

LooknLearn: An AR System of Linked Video

Sujeetkumar Umbarkar (Corresponding author)

Department of Computer Engineering,
AISSMS College of Engineering, Pune -411001
University of Pune, India

Tel: +91-8149597979 E-mail: sujeet.umbarkar@gmail.com

Rahul Patil

Department of Computer Engineering,
AISSMS College of Engineering, Pune -411001
University of Pune, India

Tel: +91-9579578784 E-mail: patilrahul.sangli@gmail.com

Deepak Yadav

Department of Computer Engineering,
AISSMS College of Engineering, Pune -411001
University of Pune, India

Tel: +91-8087602473 E-mail: dpk.yadav@yahoo.com

Abstract

This paper presents an augmented reality system using video as the primary interface. Using the authoring component of this system, individuals can leave "video streams," consisting of several layers of film, movies, moving pictures, animation, video clips at a location outdoors. Using the navigation component, individuals can watch streams by walking into the area that the stream occupies. Furthermore, stream can be linked together, whereby an individual is directed from one stream to related streams in the area.

Keywords: Augmented reality, video stream, navigation, authoring environment, video links

1. Introduction

Augmented reality (AR) systems are designed to enhance the known physical environment. Typically, this is done by get an edge of a virtual world onto the physical world. Most of such edges used in augmented reality system may consist of video, video, or a combination of both the mediums. AR differs from the completely immersive world of virtual reality (VR) in that users are able to maintain the context of their surrounding environment, while still obtaining the benefit of additional sensory input and information.

LooknLearn uses virtual overlays to allow users to leave personal streams that are associated with a particular location in physical space. An author creates a sound and embeds it at a location using an authoring toolkit. An individual can then look that video while he or she is traversing the designated location. *LooknLearn* approaches augmenting the space with an emphasis on social interactivity. Unlike other systems, the users of the system create large content of the space. A simple, yet flexible authoring system is thus an important part of *LooknLearn*. Furthermore, *LooknLearn* does not require adding any physical objects (such as tags) to the environment. This makes it particularly useful in the augmentation of public spaces.

There are various basic issues we are addressing with *LooknLearn*. How do we define a video stream? How do we embed it in a space? How are we made aware that it exists? How can we link videos to compose a clear ways for users to navigate freely?

2. Previous Work

Enhancing a space with video is not new. Many organizations and research centers today offer video tapes that direct a user through an exhibit while providing additional information about the various pieces of scientific discovery, may it be visual topographic maps plotted on Encyclopedia. This approach is linear and somewhat inflexible: the user follows the guide on the tape and cannot stray to a different path. James. D. Hollan built a prototype augmented reality system for museums that went beyond the traditional video-based systems by allowing for non-linear browsing. Using James's system, the curators of a museum could place electronic transmitters above pieces of art, and then store video information about each piece of art on a minidisk. As patrons using the system walked near one of the tagged pieces of art, the appropriate video was recalled from the minidisk and played back through the patron's device.

Both of these approaches have some advantages. With the cassette tape, presumably a curator has taken the time to arrange a tour in a meaningful manner, making the order of the exhibit, as well as the content, an integral part of the presentation. With Bederson's system, the material is also created beforehand. However, the user can move back and forth, linger, and create his/her respective tour. The visitor to a museum can now make the museum visit a very personal experience. Another AR system that influenced *LooknLearn* is the Touring Machine. This is a visual AR system designed to explore an urban environment, specifically, the Pune campus. Using this system, individuals are provided with information about the buildings he or she encounters on a walk through Pune. Context sensitive web pages are accessible via a hand-held display. An important contribution of The Touring Machine is its reliance on the global positioning system (GPS), which removes the need for placing physical tags in an area. GPS therefore allows greater flexibility in placement of augmented information.

One issue with the Touring Machine that we investigated in *LooknLearn* is how much attention an augmented reality system should grab from the user. The goal of AR is to provide a layer of information over the real world. Determining how much this layer should distract from the real world is an important question in any AR project. In The Touring Machine, the textual information overlaying buildings and the

hand-held device may become overly distracting. Because The Touring Machine was created as a way to provide information to newcomers, this approach may be appropriate. However, in *LooknLearn* we attempted to make the AR layer a more subtle and integral part of the environment.

Most previous work in augmented reality has focused on the experience of the end-user. Content creation has been left to the designers of the system (as in The Touring Machine) or to professionals (as in Bederson's system). In *LooknLearn*, allowing the users of the system to create content for the system was an important goal from the beginning. As a result, we have viewed the system as a tool that individuals can use to enhance an environment, rather than as a tool to provide particular information to end-users.

3. LooknLearn

LooknLearn is being developed within a larger exploration of video augmented reality. Many of the technical and design challenges, such as creating user-location triggered video or evaluating the importance of directional video, are common to all of our applications. *LooknLearn* focuses on the in situ placement of oral histories and reflections; this scenario emphasizes the development of a navigational system (for exploring a series of related stories) and an easy to use authoring environment.

LooknLearn is a location or area that is frequently used by members of the community and it allows for casual encounters for it is also a throughway of the campus near about.

Finally, it is an area that is constantly evolving: seasonally and architecturally. It is ripe for historical perspectives as well as reflection to those who visit it frequently.

4. The Video Stream

The basic unit in *LooknLearn* is the video stream. The video stream is a multi-layered, customizable collection of videos that can be placed in a space. An important design consideration for the video stream was utilizing the accuracy afforded us by technology that can provide the location of the user within tens of centimeters of his or her actual location. Video AR systems such as [1, 3] could determine if the user was directly in front of an object or in a given room. The higher resolution available in *LooknLearn* allows one to experiment with degrees of videos that were not previously possible.



Figure 1 The map and the historic building at the specific location

On the left is a GUI representation of the augmented environment. (Department of Chemistry) The semi-transparent circles represent video stream. On the right is a photograph of the area we are augmenting, from the perspective Department of Chemistry. The stream consists of a single primary video, with other video plaited into the border. These plaits of video overlap the stream, with each plaits of video shifting into and out of recognition. The use of plaited video in addition to the primary video allows for the creation of highly individualized streams. For example, the primary video may be the reflections of an individual on the nearby architecture. The author's favorite video clip or animation could then be plaited around the edge of the reflections.

The above figure depicts the graphical user interface for the authoring component. It consists of a map of the Media Lab courtyard with superimposed sound streams. This tool has two modes: a draft mode and a demo mode. The draft mode allows the author to create a new sound imprint, place it on the map, adjust the various video properties, and record video. This is done using a drag and drop interface written in Java.

The demo mode of the authoring interface allows the author to simulate the experience as if they were outdoors in the courtyard. By using the mouse to designate the navigator's position, the author can look what the navigator would look if he or she were following that path. In this manner, the author can experiment with different scenarios, video braiding, and video clips with the aid of visuals.

Allowing a person to record an imprint in the field is an important goal of this project. Because the video streams are contextualized to the space, it is reasonable for a user to create an imprint while standing or sitting in that space. To make adding streams to a space as easy as possible, the interface for recording a sound in the field will be very simple. The user will hold a small microphone, and a very simple interface will be provided on the PalmPilot (with "record," "stop" and "cancel" buttons). The imprint will have

default characteristics that can be adjusted later using the more complex GUI environment.

It is important to note that most properties of these streams are adjustable. The amount of braiding that occurs, whether the video loops or plays once, and even whether a plait exists or not is left up to the author. The intention of our project is not to determine what is or is not an ideal stream, but rather to provide the underlying structure and leave the form up to the person creating the streams. This is similar to HTML (Hyper-Text Markup Language), in which the tools are provided for individuals to make a web page, but it is up to the author to make the web page interesting.

5. Navigating

To browse the video streams, the user carries the navigation toolkit as he or she travels through the augmented space. The toolkit, shown in Figure 2, consists of a pair of headphones with a digital compass attached, a laptop, a PalmPilot, a GPS receiver, a battery, and a microphone. Using the information from the GPS and the digital compass, the *LooknLearn* system is able to determine the location and head position of the user with a high degree of accuracy. As the navigator encounters a video streams, he or she looks the stream's video. Because the head position is known, the video can also be manipulated, which makes the video "appear" to be coming from a particular location in space.

To aid the user in movement through a space, a PalmPilot is provided. The PalmPilot can show the distance and direction to streams in an area. It is important to note that the information supplied on the PalmPilot is not necessary to use the system. Rather, it provides information that can aid the traveler through the augmented space. In future implementations, we will provide other, video-only signals to the user about the presence of streams in an area (such as "video samples" emanating from the center of the streams). Both of these systems may be unnecessary for the bulk of users, as many people may wish to simply wander around an area and seek out the stream without any help from the system.

Perhaps the biggest technical challenge in the navigation system we use is accurately sensing the position of an individual. We are currently using a GPS receiver to determine positioning. Using a system such as GPS is convenient for it allows for arbitrary positioning of videos.

The space does not have to be specially prepared with IR tags or other samples. A video can be placed on a busy street, by a statue, or at an arbitrary location in a field. Some of the drawbacks to GPS are that it can only receive signals outdoors and that it does not provide a consistent unique position. Upgrading to a real-time kinematic (RTK) GPS receiver dramatically increases the precision from stand-alone GPS. Further inconsistencies in position are accounted for in software.

The emphasis of navigating with *LooknLearn* is on the perceptual interface and not the computing and hardware behind it.

6. Authoring Environment

Many AR systems have been developed for navigating spaces. It is rare to find one that also allows the user to easily create or modify a space. With most systems, one must view or listen to material created by others. It would require programming experience and extensive knowledge of the internal structure of the AR system to create a new environment.



Figure 2 The wearable devices

At CeBit a group of German companies including Interactive got together and created the next generation of wearable gadgets starting with a "multimedia backpack with remote control functions for a video camera", then a "jacket with integral GPS, GSM, Bluetooth and mp3 capabilities" and finally an "Augmented Reality Vest which is used by the automotive industry to link the virtual world with the real world."

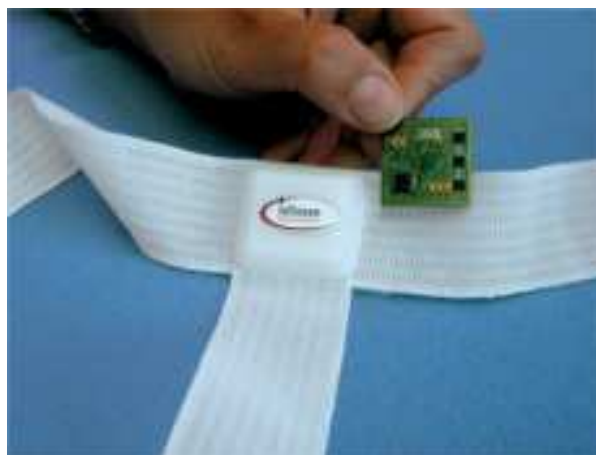


Figure 3 Electronics into clothing - Infineon has researched and developed since 2000.

Electronics products to penetrate more and more aspects of daily life such as Microprocessor-based systems are increasingly networked, wireless communication is made at different levels by more and more - GSM, GPRS, UMTS on the one hand, Bluetooth, RFID and wireless networks on the other side. Due to the high pressure to innovate, the electronics systems penetrate into areas that had previously nothing to do with HiTech. The integration of electronics can open up new market opportunities for the textile industry in the areas of sports, medicine and health, and safety in the area.

7. Video Linking

In order to enhance the sociable and informational aspects of the augmented space, the *LooknLearn* system allows authors to link streams together. On the World Wide Web (WWW), a user can create a page containing links to other sites he or she finds interesting or useful. Using this as inspiration, we added the ability for an author to link his or her streams to any other streams in the system.

As an example of how links could be useful in this environment, consider a hypothetical user from Brazil. This user decides to create an imprint near some trees in the Media Lab courtyard reflecting on the colorful New England autumn, and comparing it to the lack of season changes in her native country. She can then link this imprint to an imprint she created about the cold New England winter, and an imprint one of her Brazilian friends made.

Although the WWW was an inspiration for this project, the notion of linking in an augmented space poses a challenge. This is the fact that an augmented space is a true three-dimensional world, whereas the WWW is, for all intents and purposes, represented by a two-dimensional monitor. That is, when an individual viewing a web page follows a link, the linked page is loaded onto the same screen the individual was already using. In a system such as *LooknLearn* where the streams are contextualized to a particular area, the user must move to a new location in space to follow a link.

This poses a challenge in the design of the system, as mechanisms must be put in place to direct the user to the location of the linked streams. Because we are exploring the use of video only interfaces as well, the video beacon idea described above would also work in this situation.

8. Future work

One of the most important questions we are attempting to address with this system is the feasibility of navigating an augmented space using sound exclusively. Current AR technology could be used to provide a visual indication of imprint location. However, because the focus of *LooknLearn* is on creating video streams in a space, providing such rich visual indicators would distract from the content. As the project develops, we will explore different ways of orienting a user in a space with video. As a result, the use of this system by the visually impaired will become an interesting area to explore.

We also intend to more fully explore various notions of video augmentation. In the current avatar, streams

exist only as spots of video clips. It will be interesting to explore other ideas of augmentation. One example we have considered is the concept of a “video stream” in which a particular location consists of ever-moving snippets of sound. While standing in this stream, the user would look videos rushing past (using the spatialized video system described above).

One of the limitations we have encountered in this project is the bulk of the system. To obtain the accuracy that we need, the GPS requires two large antennas and a bulky receiving unit. As a result, most of the system is contained on a luggage cart that the user can move from spot to spot. (We chose not to use a backpack to avoid the issue of user fatigue from the weight.)

In the future, we hope that more accurate GPS systems will approach the size of current off-the-shelf, but less accurate, implementations. Once this goal is realized, the entire *LooknLearn* system could conceivably be enclosed in a Walkman sized device, allowing more natural use of the system.

9. References

Bederson, B. Video Augmented Reality: a prototype automated tour guide. Conference companion on Human factors in computing systems, 1995, Page 210.

Feiner, S., MacIntyre, B., Höllerer, T., Webster, A. A Touring Machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. Proceedings of the First Annual International Symposium on Wearable Computing, Cambridge, MA, October 13-14, 1997. Page 74-81.

Mynatt, E., Back, M., Want, R., Baer, M., and Ellis, J. Designing Video Aura. Conference proceedings on Human factors in computing systems, 1998, Page 566 - 573.

Rekimoto, J., Ayatsuka, Y., and Hayashi, K. Augment-able Reality: situated communication through physical and digital spaces. ISWC'98, 1998.

This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE's homepage:

<http://www.iiste.org>

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. **Prospective authors of IISTE journals can find the submission instruction on the following page:**

<http://www.iiste.org/Journals/>

The IISTE editorial team promises to review and publish all the qualified submissions in a fast manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar

