual search on the specimen collection [28]. The algorithm receives a new leaf image as input and uses statistics based on the lengths of the shortest paths between landmark points on the boundary of the shape silhouette to output a ranked list of closest matches from the specimen collection. The algorithm has been tested on a set of 93 species of plant native to our test location. Although quite accurate, the search results are not perfect, so the botanist must be able to closely inspect individual matches. The EFG must also make possible text search of data associated with the specimen collection.

The EFG must support the workflow of the botanist in the context of the system and tasks. This includes acquisition of the leaf image and clear presentation of the ranked results from the visual search. It also includes supporting the matching task through easy database inspection of stored species and any relevant contextual information about the species and previous samples. Finally, the EFG must provide simple means of recording data, once the match has been verified.

In developing the EFG user interface, we have considered a number of additional requirements. Visual search results must be displayed when the user is mobile and outdoors. The inspection and comparison task requires the ability to explore the visual representation of the leaf in the database at multiple levels of detail and magnification. Additional information about the plant species—including overall plant and root structure, habitat, and historical information—aids in the matching task and should be readily accessible. The information must be situated in the context in which it will be most useful to field botanists and in a manner that supports existing process and skills. To address these issues, we have developed three user interfaces; a fourth has been developed by our collaborators [4].

5 USER INTERFACE

In the context of the model-view-controller paradigm [26], each of our prototypes can be considered a different exploration of a control-view combination using the same model of virtual vouchers. The first prototype uses a see-through, head-worn display to view information in context and provides a tangible user interface for the manipulation of search results. The second prototype provides a hands-free user interface using a see-through, head-worn display that presents search results and enables control of the results through head movement and a single button. A third prototype presents results on a hand-held computer and uses the orientation of the hand-held to navigate through and inspect visual search results.

In considering different user interfaces, we wanted to explore how quickly and easily a user could see results, select a species for inspection, and inspect the species sample or examine necessary contextual information about the species sample. To that end, a key component of our prototype is the ability to explore many levels of detail of the virtual voucher, beyond what one might even expect from the physical leaf. We next provide an overview of related work in zooming and level of detail, followed by prototype descriptions.

5.1 Level of Detail and Semantic Zooming

One of the earliest examples of visualizing a change in levels of detail is Ray and Charles Eames' short film, *Powers of Ten* [13]. As the camera moves away from the Earth, level of detail changes are noted as order of magnitude changes in the width of a square frame. Donelson's Spatial Data Management System [12] made the experience of level of detail changes interactive by giving the user joystick control over the "flight" around a large visual data surface. Zooming into the surface could reveal greater levels of detail or switch to alternative representations. Text could be revealed by zooming into items that represented text. Gurwitz and colleagues' MIDAS supported smooth continuous pan and zoom of an animated microprocessor simulation, changing the level of detail displayed as the user zoomed in and out [18]. Herot and colleagues further developed the spatial data management concept, using a hierarchy of icons and graphical data spaces [19, 20]. Friedell and colleagues extended this work with the View System, which dynamically generated graphics based on database queries and motion through space [15]. Furnas [17] formalized the idea of generalized fisheye views to use focus and degree of interest to change representations. In more recent years, Perlin and Fox introduced the term *semantic zooming* with the Pad system [32] which has been expanded by Bederson and colleagues with Pad++ [6] and PhotoMesa [5].

Our prototypes build on this tradition by considering the set of virtual vouchers as spatial data in which each individual virtual voucher can be explored in both level of detail and semantic zooming.

5.2 Tangible AR User Interface

Motivated by the way botanists in the field manipulate samples of species, we first prototyped a tangible AR user interface [22, 24] that provides a physical representation for inspecting visual search results and individual virtual vouchers. The visual representation of the virtual voucher is displayed in context and changes based on spatial modalities manipulated by moving a tangible handle associated with the virtual voucher.

The use of spatial modalities to change the user interface has been investigated in a variety of ways. Bier and colleagues intro-
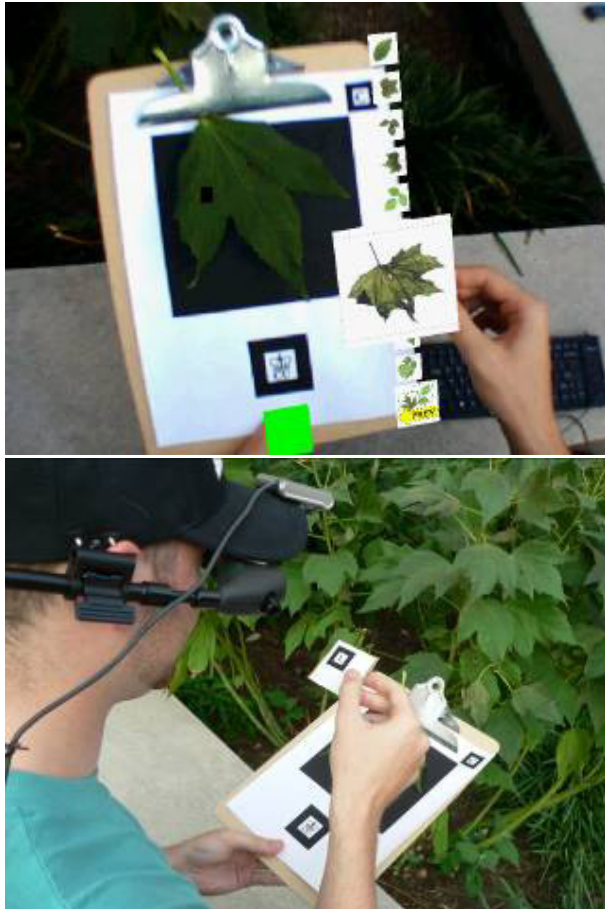
Figure 3: (a) View through a video see-through display of results and a virtual voucher in hand. (b) Third-person view of user and fiducials.

duced Toolglass widgets and Magic Lens filters as a class of see-through tools and as a means of creating spatial modes in user interface systems [8, 9]. Viega extended the Magic Lens concept into 3D by considering a volumetric lens [39] and Rekimoto developed a variant for AR [34]. Looser and colleagues also developed techniques for 3D Magic Lenses in AR and examined fundamental interactions in magnification, object selection and manipulation, and information filtering [29].

We build on this work by inverting the magic lens user interface concept to attempt to take advantage of the existing spatial intuition that botanists use for inspecting physical leaf samples. Instead of changing what the user sees through a lens based on the position of the lens, we change the semantics, modality, or magnification of the object based on the 3D spatial location of the object.

In our prototype, a leaf is placed on a clipboard to provide a consistent background for the computer vision algorithm. A card containing a visual fiducial (not specific to the particular leaf) is then placed below the leaf to trigger image acquisition and initiate the visual search (CP 2). The results of the search are displayed along the side of the clipboard next to the original leaf sample in ranked relevancy order. The card can then be placed in the same location as one of the search results images so that the card morphs into that image and can now be manipulated to inspect the

virtual voucher (Figure 3a–b and CP 3). While this is similar to the picking mechanism described by Kato and colleagues [24], our intention is to provide a conceptual model in which the card transforms into a virtual voucher.

For the inspection task, the user can magnify the leaf and inspect venation or edge details by moving the card towards the user. The leaf image is magnified disproportionately relative to the actual distance travelled, as if the object was growing in size as it moves towards the user (CP 4). Semantic changes are based on distance from the user, spatial zones, or orientation of the card. For instance, if a card is held towards the left, the image of the full plant is shown, and if the card is held towards the right, the image of the sample leaf is shown. An alternative interaction changes the modality by flipping the card.

5.2.1 Spatial Morphing

We have considered a number of different spatial mappings for combining level-of-detail and semantic morphing. One approach uses continuous subspaces of magnification or level-of-detail within contiguous zones of modality, semantics similar to spatial modalities found in the *n*–Vision [14] system. For example, when the virtual voucher is held close to the user, they can magnify and examine a single leaf by pulling it closer or moving it farther away; however, when the virtual voucher is held farthest away, it morphs into the full plant, which also changes in size based on distance. Another example creates zones in the quadrants of space in front of the user. Distance changes level of detail, and the specific zone changes semantics.

In early, informal trials by our botanist colleagues, we found that when the mapping of zones is user-centric, it can be confusing to map too much semantic information along the outward *z*-axis. Changing to another representation as the object moves away from the user does not provide enough information if the size continues to decrease and the representation is too small to see. The projected view of a 3D object that itself remains constant size in three-space does not provide enough of a size change, so we exaggerate the increase in scale as the object gets closer.

Once the botanist has decided on the identification, the botanist places the selected virtual voucher below the actual leaf, triggering a match. The new sample is recorded along with contextual data about the sample.

5.3 Head-Movement Controlled AR User Interface

Our second design explores providing a hands-free user interface to the specimen collection that can be used while the user's hands are otherwise occupied. Instead of tangible objects, we use head movement to control inspection of the virtual vouchers.

There are many ways in which head movement can be mapped to control of a user interface. Chung's comparison of head-tracked steering modes found that "orbital" mode, in which an object rotated in place as the subject's head rotated, produced the highest scores. He attributed this to maintaining the object in the center of vision and muscle memory [11]. Koller and colleagues built on this work to examine additional issues and applications of orbital viewing [25]. Hix and colleagues explored concurrently panning and zooming a flat screen display using head position. Moving the head closer to the screen would zoom, while moving the head from side-to-side would pan [21]. Fuhrmann and colleagues carried this beyond orbital viewing to develop head-directed navigation in which the pitch moves the subject forward or backward in a virtual world while steering is done through head rotation [16]. Schmandt [36] and Brewster and colleagues [10] used head rotation to control channels of audio. We build on existing head-movement control by exploring our own mapping for moving objects and adding the notion of look-and-lock, described later in this section.
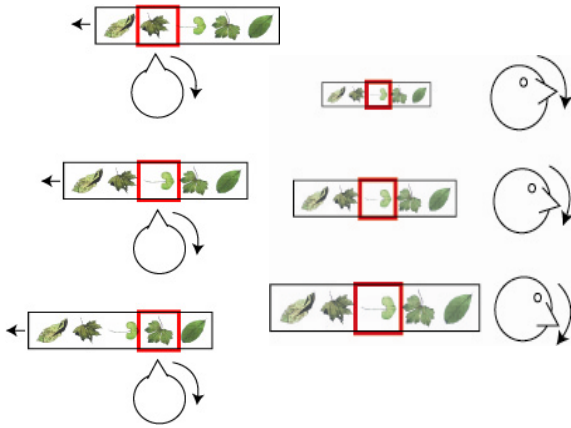
Figure 4: (a) Movement of virtual vouchers as the head rotates right. (b) Scale change as the head angles down.

In this prototype, we developed two variations for head movement control. In the first variation, which we refer to as *free-moving*, the results of a visual search query are displayed in a row perpendicular to the head direction, floating in space and centered in the user's field of view. Rotating the head to the right continuously moves the row of virtual vouchers to the left, while maintaining the row perpendicular to the ray extending from the center of the user's vision, as illustrated in Figure 4 and CP 5. Rotating the head to the left has the opposite effect. In this way, head rotation is used to quickly move between images in the results. The image that is currently in the center of the user's view is highlighted. The second variation, which we call *rotating*, is similar, but instead of sliding left or right, the virtual vouchers rotate position so that the set is always centered, but the voucher in focus changes with rotation. The ordering always stays the same. The effect is similar to Chung's orbital mode [11] and was developed as an alternative to explicitly centering the vouchers.

In both variations, magnification is controlled by head pitch. Looking down increases the object size, while looking up reduces to the overview of all results. This head pitch mechanism is similar to Bell and colleagues' use of head-pitch–controlled scaling for World in Miniature user interfaces [7]. We have experimented informally with multiple modalities, but have found thus far that the best mapping is to magnify the object and increase detail when angled down, and to provide an overview of all results when looking straight ahead or slightly up.

5.3.1 Look-and-Lock

If the user changes head orientation drastically, the results overview can be centered on the current head orientation, and subsequent changes are relative to that orientation. The user can also lock the position of search results at any given time, so that the view does not change based on head movement. The lock is currently implemented by pressing a wireless button worn on the body or hand. For example, a user can select a particular leaf with head rotation, and then magnify the venation by angling the head down slightly. They can then lock the image in place, so they can move their head without changing the magnified view of the leaf. Figure 5 shows a locked voucher viewed through a video see-through display. When the visual results are unlocked, the image is once again controlled by head movement. Pierce and colleagues allowed users to hold objects relative to their head position while using a 3D desktop applications to drop objects in a toolspace [33]. In contrast, we lock the entire view to allow free head movement relative to the world without changing the displayed objects.

We call this clutching mechanism *look-and-lock,* and we have found the technique useful for quickly finding a point of inspection and then comparing that image with the physical species under consideration. At all times, the visual results are displayed in an egocentric manner, such that a mobile user will always have the results ready at hand.

Once the species has been correctly identified, the lock is held down, which triggers a match, and the new sample is recorded, along with contextual data about the sample.

5.4 Hand-Held Computer and Orientation Tracker

We developed a third prototype that uses the software from our head-movement controlled AR and displays the visualization on a hand-held computer. The interface is controlled by tilting the hand-held computer, instead of by head movement. In large part, this user interface was developed to understand the trade-offs between displaying information in context using AR and 2D screen-based presentation of the user interface.

5.5 Hand-Held Computer and Zoomable User Interface

A fourth prototype was developed earlier by our colleagues using the PhotoMesa [5] zoomable user interface on a hand-held computer, and is discussed in more detail by Agarwal [4]. A camera was used to collect leaf images in the field. The images are then downloaded to the computer. The vision algorithm is run on the image and the results are displayed in PhotoMesa as a 2D grid of images. Clicking the left mouse button on an image zooms into that image and clicking the right mouse button zooms out of the image.

6 HARDWARE AND SOFTWARE PLATFORM

Our system runs on a Sony U750 hand-held computer under Windows XP, connected to a Liteye-500 800×600 resolution, color, see-through, head-worn display, mounted on a baseball cap (Figure 6 and CP 1). Our system is intended for use in outdoor environments that require a clear view of the walking path; therefore, we wanted to use an optical see-through display (to transmit the real world at full resolution), with high transparency and brightness, and minimal obstruction of the user's view below the display. Although we would have preferred to use a stereo display, we chose the monocular Liteye-500 because it is significantly brighter and more transparent than the other displays we had available. For comparison, we also used a stereo Sony LDI-D100B 800×600 resolution, color, see-through, head-worn dis-

Figure 5: A locked image seen through an optical see-through display.

Figure 6: Head-worn display, video camera, and orientation tracker mounted on baseball cap.

play. The tangible AR user interface uses a clipboard and small cardboard cards with printed fiducials, which are tracked using a Creative Labs video camera, also mounted on the baseball cap. We use ARToolkit [23] to identify and track the fiducials and OpenGL for image display. The head-movement controlled AR user interface uses an InterSense InertiaCube 2 hybrid inertial orientation tracker to track head orientation, along with a simple wireless two button input device that is worn on the body to initiate lock, center, and select. The third user interface uses the InertiaCube 2 and the Sony U750.

7   OBSERVATIONS AND EVALUATIONS

The prototypes were tried indoors by four of our colleagues in the Department of Botany at the Smithsonian Institution to elicit initial feedback on the conceptual model, process, user interface, and hardware configurations. Although we have tested the functionality of the system outdoors, we wanted a controlled environment that would allow us to focus on the differences between the prototypes. The botanists represented a wide range of expertise with computers and technology. Each person used the three prototype user interfaces: tangible AR, head-movement controlled AR, and hand-held computer with orientation sensor. At the end of the day, they gathered to discuss their experiences and also tried the head-movement control with a full stereo head-worn display. Some had also used the PhotoMesa user interface and could compare our user interfaces to it.

7.1.1   Conceptual Model

The EFG and virtual voucher conceptual models fit well with their existing model and practice. Since they already use a paper field guide as a tool for identification, the concept of an EFG that aids in the same task made sense. The virtual voucher extended their existing concept of a specimen voucher. They were aware that the virtual voucher represented depth of information beyond the image that was on the screen at any given point in time. The model also provided a close association with the purpose of a physical voucher specimen as both a tool for identification and proof that the specimen had been collected. A number of the botanists requested specific information that should be displayed to aid in identification. Collectively, they felt that the characteristics that would be most useful for identification were images (leaves, leaves on branches, full tree or plant, trunk, fruit, and bark), locality, region or distribution maps, genus, species, common name, and text description. Although we had included it in the prototype, the botanists did not see the taxonomic hierarchy as being useful to them in the field.

7.1.2   Process

We also discussed the process of identification and collection of specimens. One of the more interesting observations was that identification is as much a process of elimination as it is one of focus. This came up with regard to the tangible AR user interface when one of our colleagues said he wanted to remove some of the virtual vouchers from the display because he had already eliminated them from the set of possibilities. This decision was based on visual traits, locale, and seasonal traits. For example, a plant that flowers in late August would be excluded when identifying a plant that is flowering in early April. One botanist referred to this as "more a process of elimination than just choosing a winner."

7.1.3   Reactions to Tangible AR

The botanists liked the tangibility of this user interface and the sense that the virtual vouchers were connected to physical objects. The simple action of flipping the card to see a new aspect of the virtual voucher such as the full plant image was also appealing. We observed them making up their own physical language for manipulating the vouchers. One of our colleagues wanted to tap a voucher to make it disappear if he had eliminated it from the possible matches. There were also points when the user interface seemed to confuse them. They all learned to raise the voucher closer to change magnification, but when some of them wanted to change the size of the overview images on the clipboard, none of them realized they could lift the clipboard to bring the set of images closer. After they learned that they could do this, they used that to take a closer look at overview images.

Layout was also brought up during use of the system. One colleague asked that the results be placed in an arc around the leaf rather than just along the side of the leaf to centralize the comparison focus. Another botanist found the ten results we presented overwhelming and remarked that this was a similar problem he had with the large number of results presented in the earlier PhotoMesa prototype they had tried previously. They would prefer seeing only five results at a time. One botanist was concerned that the fiducial cards would get lost because she "loses things in the field all the time." Another botanist asked if the fiducial could be replaced by a pen, so he could use the tip of his pencil or pen and put the leaf on his field notebook. This was brought up again as a concern that the user interface might keep their hands busy when they wanted them available for recording information.

7.1.4   Reactions to Head-Movement Controlled AR

Our colleagues used the two versions of the head-movement controlled AR user interface introduced in Section 5.3: free-moving and rotating. Starting with the free-moving variations, all of the botanists were able to learn the head-movement controls very quickly after a minute or two of experimentation. They all remarked positively on the speed with which they could inspect different plants. The look-and-lock mechanism was seen as particularly useful for comparison of details. However, they wanted the locking mechanism either to be head-controlled or to have the button mounted on the head-worn display. Two of the botanists also wanted to be able to lock into a leaf even as they moved around the leaf, so that they would not accidentally slip into another voucher, like picking a direction or "channel" on which to focus. We were surprised to find that they liked the centering mechanism, not for centering, but because it let them remove the user interface from view and then bring it back to center, no matter where they were looking. They did not like the rotating interface because it changed the spatial arrangement of vouchers. This interfered with their comparison process because vouchers were not always in the same location. The system was also found to be too sensitive to small movements when the vouchers were highly magnified.

7.1.5 Reactions to the Hand-Held with Orientation Control

The hand-held computer was viewed as a familiar tool to use, but the botanists expressed concern over having to carry or hold it. The tilting mechanism was easy for them to learn, and multiple botanists commented that they liked how fast they could move around. Since the user interface was similar to that of the head-movement controlled AR, the comments made about the interaction were similar.

7.1.6 Additional Reactions

The cap-mounted monocular display was seen as uncomfortable and bulky relative to the hand-held computer or the more sturdily mounted binocular display. At times, the botanists were unsure whether to look into the monocular display with their right eye or focus on the world with their left eye. They described the existing headgear that they wore as baseball caps or "Indiana Jones" style hats.

In discussing hands-free interaction, a number of them asked about including audio notes. One wanted audio notes recorded if he could have them automatically transcribed and another was interested in audio notes even without transcription, since he always transcribed notes from his field notebook to a database after every trip.

In comparison to PhotoMesa on the hand-held and the hand-held with orientation control, one botanist commented that he preferred the AR user interfaces because they "make it a part of you."

7.1.7 Lessons Learned

Based on these initial trials, we have learned a number of lessons that will inform our future work. First, we note that each of the different prototypes provided benefits and costs in usage. We want the benefits of physical affordances provided by tangible user interfaces without the extra requirements of a clipboard and fiducial cards. One way to address this will be to use the notebook and pencil that are already carried around by the botanists on field explorations. Similarly, head-movement control provides a useful means of controlling the interface when the hands are busy but is not always necessary and provides less precise control than the tangible interface. We will integrate these modalities together so the botanist has the option of manipulating virtual vouchers with modalities that match different modes of use.

See-through displays have the benefit of representing the information in context. They also provide the experience of directly manipulating virtual vouchers, which was positively viewed in our trials, but they are still bulky to wear. In our trials, the botanists asked about accessing the same functionality from either the head-worn or hand-held system, depending on the local environment. We believe we can address this and carry the same conceptual model across both head-worn and hand-held user interfaces.

For comparison tasks, we observed that drastically changing the spatial layout of results interfered with the comparison task because objects of comparison were not positioned in their expected locations. We found this to be problematic in comparing the rotating versus centered head-movement controlled user interfaces.

While our current application represents visual search results of botanical samples, we believe the techniques for inspection and comparison of visual search results may be able to be generalized to other visual search results in which objects and their holistic characteristics need to be inspected. In particular, we believe the use of physical manipulation coupled with spatial modalities could provide a quick and easy way to browse through a wide range of visual information.

We also believe the look-and-lock mechanism provides a useful means of supporting hands-free manipulation of objects. Based on our observations thus far, the technique is intuitive to learn and supports a combination of head movement and head-movement control.

8 CONCLUSIONS AND FUTURE WORK

We have presented three prototype EFG user interfaces. One provides a tangible AR user interface that uses spatial modalities for semantic zooming and level-of-detail. Another user interface provides a hands-free, head-movement-controlled AR that supports look-and-lock techniques for inspection of details. A third uses orientation of a hand-held computer to control the user interface. We contrasted these with an earlier prototype developed by our colleagues that uses a zoomable interface on a hand-held computer to view and inspect results. Our user interfaces explore new techniques for displaying visual search results and inspecting virtual vouchers in an EFG.

Based on feedback from our initial trials, there are many directions that we would like to explore in the future. In our next iteration of the user interface, we intend to integrate head-movement control with tangible AR and the field notebook, so that the same user interface can be used with or without tangible components. We also want to experiment with completely hands-free interactions by substituting head gestures, such as nods, for the buttons we use currently. Finally, we want to experiment more with continuous and quantized movement to create "channels," as described in Section 7.1.4. In terms of visual search results, we want to explore ways of displaying large search results and techniques for handling search when the initial results do not contain a correct match. While this is not necessary for the current project, we believe it will apply to more generalized use.

As part of our joint research, we plan to expand the EFG to eventually support the entire process of "voucher" collection across multiple trips. It will include:

- integration of GPS and dead reckoning mechanisms (deducing location based on speed and direction of movement) to provide geotemporal coding that enables contextual information;
- geotemporally-situated visualization of sites where a species has been found, as well as situated visualization of prior paths followed in the environment; by situated *visualization*, we mean a visualization that is located and overlaid in a place where it is meaningful and where the place adds to the visualization; and
- and tools for recording notes (most likely audio).

As the EFG and its user interface mature, our botanist colleagues will field test it in their work in Panama or Costa Rica. Finally, we would like to see this system be used on a daily basis by amateur botanists who simply want to identify a plant they encounter on a morning walk.

10 VIDEO

Digital video of this work can be viewed at http://www.cs.columbia.edu/graphics/projects/efg/

REFERENCES

[1] CyberTracker. http://www.cybertracker.co.za/

[2] The Green River Paleobotany Project (GRPP). http://greenriver.dmns.org/

[3] Gregory D. Abowd, Christopher G. Atkeson, Hong Hong, Sue Long, Rob Kooper, and Mike Pinkerton, "Cyberguide: a mobile context-aware tour guide," *Wireless Networks*, vol. 3, pp. 421-433, 1997.

[4] Gaurav Agarwal, "*Presenting Visual Information to the User : Combining Computer Vision and Interface Design.*" Master's thesis, University of Maryland, College Park, 2005.

[5] Benjamin B. Bederson, "PhotoMesa: a zoomable image browser using quantum treemaps and bubblemaps," UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology, Orlando, Florida, 2001.

[6] Benjamin B. Bederson and James D. Hollan, "Pad++: a zooming graphical interface for exploring alternate interface physics," Proceedings of the 7th annual ACM symposium on User interface software and technology, Marina del Rey, California, United States, 1994.

[7] Blaine Bell, Tobias Hollerer, and Steven Feiner, "An annotated situation-awareness aid for augmented reality," UIST '02: Proceedings of the 15th annual ACM symposium on User interface software and technology, 2002.

[8] Eric A. Bier, Maureen C. Stone, Ken Fishkin, William Buxton, and Thomas Baudel, "A taxonomy of see-through tools," CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems, Boston, Massachusetts, 1994.

[9] Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony DeRose, "Toolglass and magic lenses: The see-through interface," *Proc. SIGGRAPH '93*, vol. 27, pp. 73–80, 1993.

[10] Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker, "Multimodal 'eyes-free' interaction techniques for wearable devices," CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems, Ft. Lauderdale, Florida, 2003.

[11] James C. Chung, "A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams," Proceedings of the 1992 symposium on Interactive 3D graphics, Cambridge, Massachusetts, United States 1992.

[12] William C. Donelson, "Spatial management of information," Proceedings of the 5th annual conference on Computer graphics and interactive techniques, 1978.

[13] Charles Eames and Ray Eames, "*Powers of Ten*," Eames Office, 1968, pp. 9 min.

[14] Steven Feiner and Clifford Beshers, "Worlds within worlds: metaphors for exploring n-dimensional virtual worlds," UIST '90: Proceedings of the 3rd annual ACM SIGGRAPH symposium on User interface software and technology, Snowbird, Utah, 1990.

[15] Mark Friedell, Jane Barnett, and David Kramlich, "Context-sensitive, graphic presentation of information " *SIGGRAPH Comput. Graph.* , vol. 16, pp. 181-188, 1982.

[16] Anton Fuhrmann, Dieter Schmalstieg, and Michael Gervautz, "Strolling Through CyberspaceWith Your Hands In Your Pockets:Head Directed Navigation In Virtual Environments," Proceedings of the 4th EUROGRAPHICS Workshop on Virtual Environments, 1998.

[17] G. W. Furnas, "Generalized fisheye views," CHI '86: Proceedings of the SIGCHI conference on Human factors in computing systems, 1986.

[18] R F Gurwitz, R T Fleming, and A Van Dam, "MIDAS: A Microprocessor Display and Animation System," *IEEE Transactions on Education*, vol. E-24, pp. 126-133, 1981.

[19] Christopher F. Herot, "Spatial management of data," *ACM Trans. Database Syst.* , vol. 5, pp. 493-513, 1980.

[20] Christopher F. Herot, Richard Carling, Mark Friedell, and David Kramlich, "A prototype Spatial Data Management System," Proceedings of the 7th annual conference on Computer graphics and interactive techniques, Seattle, Washington, United States, 1980.

[21] Deborah Hix, James N. Templeman, and Robert J.K. Jacob, "Pre-Screen Projection: From Concept to Testing of a New Interaction Technique," Proceedings of ACM CHI'95 Conference, Denver, 1995.

[22] Hiroshi Ishii and Brygg Ullmer, "Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms," Proceedings of the SIGCHI conference on Human factors in computing systems, Atlanta, GA, 1997.

[23] H. Kato and M. Billinghurst, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," 2nd IEEE and ACM International Workshop on Augmented Reality 1999.

[24] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana, "Virtual object manipulation on a table-top AR environment," Proc. Int'l Symp. Augmented Reality 2000, Los Alamitos, Calif., 2000.

[25] David R. Koller, Mark R. Mine, and Scott E. Hudson, "Head-tracked orbital viewing: an interaction technique for immersive virtual environments " Proceedings of the 9th annual ACM symposium on User interface software and technology, Seattle, Washington, United States 1996.

[26] Glenn E. Krasner and Stephen T. Pope, "A Description of the Model-View-Controller User Interface Paradigm in the Smalltalk-80 System," *J. Object-Oriented Programming*, vol. 1, pp. 26-49, 1988.

[27] W. John Kress, "Paper floras: how long will they last? A review of Flowering Plants of the Neotropics," *Am. J. Bot.*, vol. 91, pp. 2124-2127, 2004.

[28] Haibin Ling and D. W. Jacobs, "Using the Inner-Distance for Classification of Articulated Shapes," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

[29] Julian Looser, Mark Billinghurst, and Andy Cockburn, "Through the looking glass: the use of lenses as an interface tool for Augmented Reality interfaces," GRAPHITE '04: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia, Singapore, 2004.

[30] Ian Oliver, 1997. "A Virtual Catalogue of Invertebrate Life on Earth," accessed from http://www.iupac.org/symposia/proceedings/phuket97/oliver.html

[31] Jason Pascoe, Nick Ryan, and David Morse, "Using While Moving: HCI Issues in Fieldwork Environments," *ACM Transcations on Computer Human Interaction*, vol. 7, pp. 417-437, 2000.

[32] Ken Perlin and David Fox, "Pad: an alternative approach to the computer interface," Proceedings of the 20th annual conference on Computer graphics and interactive techniques, Marina del Rey, California, 1993.

[33] Jeffrey S. Pierce, Matthew Conway, Maarten van Dantzich, and George Robertson, "Toolspaces and glances: storing, accessing, and retrieving objects in 3D desktop applications," SI3D '99: Proceedings of the 1999 symposium on Interactive 3D graphics, 1999.

[34] Jun Rekimoto, "The magnifying glass approach to augmented reality systems," International Conference on Artificial Reality and Tele-Existence '95 / Conference on Virtual Reality Software and Technology '95 (ICAT/VRST'95), 1995.

[35] Nick Ryan and Jason Pascoe, "FieldNote: a Handheld information System for the Field," Workshop on TeloGeoProcessing, Lyon, 1999.

[36] Chris Schmandt, "Audio hallway: a virtual acoustic environment for browsing," Proceedings of the 11th annual ACM symposium on User interface software and technology, San Francisco, CA, 1998.

[37] R. D. Stevenson, William A. Haber, and Robert A. Morris, 2003. "Electronic Field Guides and the User Communities in the Eco-informatics Revolution," accessed from http://www.consecol.org/vol7/iss1/art3

[38] Barbara Tinker, Carolyn Staudt, and Dick Walton, 2002. "The Handheld Computer as Field Guide," accessed from http://www.concord.org/publications/newsletter/2002winter/monday _lesson.html

[39] J. Viega, M. J. Conway, G. Williams, and & R. Pausch, "3D Magic Lenses," UIST '96, Seattle WA, 1996.

[40] Ron B. Yeh and Scott R. Klemmer, "ButterflyNet: Mobile Capture and Access for Biologists," to be published in Extended Abstracts of UIST 2005, Seattle, WA, 2005.
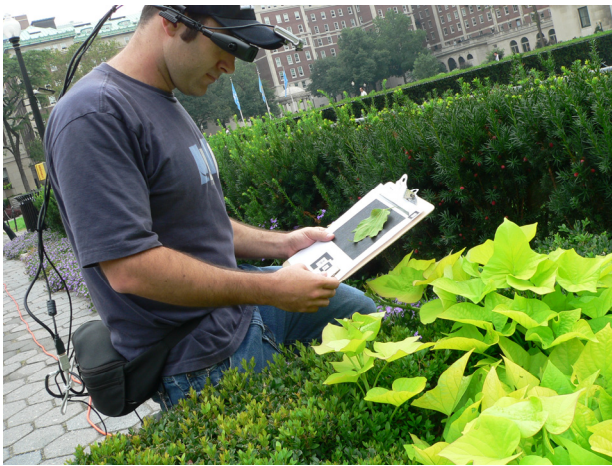
Color Plate 1: Mobile AR prototype being used outside (the orange power cable is for our video system, not the prototype).
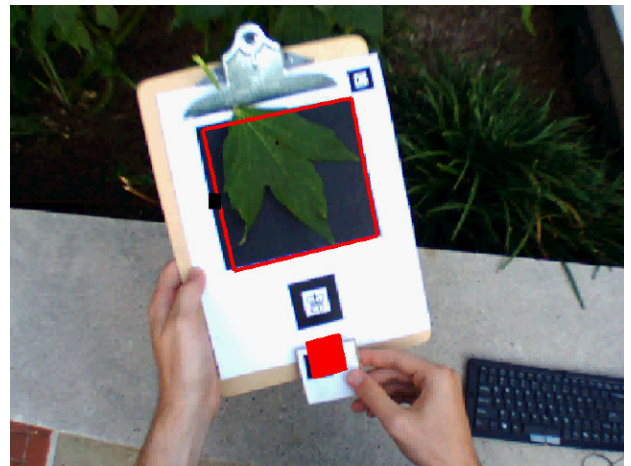
Color Plate 2: View through a video see-through display outdoors as a leaf image is acquired for visual search.
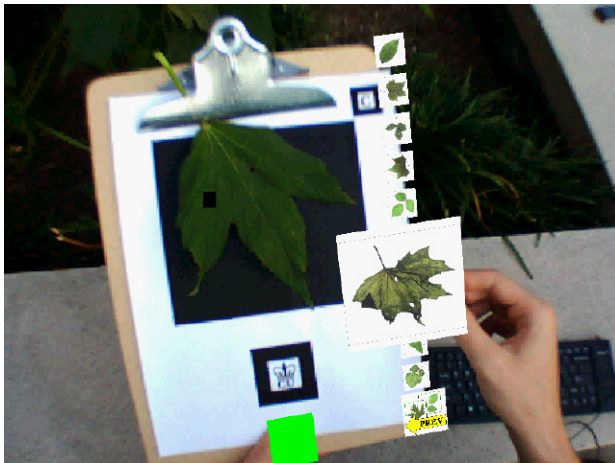
Color Plate 3: View through a video see-through display outdoors showing browsing results of a search query.

Color Plate 4: View through a video see-through display of inspection of a tree view of a virtual voucher.

Color Plate 5: A locked image as viewed through an optical see-through display outdoors.