Augmented Reality Application

For the Educational System

Duagi T. Bulent

A Report Paper submitted in partial fulfillment of the requirements for the degree of

Masters of Advanced Computer Graphics, Multimedia and Virtual Reality

University "Politehnica" Bucharest
2010

Report Paper Structure:

1. Abstract

2. Keywords

3. Theoretical background

    3.1 Augmented Reality

        3. 1.1. Domains for Augmented Reality

        3.1.2. Augmented Reality and Learning

    3.2. Learning Styles

        3.2.1. Visual Learning Style

        3.2.2. Auditory Learning Style

        3.2.3. Kinesthetic Learning Style

4. Similar Work

5. Implementation

    5.1. Technologies

        5.1.1. ARToolkit and FLARToolkit

        5.1.2. FLARManager

        5.1.3. Papervision3D

        5.1.4. Adobe Flex

        5.1.5. Adobe Flash

        5.1.6. Actionscript

        5.1.7. Adobe Flash Builder

    5.2. Toolkit Processes

        5.2.1. The FLARToolkit process

        5.2.2. The FLARManager toolkit

    5.3. Interaction proposal

6. Further development

7. References

1. Abstract:

Augmented Reality (AR), which overlays virtual objects onto real scenes, has large potential to provide learners with a new type of learning material. There are at least three different learning styles, each style can be enhanced using AR, especially the kinesthetic learning style, which is the one most ignored in the actual educational system. Although many AR systems have been developed for demonstration, there is a gap between their ideal and practical use. This paper presents an application that combines augmented reality and traditional printed materials. Improvement of human-computer interface is considered to serve as a bridge for the gap. Open-source AR technology facilitates the development of educational AR applications.

2. Keywords:

Augmented Reality, learning, human-computer interaction, educational

3. Theoretical background:

3.1. Augmented Reality

Augmented Reality (AR) is a growing area in Mixed Reality research. Mixed Reality combines the content from the real world with virtual imaginary. Augmented Reality is a subset of this where virtual content is overlaid into real objects of the world. Extending the concept of AR, it includes virtual graphics and audio.

An Augmented Reality system supplements the real world with virtual objects. It means that virtual (computer-generated) content is added to the real world. An AR system has the following three main characteristics:

- Combines real and virtual objects in a real environment
- Runs interactively, and in real time
- Registers virtual objects onto the real world.

Milgram [1][2] describes a taxonomy that identifies how augmented reality and virtual reality are related. He defines the Reality-Virtuality continuum shown as Figure 1. The real world and a totally virtual environment are at the two ends of this continuum with the middle region called Mixed Reality. Augmented reality lies near the real-world end of the spectrum with the predominate perception being the real-world augmented by computer generated data. Augmented virtuality is a term created by Milgram to identify systems that are mostly synthetic with some real world imagery added—such as texture mapping video onto virtual objects. An example of this technology is live video texture-mapped onto the graphics image of an avatar—a computer-generated virtual character that is a stand-in for the user's presence within a virtual environment [3].
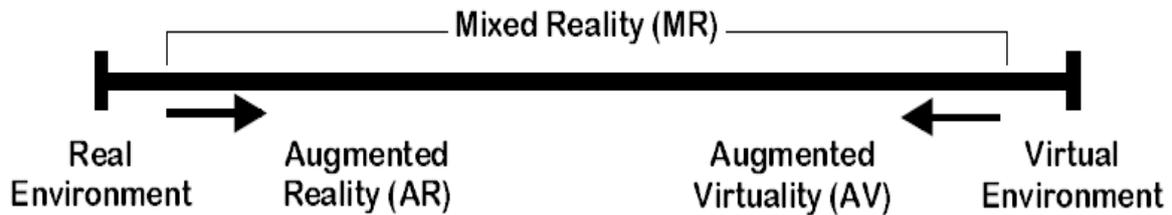
Figure 1. Milgram's Reality-Virtuality Continuum

Milgram further defines a taxonomy for the Mixed Reality displays. The three axes he suggests for categorizing these systems are: Reproduction Fidelity, Extent of Presence Metaphor and Extent of World Knowledge. Reproduction Fidelity relates to the quality of the computer generated imagery ranging from simple wireframe approximations to complete photorealistic renderings. The real-time constraint on augmented reality systems forces them to be toward the low end on the Reproduction Fidelity spectrum. The current graphics hardware capabilities cannot produce realtime photorealistic renderings of the virtual scene. Milgram also places augmented reality systems on the low end of the Extent of Presence Metaphor. This axis measures the level of immersion of the user within the displayed scene. This categorization is closely related to the display technology used by the system. There are several classes of displays used in augmented reality systems an each gives a different sense of immersion in the virtual environment presented to the user. In an augmented reality system, some display technologies utilize the user's direct view of the real world. Immersion in that environment comes from the user simply having his eyes open. It is contrasted to systems where the merged view is presented to the user on a separate monitor for what is sometimes called a "Window on the World" [4] view.

The third, and final, dimension that Milgram uses to categorize Mixed Reality displays is Extent of World Knowledge. Augmented reality does not simply mean the superimposition of a graphic object over a real world scene. This is technically an easy task. To do a realistic merging of virtual objects into a real scene, knowledge about the world is needed. One difficulty in augmenting reality, as defined here, is the need to maintain accurate registration of the virtual objects with the real world. This requires detailed knowledge of the relationship between the frames of reference for the real world, the camera viewing it and the user . To properly compute the visual interactions between real and virtual objects, data about their locations and orientations in three-dimensional space are needed. Correct dynamic interactions with virtual objects require knowledge of the dynamic characteristics of both the real and virtual objects in the augmented environment. In some domains models for the real and virtual worlds are well known, which makes the task of augmenting reality easier or might lead the system designer to use a completely virtual environment.

3.1.1. Domains for Augmented Reality

3.1.1.1. Medical

Because imaging technology is so pervasive throughout the medical field, it is not surprising that this domain is viewed as one of the more important for augmented reality systems. Most of the medical applications deal with image-guided surgery. Pre-operative imaging studies, such as CT or MRI scans, of the patient provide the surgeon with the necessary view of the internal anatomy. From these images the surgery is planned. The surgeon visualizes the path through the anatomy to the affected area where, for example, a tumor must be removed, by first creating a 3D model from the multiple views and slices in the preoperative study. Through their extensive training, surgeons become very adept at mentally creating the three-dimensional visualization that is needed to render a diagnosis. Some newer systems do have the ability to create 3D volume visualizations from the imaging study. Figure 2 shows how augmented reality can be applied so that the surgical team sees the CT or MRI data correctly registered on the patient in the operating theater while the procedure is progressing. Being able to accurately register the images at this point will enhance the performance of the surgical team and may eliminate the need for the painful and cumbersome stereotactic frames [5] currently used for registration. Descriptions of other work in the area of image-guided surgery using augmented reality can be found in [6][7][8].

Figure 2. Image guided surgical procedure

Another application for augmented reality in the medical domain is in ultrasound imaging [9]. Using an optical see-through display the physician can view a volumetric rendered image of the fetus overlaid on the abdomen of the pregnant woman. The image appears as if it were inside the abdomen and is correctly rendered as the user moves. Figure 3a shows an image from the system along with a second application in ultrasound imaging seen in Figure 3b. In this second application, the augmented reality image helps the surgeon guide a biopsy needle to the site of a suspected tumor during a mock breast biopsy procedure. The V-shaped object in the left part of the image is used for registering the ultrasound image with the view of the real scene that the surgeon is seeing.
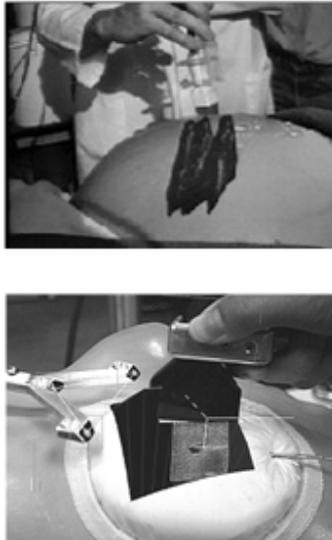
Figure 3. Ultrasound imaging using augmented reality displays

Finally, a future application is the extension of the virtual reality based craniofacial surgery visualization and simulation systems [10] by adding an augmented reality display. These systems currently allow the surgeon to experiment, within the safety of a virtual environment, with different treatment approaches for the reconstructive work to be done. The model that the surgeon works on is a 3D volume visualization derived from a pre-operative CT or MR study. Augmented reality would allow the surgeon to see the final results directly on the patient rather than only with the volume visualization.

3.1.1.2. Entertainment

A simple form of augmented reality has been in use in the entertainment and news business for quite some time. When you watch an evening weather report the weather reporter often stands in front of changing weather maps. In the studio the reporter is actually standing in front of a blue or green screen. This real image is augmented with computer generated maps using a technique called chroma-keying. It is also possible to create a virtual studio environment so that the actors appear to be positioned in a studio with computer generated decorating [11] an example of which is shown in Figure 9. The work with virtual studios goes beyond the simple chroma-keying of an actor over a flat two-dimensional background image. It extends the traditional keying to define the background in three-dimensions and key the live action as a three-dimensional image also. Balcisoy and Thalmann [12] present a virtual studio populated by virtual humans that interact with human actors. The interactions are very carefully scripted and the human actor does not have the ability to feel a virtual actor. With tracking of the human, appropriate visual interactions can be accomplished. This system uses calibrated cameras and careful measurements of the location of objects in the real studio.

Movie special effects make use of digital compositing to create illusions [13]. Strictly speaking, with current technology, this is not considered augmented reality because it is not generated in real-time. Most special effects are created off-line, frame by frame with a substantial amount of user interaction and computer graphics system rendering. But some work is progressing in computer analysis of the live action images to determine the camera parameters and use this to drive the generation of the virtual graphics objects to be merged [14]. Princeton Electronic Billboard has developed an augmented reality system that allows broadcasters to insert advertisements into specific areas of the broadcast (National Association of Broadcasters 1994). For example, while broadcasting a baseball game this system places an advertisement in the image so that it appears on the outfield wall of the stadium. The electronic billboard requires calibration to the stadium by taking images from typical camera angles and zoom settings in order to build a map of the stadium including the locations in the images where advertisements will be inserted. By using pre-specified reference points in the stadium, the system automatically determines the camera angle being used and referring to the pre-defined stadium map inserts the advertisement into the correct place. The approach used for mapping these planar surfaces is similar to that which is presented in Chapter 4 of this thesis. A French company, Symah Vision, has also developed a similar application. Another application in sports broadcasting is Fox network's FoxTrax system [15] for tracking the path of a hockey puck during a game. The path of the puck is overlaid on the image of the hockey rink as a blue streak. The streak changed color based on the speed of the puck. This system requires a detailed calibration process for each television camera and the ice ring itself. Augmented reality can be applied to enhance games that people play. A system [16] developed for pocket billiards players uses a head-mounted display and wearable computer to analyze the layout of the table and suggest possible shots for the player to take. The trajectory of the shot is displayed as augmented graphics over the image of the pool table. Or consider a futuristic game of paintball where players wear augmented reality headsets. The image that the players see is not only of the game area and their real opponents but virtual players are also playing along with them.
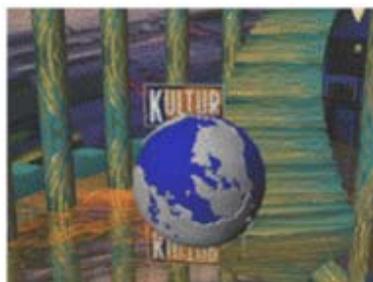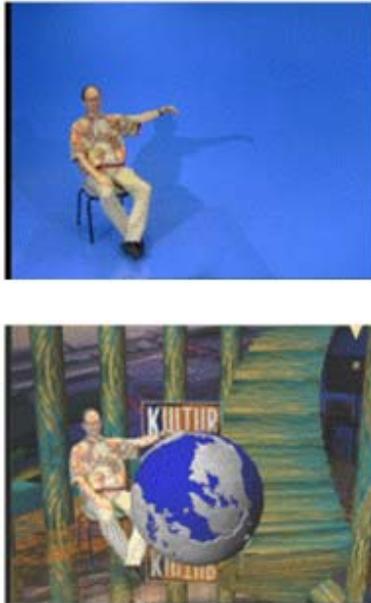
Figure 4 - Virtual set technology. (a) virtual set background, (b) live action, (c) combined video [11].

3.1.1.3. Military

The military has been using displays in cockpits that present information to the pilot on the windshield of the cockpit or the visor of their flight helmet. This is a form of augmented reality display. SIMNET, a distributed war games simulation system, is also embracing augmented reality technology. By equipping military personnel with helmet mounted visor displays or a special purpose rangefinder [17] the activities of other units participating in the exercise are seen. While looking at the horizon, for example, the display-equipped soldier sees a helicopter rising above the tree line. Actually, another participant is flying this helicopter in simulation. In wartime, the display of the real battlefield scene could be augmented with annotation information or highlighting to emphasize hidden enemy units.

The University of Rochester is participating in the Video Surveillance and Monitoring (VSAM) project funded by the Defense Advance Research Projects Agency (DARPA). Figure 5 shows the scenario for using augmented reality in this project. Aerial reconnaissance units fly overhead and generate reference marks for registration. Autonomous ground units with video surveillance equipment monitor sections of the same area. Aerial views augmented by the information from the surveillance units are generated for strategic command and control. Another scenario equips ground level warfighter or urban crime/terrorist fighters with special displays. Aerial reconnaissance units identify suspect objects and transmit the location of these objects to the ground units. The suspect objects may be hidden from the view of the ground forces but will appear in the augmented view on their displays.
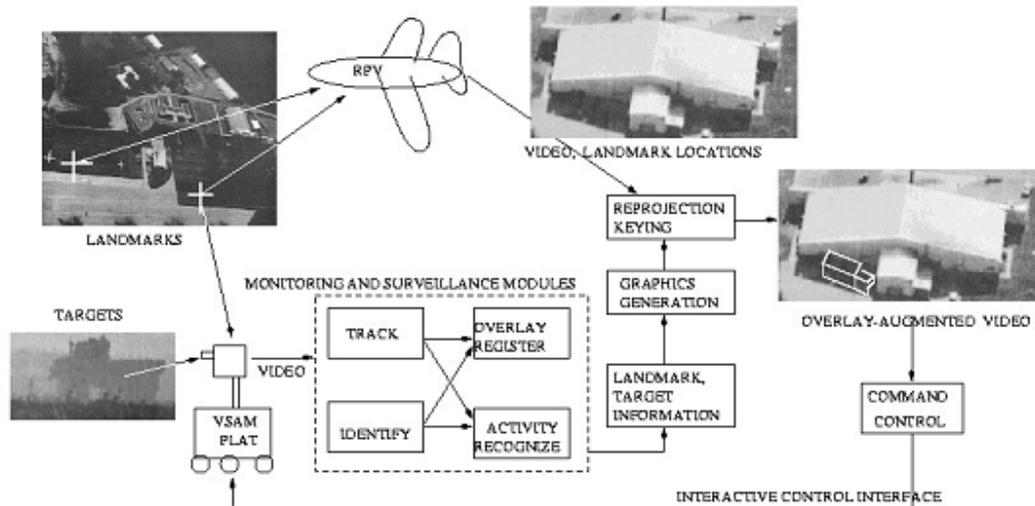
Figure 5. Video Surveillance and Monitoring (VSAM) scenario

3.1.1.4. Engineering Design

Imagine that a group of designers is working on the model of a complex device for their clients. The designers and clients want to do a joint design review even though they are physically separated. If each of them had a conference room that was equipped with an augmented reality display this could be accomplished. The physical prototype that the designers have mocked up is imaged and displayed in the client's conference room in 3D. The clients walk around the display looking at different aspects of it. To hold discussions the client points at the prototype to highlight sections and this is reflected on the real model in the augmented display that the designers are using (Figure 6). Or perhaps in an earlier stage of the design, before a prototype is built, the view in each conference room is augmented with a computer generated image of the current design built from the CAD files describing it. This allows real time interaction with elements of the design so that either side can make adjustments and changes that are reflected in the view seen by both groups [18]. A technique for interactively obtaining a model for 3D objects called 3D stenciling that takes advantage of an augmented reality display is being investigated in our department by Kyros Kutulakos [19].

Figure 6. Engineering design using an augmented reality display

3.1.1.5. Robotics and Telerobotics

In the domain of robotics and telerobotics an augmented display can assist the user of the system [20][21]. Corby and Nafis [22] describe an augmented reality telemanipulation system for nuclear reactor inspection operations. In a telerobotic system, the operator uses a visual image of the remote workspace to guide the robot. Annotation of the view would still be useful just as it is when the scene is in front of the operator. There is an added potential benefit. Since often the view of the remote scene is monoscopic, augmentation with wireframe drawings of structures in the view facilitates visualization of the remote 3D geometry. If the operator is attempting a motion he first practices it on a virtual robot that he sees as an augmentation to the real scene (Figure 7). The operator can decide to proceed with the motion after seeing the results. The robot executes the motion pattern directly which in a telerobotics application eliminates the oscillations often present due to long communication delays to the remote site.
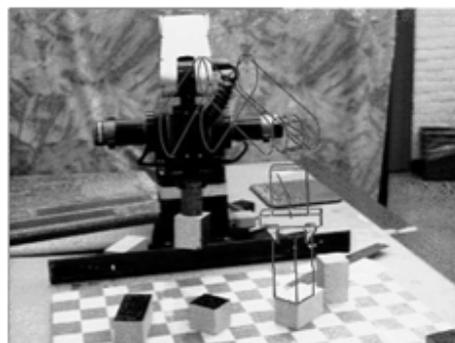

Figure 7. Augmented reality in robotics

3.1.1.6. Manufacturing, Maintenance and Repair

When a maintenance technician approaches a new or unfamiliar piece of equipment instead of opening several repair manuals he simply puts on an augmented reality display. In this display an augmented reality system shows the image of the equipment augmented with

annotations and information pertinent to the repair. For example, the location of fasteners and attachment hardware that must be removed are highlighted. As part of the next step an inside view of the machine highlights the boards that need to be replaced [23]. Figure 8 shows an example of this. The military has developed a wireless vest worn by personnel that is attached to an optical see-through display [17]. The wireless connection allows the soldier to access repair manuals and images of the equipment. Future versions might register those images on the live scene and provide animation to show the procedures that must be performed.
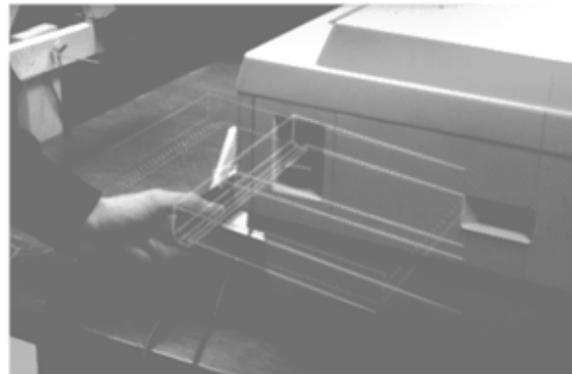

Figure 8. Equipment maintenance application for augmented reality [23]

Aircraft manufacturers are particularly active in incorporating augmented reality systems into their manufacturing and maintenance activities. Boeing researchers developed an augmented reality display to replace the large work frames used for making wiring harnesses for their aircraft [24]. Using this experimental system (Figure 9b), the technicians are guided by the augmented display that shows the routing of the cables on a generic frame used for all harnesses. The augmented display allows a single fixture to be used for making the multiple harnesses. Research undertaken by the University of Southern California [25] is using an augmented reality system to guide technicians through maintenance and repair sequences Figure 9a.
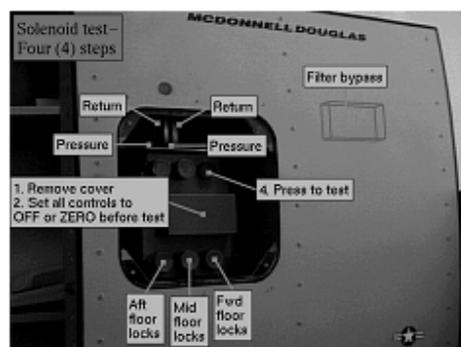
Figure 9 a) b). Aircraft manufacturing use of augmented reality

3.1.1.7. Consumer Applications

Virtual reality systems are already used for consumer design. There are programs available for home computers that assist the homeowner with remodeling projects, such as building a new deck. These computer graphics programs allow you to graphically build the new deck attached to the back of a general model of a house. It is conceivable that a future system would allow you to make a video tape of your house shot from various viewpoints in your backyard. The video is input to your computer so that the design program could show the new deck in its finished form attached to your house. Or use a tape of your current kitchen and the augmented reality program replaces your current kitchen cabinetry with virtual images of the new kitchen that you are designing.

Applications in the fashion and beauty industry that would benefit from an augmented reality system can also be imagined. If the dress store does not have a particular style dress in your size an appropriate sized dress could be used to augment the image of you. As you look in the three sided mirror you see an image of the new dress on your body. You view changes in hem length, shoulder styles or other particulars of the design before you place the order. When you head into some hightech beauty shops today you see what a new hair style will look like on a digitized image of yourself. But with an advanced augmented reality system you see the new styling as you move. If the dynamics of hair are included in the description of the virtual object you also see the motion of your hair as your head moves.

3.1.2. Augmented Reality and Learning

The augmented reality interface combines aspects of virtual reality and the realworld environment by providing a person a chance at viewing one or more virtual 3D objects in real-world space. One of the disadvantages of a completely immersive virtual environment is the loss of "sense of self" in space. That is, in immersive virtual reality one's self is composed of small components of the physical self through the use of avatars or portions of one's body, acting in a space consisting of virtual objects. This has often resulted in feelings of

disorientation and difficulty in movement and intended behaviors in the virtual environment. Active vision theory describes the importance of maintaining a sense of physical presence in order to make behaviors that are based upon information obtained moment-to-moment. In augmented reality, the person is able to combine a virtual 3D object into the normal viewing perspective without losing any of the advantages of object movement and individual movement that creates the behaviors that help us perform activity (gain sensorial-based knowledge) in real-world environments.

Why can't this be done through conventional manipulation in computer terminal virtual object manipulation? In virtual environments experienced through a computer desktop-based application, manipulation of an object is performed through operations that "filter" the manipulation of an object. In other words, in order to rotate or move a virtual object, the person must cognitively "transform" these operations into 1) move mouse cursor over appropriate button, 2) click button, 3) see object orientation change, and 4) process the result in order to create additional mouse clicks. This brief list is greatly simplified to explain the complicated cognitive and motor processes needed in order to make a desktop virtual object change its orientation. The point is that such processes may inhibit the acquisition of visual information. Active vision theory advocates the direct physical manipulation of an object for the effective computation of object recognition--and eventual understanding in accordance with this recognition. This "filtering" effect of desktop interfaces has been postulated to inhibit the effective cognitive processes involved in assimilating and accommodating information. Early studies of this interface have indeed shown that some "visual affordances" are more effectively produced with an augmented reality system than with virtual reality [26].

AR has yet to be taken out of the laboratory and into the classroom and implemented as a potentially powerful educational interface. Hedley's research [26] is based upon students' ability to formulate opinions and remember aspects of spatial information by directly comparing a 2D interface (both paper and computer screen) with AR. This research indicates that AR can improve students' ability to:
• quickly make decisions about relative distances between objects;
• accurately make judgments about relative locations of "landmarks;"
• recall detailed information about an environment, including the color, shape, and texture of objects.

3.2. Learning styles

The three most important learning styles are visual, auditory and kinestezic. The paper describes each of them in the following sections, based on the works of [27] and [28].

3.2.1. Visual Learning Style

Visual learning is a teaching and learning style in which ideas, concepts, data and other information are associated with images and techniques. It is one of the three basic types of learning styles that also includes kinesthetic learning and auditory learning.

Visual learners also prosper when shown graphs, graphic organizers, such as webs, concept maps and idea maps, plots, and illustrations such as stack plots and Venn plots, are some of the techniques used in visual learning to enhance thinking and learning skills.

Visual learners usually possess these qualities:
- have great instinctive direction
- can easily visualize objects

3.2.2. Auditory Learning Style

Auditory learning is a learning style in which a person learns through listening. They may struggle to understand a chapter they've read, but then experience a full understanding as they listen to the class lecture. An auditory learner may benefit by using the speech recognition tool available on many PCs.

Auditory learners may have a knack for ascertaining the true meaning of someone's words by listening to audible signals like changes in tone. When memorizing a phone number, an auditory learner will say it out loud and then remember how it sounded to recall it.

Auditory Learners are good at writing responses to lectures they've heard. They're also good at oral exams, effectively by listening to information delivered orally, in lectures, speeches, and oral sessions. Auditory learners make up about 20% of the population.

It is believed that when an auditory/verbal learner reads, it is almost impossible for the learner to comprehend anything without sound in the background. In these situations, listening to music or having different sounds in the background (TV, people talking, etc) will help learners work better.

Auditory learning can also be considered a different style of learning. Some people are visual learners, some kinesthetic learners, some tactile learners, and some are auditory learners. This is explained as a person who depends on hearing and speaking as a main way of learning. Auditory learners must be able to hear what is being said in order to understand and may have difficulty with instructions that are written. They also use their listening and repeating skills to sort through the information that is sent to them.

Auditory learners are good at storytelling. They solve problems by talking them through. Speech patterns include phrases "I hear you; That clicks; It's ringing a bell", and other sound or voice-oriented information. These learners will move their lips or talk to themselves to help accomplish tasks.

3.2.3. Kinesthetic Learning Style

Kinesthetic learning is a teaching and learning style in which learning takes place by the student actually carrying out a physical activity, rather than listening to a lecture or merely watching a demonstration. It is also commonly known as a "do-er". Some people are visual learners, some kinesthetic learners, and some are auditory learners. Students associated with

this predominant learning style are thought to be natural discovery learners; they have realizations through doing, as opposed to having thought first before initiating action. The evidence on kinesthetic learners benefiting from specialized instruction or targeted materials appears mixed, because the diagnosis of learning preference is itself problematic. However researchers on both sides of the debate agree that there is data showing "that a teaching strategy based on a 'programmed learning sequence' and designed to favor visually- and tactilely-oriented students increased attainment for all students in the experimental group." Other studies also show that mixed modality presentations, for instance using both auditory and visual techniques, improve results for subjects across the board.

Kinesthetic learning is when someone learns things from doing or being part of them. They make up about 15% of the population and struggle to pick things up by reading/ listening to things. Many people mistake themselves for kinesthetic/ tactile learners because they have not used the full variety of learning options, which means they cannot find the right learning state for them. The kinesthetic learner usually does well in things such as chemistry experiments, sporting activities, and acting. They also may listen to music while learning or studying. It is common for kinesthetic learners to focus on two different things at the same time. They will remember things by going back in their minds to what their body was doing. They also have very high hand-eye coordination and very quick receptors. They use phrases such as "I can see myself doing that" and "It's starting to come alive".

4. Similar Work

AR enhances real scenes viewed by the user, overlaying virtual objects over the real world, and works to improve the user's performance in and perception of the world. AR has the following advantages: 1) one can get three-dimensional (3D) information based on a real scene, 2) one can see objects from their own viewpoint, and 3) one can interact with both virtual and real objects. Using the above features, many AR systems have been developed for demonstrations, and some of them have targeted their applications on education.

Earth-Sun Relationships [29] presents seasonal variation of light and temperature, and the virtual sun and earth are manipulated on a small handheld platform that changes its orientation in coordination with the viewing perspective of the student.

Figure 10. Earth-Sun Relationships

Construct3D [30] has designed as a 3D geometry construction tool for mathematics and geometry education, and provided interactive learning environments through various scenarios.

Figure 11. Construct 3D

Augmented Chemistry [31] was a kind of virtual chemistry laboratory where students could view simple atoms and acquire their own complex molecules while being bound by the subatomic rules.

Multimedia Augmented Reality Interface for E-learning (MARIE) [32] has been developed as an application for engineering education in order to enhance traditional teaching and learning methods.

Figure 12. Multimedia Augmented Reality Interface for E-learning (MARIE)

The MagicBook [33] basically used a headmounted display (HMD) as a device of information presentation, though it was adapted for a handheld type.

Figure 13. MagicBook

5. Implementation

5.1. Technologies

5.1.1. ARToolkit and FLARToolkit

ARToolKit is a free, open-source C software library that uses computer vision techniques to calculate camera pose (position and orientation) relative to a black square marker. It is fast enough to perform this calculation at 30 frames per second on a normal desktop. ARToolKit is typically used for applications that augment video of the real world with computer- generated objects. FLARToolkit is the latest in a series of ports of ARToolkit

5.1.2. FLARManager

FLARManager is a lightweight framework that makes building FLARToolkit augmented reality applications easier. When I first came across FLARToolkit, I was simultaneously excited and stymied -- there seemed to be so much potential, but it was difficult to use and even harder to understand. Augmented reality is still a long way from maturity, and leveraging FLARToolkit on the ubiquity of Flash Player provides a great opportunity for further exploring its possibilities. FLARManager aims to provide quicker entry to the technology, to allow a focus on new forms of interaction via marker detection and augmented reality.

FLARManager comes bundled with a series of examples and tutorials. Documentation and walkthroughs can be found in the Inside FLARManager section of my blog. The tutorials move from a quickstart through simple 2D and 3D implementations of FLARManager. This article will provide a more detailed look into the process of marker detection and recognition in FLARToolkit, and how FLARManager can improve the accuracy and reliability of those processes via simple configuration changes.

Computer vision in a web context has many inherent difficulties. The main problems arise from a lack of control over the end users' environmental conditions. Poor or uneven illumination makes software analysis of an image much trickier, and these problems affect FLARToolkit substantially. The last couple of FLARManager releases have focused on functionality that gives developers more control over marker detection and tracking by attempting to mitigate the effects of suboptimal conditions.

5.1.3. Papervision3D

Papervision3D is an open source 3D engine for the Flash platform.

5.1.4. Adobe Flex

Adobe Flex is a software development kit released by Adobe Systems for the development and deployment of cross-platform rich Internet applications based on the Adobe Flash platform. Flex applications can be written using Adobe Flex Builder or by using the freely available Flex compiler from Adobe.

Traditional application programmers found it challenging to adapt to the animation metaphor upon which the Flash Platform was originally designed. Flex seeks to minimize this problem by providing a workflow and programming model that is familiar to these developers. MXML, an XML-based markup language, offers a way to build and lay out graphic user interfaces. Interactivity is achieved through the use of ActionScript, the core language of Flash Player that is based on the ECMAScript standard.

The Flex SDK comes with a set of user interface components including buttons, list boxes, trees, data grids, several text controls, and various layout containers. Charts and graphs are available as an add-on. Other features like web services, drag and drop, modal dialogs, animation effects, application states, form validation, and other interactions round out the application framework.

In a multitiered model, Flex applications serve as the presentation tier. Unlike page-based HTML applications, Flex applications provide a stateful client where significant changes to the view don't require loading a new page. Similarly, Flex and Flash Player provide many useful ways to send and load data to and from server-side components without requiring the client to reload the view. Though this functionality offered advantages over HTML and JavaScript development in the past, the increased support for XMLHttpRequest in major browsers has made asynchronous data loading a common practice in HTML-based development as well.

5.1.5. Adobe Flash

Adobe Flash (formerly Macromedia Flash) is a multimedia platform originally acquired by Macromedia and currently developed and distributed by Adobe Systems. Since its introduction in 1996, Flash has become a popular method for adding animation and interactivity to web pages. Flash is commonly used to create animation, advertisements, and

various web page Flash components, to integrate video into web pages, and more recently, to develop rich Internet applications.

Flash can manipulate vector and raster graphics, and supports bidirectional streaming of audio and video. It contains a scripting language called ActionScript. Several software products, systems, and devices are able to create or display Flash content, including Adobe Flash Player, which is available free for most common web browsers, some mobile phones and for other electronic devices (using Flash Lite). The Adobe Flash Professional multimedia authoring program is used to create content for the Adobe Engagement Platform, such as web applications, games and movies, and content for mobile phones and other embedded devices.

Files in the SWF format, traditionally called "ShockWave Flash" movies, "Flash movies" or "Flash games", usually have a .swf file extension and may be an object of a web page, strictly "played" in a standalone Flash Player, or incorporated into a Projector, a self-executing Flash movie (with the .exe extension in Microsoft Windows or .hqx for Macintosh). Flash Video files have a .flv file extension and are either used from within .swf files or played through a flv-aware player, such as VLC, or QuickTime and Windows Media Player with external codecs added.

5.1.6. Actionscript

ActionScript is a scripting language based on ECMAScript. ActionScript is used primarily for the development of websites and software using the Adobe Flash Player platform (in the form of SWF files embedded into Web pages), but is also used in some database applications (such as Alpha Five), and in basic robotics, as with the Make Controller Kit. Originally developed by Macromedia, the language is now owned by Adobe (which acquired Macromedia in 2005). ActionScript was initially designed for controlling simple 2D vector animations made in Adobe Flash (formerly Macromedia Flash). Later versions added functionality allowing for the creation of Web-based games and rich Internet applications with streaming media (such as video and audio).

5.1.7. Adobe FlashBuilder 4

Flash Builder 4 is Adobe's professional Flex IDE built on Eclipse. It can run either as a standalone tool or as a plug-in to an existing Eclipse installation. Flex SDK includes the Flex framework (also known as the Flex class library), Flex command-line compilers, Adobe Integrated Runtime (AIR) framework, Adobe AIR command-line compilers, the Flex debugger, the ASDoc utility, and the debugger version of Flash Player. Use the Flex SDK to develop, compile, and deploy Flex applications that connect to XML and SOAP web services with no additional charges or server licensing required. This release also includes Data Visualization features.

5.2.1. The FLARToolkit Process

5.2.1.1. Thresholding

The first step in many computer vision applications that rely on edge detection is to threshold the source image. A binary image is made by changing pixels brighter than a threshold value to one color, and pixels darker than the threshold to another.

Figure 14. Thresholding

Thresholding separates the source image into a binary image, making analysis less computationally expensive.

5.2.1.2. Labeling

FLARToolkit's next step is to find contiguous areas in the thresholded image, speficially within the areas below the threshold (darker areas). Using BitmapData.getColorBoundsRect and BitmapData.floodFill, contiguous areas are 'labeled' with unique colors, used later to id the areas.
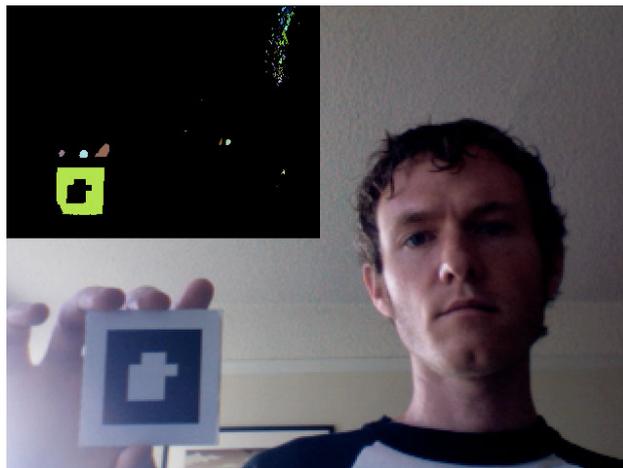

Figure 15. Labeling

Each contiguous area of white (corresponding to dark areas of the source image) is 'labeled' with a different color.

5.2.1.3. Marker outline detection

With candidates for marker locations, FLARToolkit then proceeds to search the labeled areas for shapes that could be transformed squares (i.e. marker outlines).

5.2.1.4. Pattern matching

Once all marker outline possibilities have been established, FLARToolkit analyzes the areas of the image within the outlines and compares the contents with the list of patterns the developer has asked FLARTookit to detect. FLARToolkit assigns a 'confidence' value to all of the matches; matches that are at or above the confidence level specified by the developer are reported as pattern matches.

5.2.2. The FLARManager process

One of the central design philosophies of FLARManager is to avoid modification to the FLARToolkit source as much as possible. By remaining loosely coupled with FLARToolkit, each project can continue development separately, and FLARManager can theoretically be implemented for other Flash marker detection engines that arrive in the future.

To this end, the opportunities for FLARManager to improve on FLARToolkit's process are limited. However, FLARToolkit has free access to both the source image that is passed into FLARToolkit, and to the results that FLARToolkit delivers after analysis. These are the areas on which FLARManager focuses.

All of the settings discussed below can be specified either as member variables of a FLARManager instance, or as attributes of the <flarManagerSettings> node in FLARManager's external xml configuration file.

5.2.2.1. Thresholding

FLARToolkit uses static thresholding: the threshold value does not change from frame to frame. This causes problems when lighting is variable or dim; these are both often the case for web end-users. FLARManager now uses 'adaptive' or 'automatic' thresholding.

Adaptive thresholding is enabled by default in FLARManager. To toggle it, set adaptiveThresholding to either true or false. When adaptive thresholding is active, the threshold value used by FLARToolkit will be changed semi-randomly whenever no markers are detected. There are two values that can be adjusted to tweak the threshold changes.

The first is adaptiveThresholdingSpeed. This value determines the speed at which the threshold is adjusted every frame. Once markers become absent from the source image, the

threshold begins to accelerate away from its current value. If adaptiveThresholdingSpeed is set to a low value (0.0 - 0.3), more gradual changes in the threshold result in an initial lag in marker detection, but a better retention of markers once detected. If adaptiveThresholdingSpeed is set to a higher value (> 0.5), an optimal threshold value may be reached more quickly, but the next threshold change may overshoot the next optimal value. This appears as quicker, but more unstable, marker detection.

The second value is adaptiveThresholdingBias. Rather than completely randomly selecting new threshold values, FLARManager applies this bias to the randomized values, to push the overall direction of change toward lower or higher threshold values. Values for adaptiveThresholdingBias below zero will cause the threshold to trend toward 0, while values above zero will push threshold toward 255. Lower thresholds favor darker environments; higher thresholds work best in brighter environments. adaptiveThresholdingBias defaults to -0.1, as darker environments tend to be more common for viewing web applications than oversaturated environments.

To see the results of experimenting with adaptive thresholding, set FLARManager.thresholdSourceDisplay to true (note: this cannot be set via the xml config file, it must be set in ActionScript). This will overlay the thresholded image on the source image.

Figure 16. FLARManager Thresholding

These images show the effects of changing thresholding. The image at far left has a threshold of 120, and the image at far right has a threshold of 0. Optimal threshold values vary with lighting conditions; adaptiveThresholding is designed to automatically seek out the optimal threshold value for every FLARManager application.

5.2.2.2. Labeling

The labeling process can be very CPU-intensive, as it has to scour the entire thresholded image for dark areas. In fact, FLARLabeling_BitmapData.labeling is the single most CPU-intensive method in FLARToolkit, even when no markers are detected. This is particularly an issue in darker environments. One technique to mitigate this is to slightly blur the source image before thresholding. This tends to increase the size and decrease the number of contiguous areas to label, resulting in higher performance.

This is where sampleBlurring comes into play. This value determines the amount of blur applied to the source image. Specifically, the BlurFilter uses 2 ^ (sampleBlurring-1) for its blurX and blurY values. (Powers of 2 perform much better than other values for BlurFilters.) Higher values result in faster performance, but because the image is more blurred, also result in worse marker detection. sampleBlurring defaults to 1. Setting to zero will turn this feature off entirely.

Figure 16. FLARManager Thresholding

These images show the effects of sampleBlurring on marker detection. The image at far left has a sampleBlurring of 0; the image at far right has a sampleBlurring of 4. Notice that the total number of contiguous black areas decreases as we move to the right, but the legibility of the marker's pattern also decreases.

5.2.2.3. Marker outline detection

Due primarily to inaccurate camera calibration, FLARToolkit has a lot of variability in its reporting of marker outline locations. This seems to be unavoidable, given that the developer has no control over the end user's camera model, or even settings (e.g. focus and zoom). One remaining option for correcting the resulting jitter is to smooth the results.

FLARManager applies smoothing by averaging detected markers' positions across a number of frames. This is smoothing. Higher values average across a greater number of frames, resulting in smoother motion and less jitter, but also in less responsiveness. smoothing defaults to 3 frames, which is just enough to remove jitter, but not so much that responsiveness of moving markers is significantly diminished.

FLARManager also provides an interface that allows developers to write and plug in their own smoothing algorithms. Custom smoothing algorithms can be written up as a class that implements the IFLARMatrixSmoother interface, and applied in a given FLARManager application by setting FLARManager.smoother to an instance of that class (this cannot be done via the config xml file; it must be set in ActionScript).

5.2.2.4. Pattern matching

As described in the labeling section above, sampleBlurring has an impact on pattern matching: higher values result in faster FLARToolkit performance, but also in poorer marker detection.

5.2.2.5. Marker retention

Maintaining reliable tracking of a marker once it's first acquired is not a part of FLARToolkit's core process, but it's worth mentioning here that FLARManager has a feature designed to address this. Developers can specify a delay between the time that FLARToolkit loses a marker and the time that FLARManager dispatches a MARKER_REMOVED event, by

23

using markerRemovalDelay. If a marker with the same pattern reappears near the same location it disappeared from, within markerRemovalDelay frames, the MARKER_REMOVED event wil not be dispatched.

Unfortunately, this is an imperfect workaround for an imperfect system. Increasing markerRemovalDelay pretty quickly results in a breakdown of the illusion of augmented reality, as the objects drawn at the marker's location get left behind when the marker moves quickly. As such, it's not recommended setting markerRemovalDelay much higher than the default of 1 frame, but some applications may be able to make more use of this than others.

5.3. Interaction Proposal

In the following sections are a workflow scenario describing the practical usage of the application, and four diagrams, for possible scenarios of use, depending primarily on the device used for processing.

Workflow scenario:

After being admitted at the university, the student receives a marker representing his faculty. He will use this marker throughout his undergraduate years to enrich his learning through augmented reality. There will be no problem in case of losing the marker, because it will be printable and also available at the faculty's library.

The learning material, for some courses in the beginning, will have some specific markers. The student will be able to combine his faculty marker with the books' marker and benefit of the augmented-reality learning experiences bound with that learning material. He will use the markers combined on the faculty's website, during the laboratories, but also at home, to better understand the complex concepts conveyed in the learning materials.
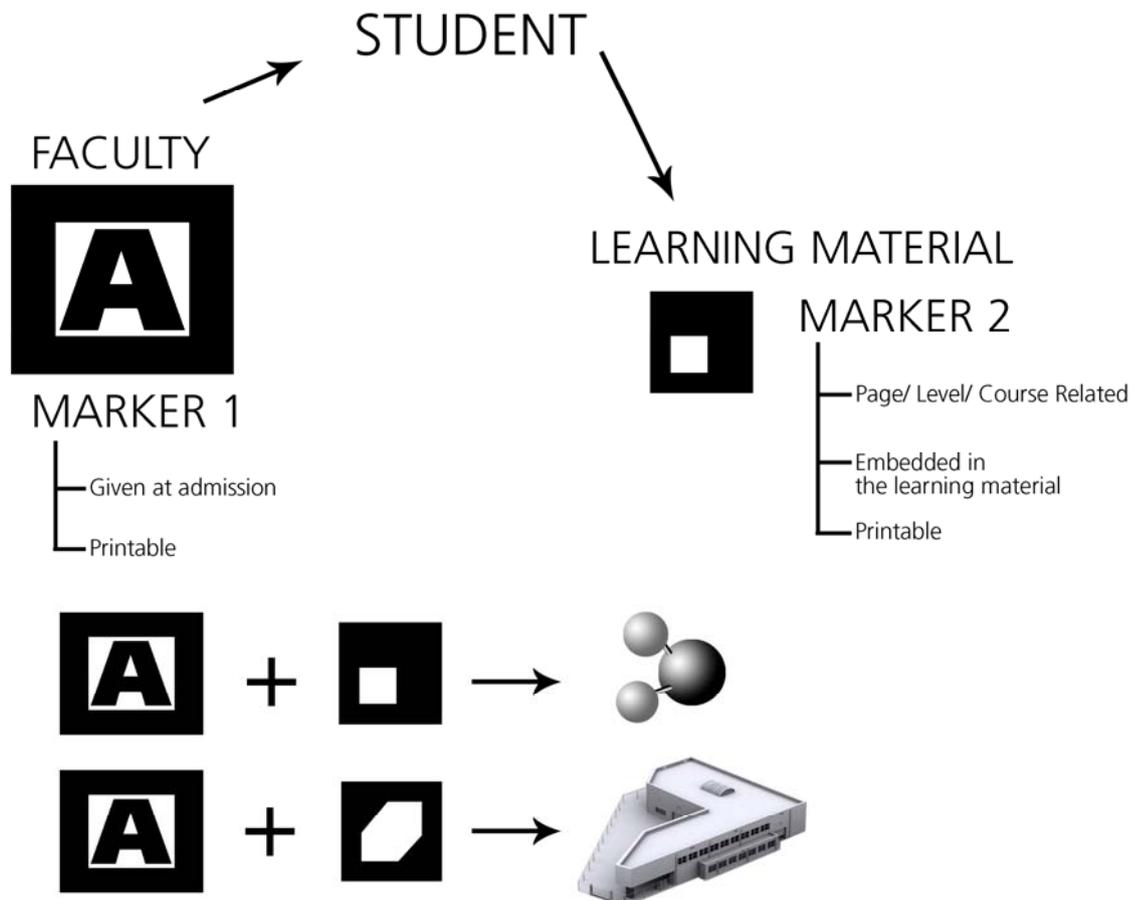

Figure 17: Workflow scenario

Different application usage setups
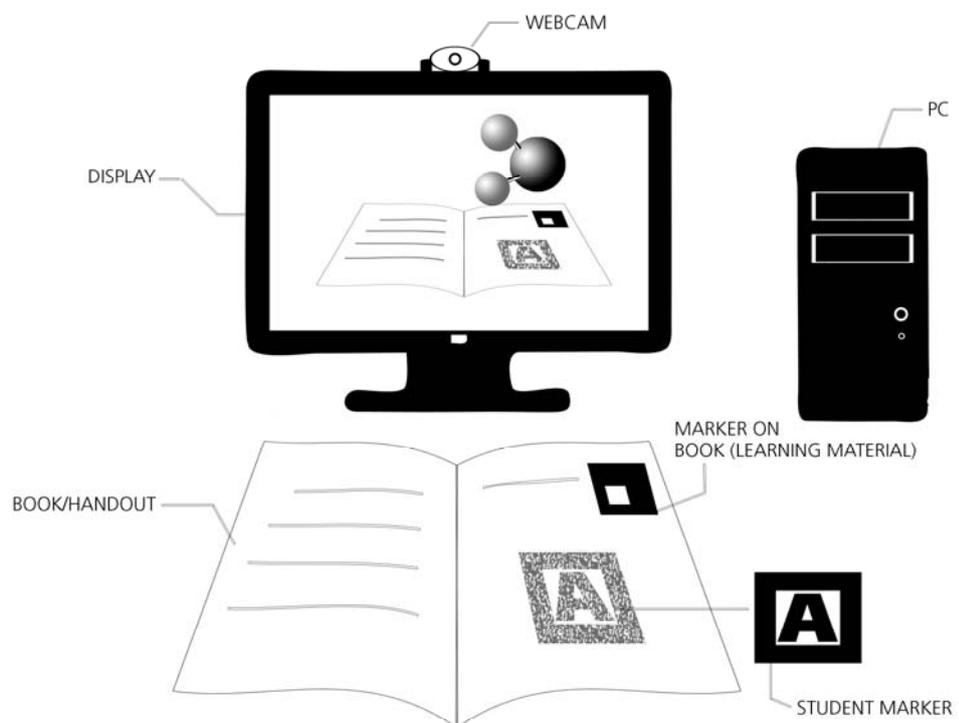
Setup 1:
Desktop PC + Display + Webcam

Figure 18: Setup 1
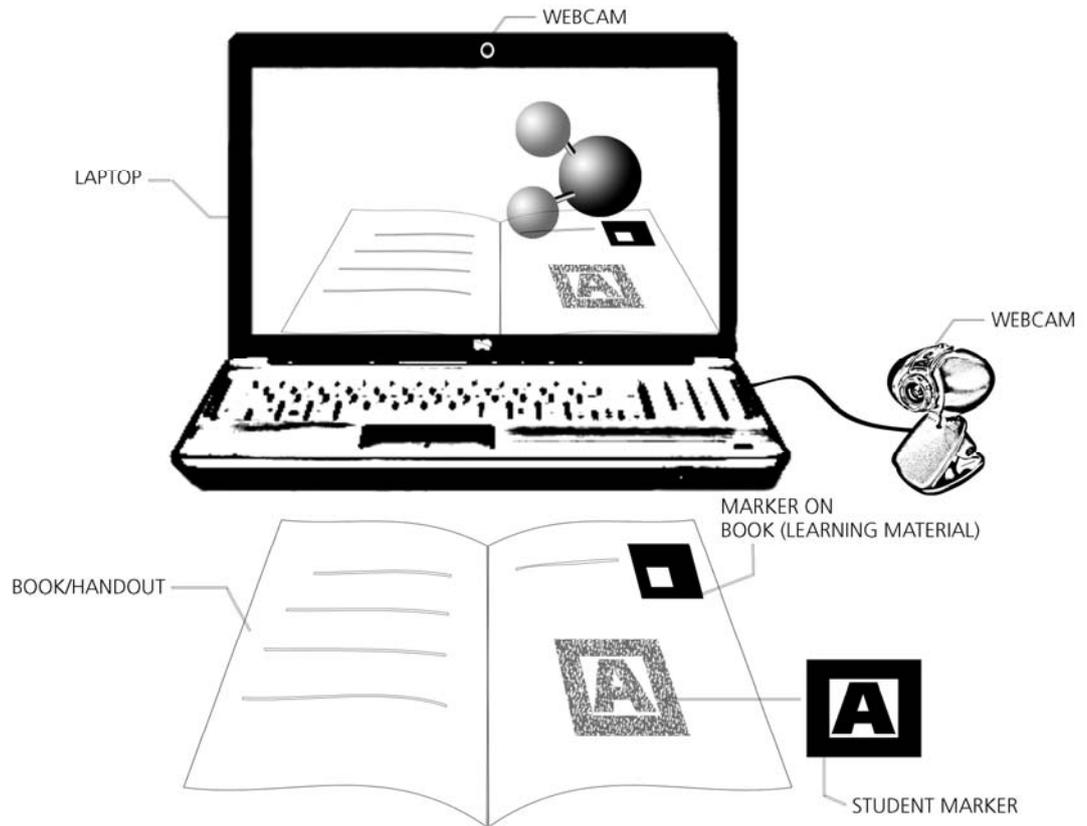
Setup 2:
Laptop (+Webcam)

Fjghre 19: Setup 2

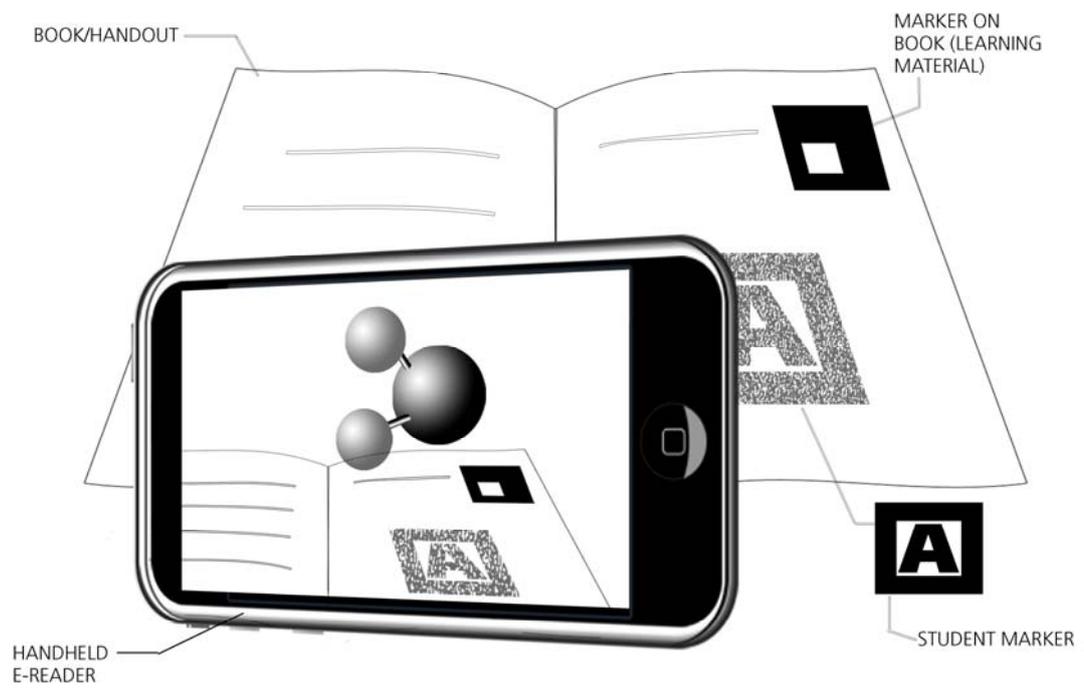Setup 3:
Handheld e-reader device

Figure 20: Setup 3
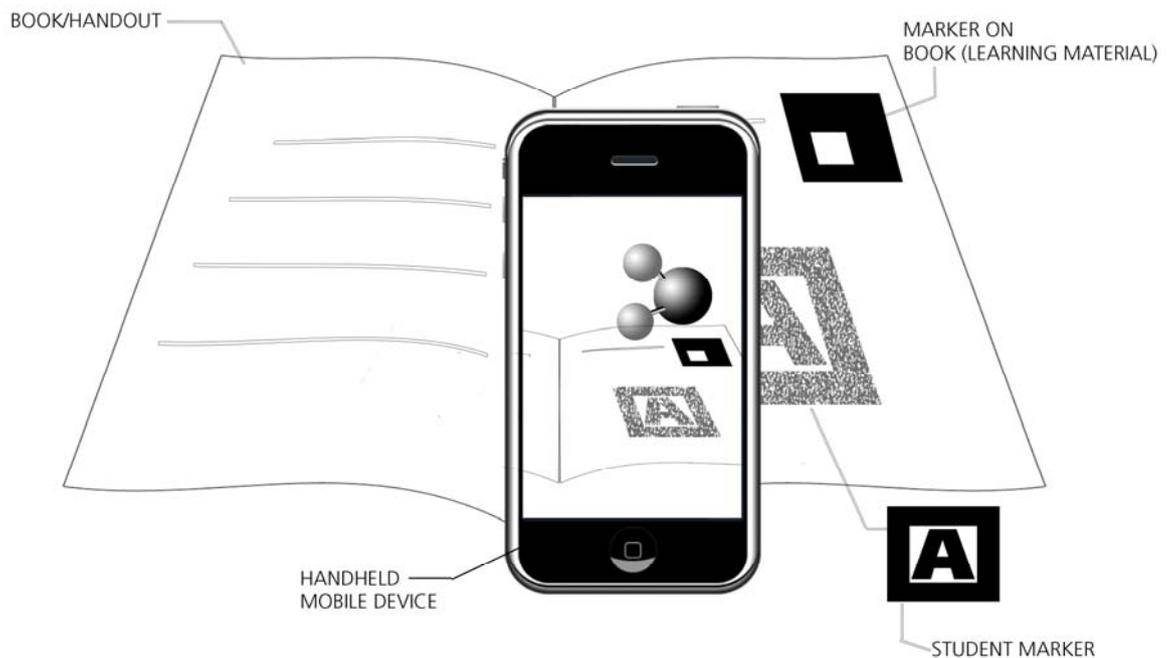
Setup 4:
Handheld mobile device / smartphone

Figure 21: Setup 4

6. Further Development

The research will continue on improving the application, along with deciding on methods for measuring the increase in learning by using the application, among students. Later, after the full development of the application, there will be a full testing period, in which the research will probe the hypotheshis that augmented-reality enables better learning.

7. References

[1] Milgram, P. and F. Kishino (1994), "A Taxonomy of Mixed Reality Visual Displays," IEICE Transactions on Information Systems, E77-D(12).

[2] Milgram, P., H. Takemura, A. Utsumi and F. Kishino (1994), "Augmented Reality: A Class of Displays on the Reality-Virutality Continuum," In Proceedings SPIE Vol. 2351: Telemanipulator and Telepresence Technologies.

[3] Capin, T. K., H. Noser, D. Thalmann, I. S. Pandzic and N. M. Thalmann (1997), "Virtual Human Representation and Communication in VLNet," IEEE Computer Graphics and Applications, 17(2).

[4] Feiner, S., B. MacIntyre, M. Haupt and E. Solomon (1993), "Windows on the World: 2D Windows for 3D Augmented Reality," In Proceedings of ACM Symposium on User Interface Software and Technology.

[5] Mellor, J. P. (1995), Enhanced Reality Visualization in a Surgical Environment, Masters thesis, AI Lab, Massachusetts Institute of Technology.

[6] Lorensen, W., H. Cline, C. Nafis, R. Kikinis, D. Altobelli and L. Gleason (1993), "Enhancing Reality in the Operating Room," In Proceedings of the 1993 IEEE Visualization Conference.

[7] Grimson, W. E. L., T. Lozano-Perez, W. M. W. III, G. J. Ettinger, S. J. White and R. Kikinis (1994), "An Automated Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visulatization," In Proceedings IEEE Conference on Computer Vision and Pattern Recognition.

[8] Grimson, W. E. L., G. J. Ettinger, S. J. White, P. L. Gleason, T. Lozano-Perez, W. M. W. III and R. Kikinis (1995), "Evaluating and Validating an Automated Registration System for Enhanced Reality Visualization in Surgery," In Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine '95.

[9] State, A., D. T. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura and H. Fuchs (1994), "Case Study: Observing a Volume Rendered Fetus within a Pregnant Patient," In Proceedings of the 1994 IEEE Visualization Conference.

[10] Patel, V. V., M. W. Vannier, J. L. Marsh and L.-J. Lo (1996), "Assessing craniofacial surgical simulation," IEEE Computer Graphics and Applications, 16(1).

[11] Gibbs, S. (1995), GMD Digital Media Lab: Virtual Studio.

[12] Balcisoy, S. and D. Thalman (1997), "Interaction between Real and Virtual Humans in Augmented Reality," In Proceedings of Computer Animation '97.

[13] Pyros, G. G. and B. N. Goren (1995), "Desktop Motion Image Processing for Episodic Television," Advanced Imaging, 10(7).

[14] Zorpette, G. (1994), "An eye-popping summer," IEEE Spectrum, 31(10).

[15] Cavallaro, R. (1997), "The FoxTrax Hockey Puck Tracking System," IEEE Computer Graphics and Applications, 17(2).

[16] Jebara, T., C. Eyster, J. Weaver, T. Starner and A. Pentland (1997), "Stochasticks: Augmenting the Billiards Experience with Probabilistic Vision and Wearable Computers," Technical Report TR-439, MIT Media Lab, October 1997.

[17] Urban, E. C. (1995), "The Information Warrior," IEEE Spectrum, 32(11).

[18] Ahlers, K. H., A. Kramer, D. E. Breen, P.-Y. Chevalier, C. Crampton, E. Rose, M. Tuceryan, R. T. Whitacker and D. Greer (1995), "Distributed Augmented Reality for Collaborative Design Applications," Technical Report ECRC-95-05, European Computer Industry Research Center 1995.

[19] Kutulakos, K. N. and J. R. Vallino (1996), "Affine Object Representations for Calibration-Free Augmented Reality," In Proceedings of 1996 IEEE Virtual Reality Annual International Symposium.

[20] Kim, W. S., P. S. Schenker, A. K. Bajczy, S. Leake and S. Ollendorf (1993), "An Advanced Operator Interface Design with Preview/Predictive Displays for Ground-Controlled Space Telerobotic Servicing," In Proceedings of SPIE Vol. 2057: Telemanipulator Technology and Space Telerobotics.

[21] Milgram, P., S. Zhai, D. Drascic and J. J. Grodski (1993), "Applications of Augmented Reality for Human-Robot Communications," In Proceedings of 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems.

[22] Corby, N. R., Jr. and C. A. Nafis (1994), "Augmented reality telemanipulation system for nuclear reactor inspection," In Proceedings SPIE Vol. 2351: Telemanipulation and Telepresence Technologies, pages 360-365, 31 October -1 November 1994.

[23] Feiner, S., B. MacIntyre and D. Seligmann (1993b), "Knowledge-Based Augmented Reality," Communications of the ACM, 36(7).

[24] Caudell, T. P. (1994), "Introduction to Augmented Reality," In SPIE Proceedings Vol. 2351: Telemanipulator and Telepresence Technologies, November 1994.

[25] Neumann, U. and Y. Cho (1996), "A Self-Tracking Augmented Reality System," In Proceedings of ACM Symposium on Virtual Reality Software and Technology.

[26] Hedley, N. R. (2003). "Three-dimensional geographic visualization and spatial mental models."

[27] Vincent, A., & Ross, D. (2001). "Learning Style Awareness". Journal of Research on Computing in Education, 33.

[28] Dunn, R, & Dunn, K (1978). "Teaching students through their individual learning styles: A practical approach". Reston, VA: Reston Publishing Company.

[29] B. E. Shelton, and N. R. Hedley (2002), "Using augmented reality for teaching Earth-Sun relationships to undergraduate geography students", Proc. The First IEEE International Augmented Reality Toolkit Workshop.

[30] H. Kaufmann (2002), "Construct3D: an augmented reality application for mathematics and geometry education", Proc. 10th ACM international Conference on Multimedia.

[31] M. Fjeld, P. Juchli, and B. M. Voegtli (2003), "Chemistry education: a tangible interaction approach", Proc. INTERACT2003.

[32] F. Liarikapis, P. Petridis, P. F. Lister, and M. White (2002), "Multimedia augmented reality interface for e-learning (MARIE)", World Transactions on Engineering and Technology Education, vol.1.

[33] M. Billinghurst, H. Kato, and I. Poupyrev (2001), "The MagicBook: a traditional AR interface", Computers & Graphics, vol.25.