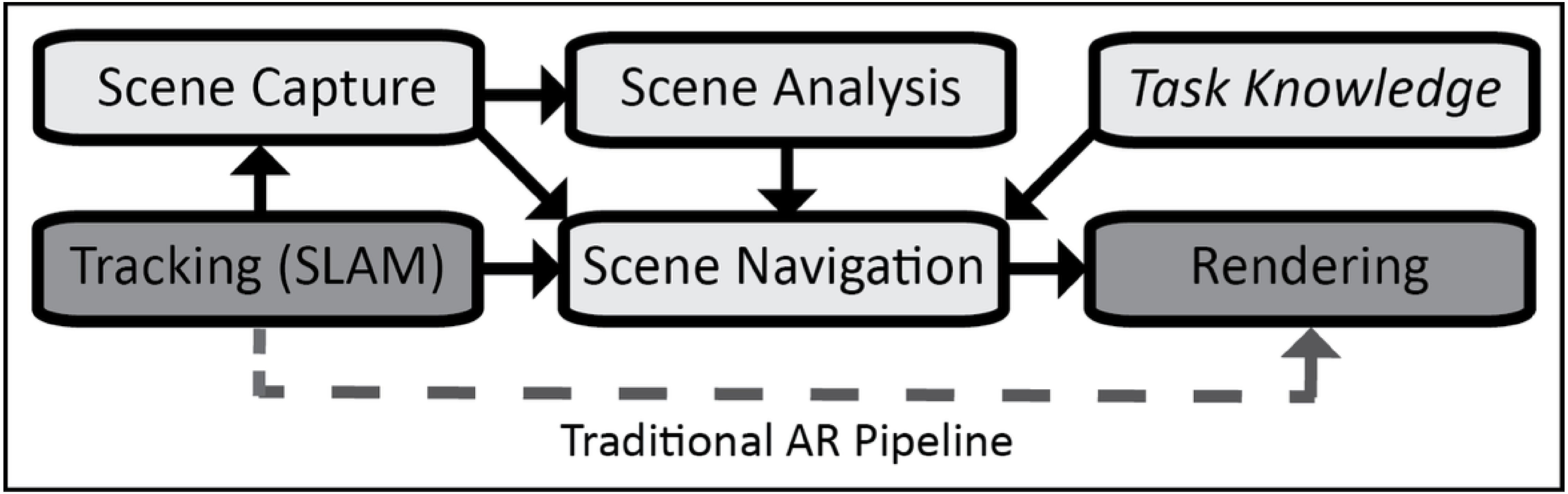
Module II
Concepts of Augmented Reality

-Scene Capture

Augmented Reality (AR) applications require knowledge about the real world environment in which they are used. This knowledge is often gathered while developing the AR application and stored for future uses of the application. Consequently, changes to the real world lead to a mismatch between the previously recorded data and the real world. New capturing techniques based on dense Simultaneous Localization and Mapping (SLAM) not only allow users to capture real world scenes at run-time, but also enables them to capture changes of the world. However, instead of using previously recorded and prepared scenes, users must interact with an unprepared environment. New interaction techniques that support users in handling captured real world environments. The techniques present virtual viewpoints of the scene based on a scene analysis and provide natural transitions between the AR view and virtual viewpoints. We demonstrate our approach with a SLAM based prototype that allows us to capture a real world scene and describe example applications of our system.

Concepts of Augmented Reality-Scene Capture

Coordinate based 3D multimedia applications benefit from cost effective, compact and easy-to-use profilers like the miniaturized 3D-Camera that works on basis of the fringe projection technique. The system uses a compact housing and is usable like a video camera with minimum stabilization like a tripod. Camera and projector are assembled with parallel optical axes having coplanar projection and imaging planes. Their axes distance is comparable to the human eyes' distance, giving a compact system of shoebox-size and allow measuring high gradient objects like the interior of tubes and delivering captured scenes with minimum loss by shadowing. Additionally, the 3D-Camera can be used for the Inverse Projection technique, allowing single-frame video rate capture and to virtually place information like virtual labels or defect maps onto the surface of objects, thus, allowing augmented reality applications. the concept and realization of the 3D-Camera is described and an overview of possible applications is given.



Traditional AR Pipeline

Augmented Reality (AR) applications traditionally rely on predefined, rigidly modelled content that only applies to those conditions of the real world that were present when the application was developed. Hence, often an AR application can be deployed only in a single physical location, which limits its flexibility. This problem is aggravated when changes in the environment cause misalignments between the previously recorded data and the real world. Such a situation can impair the correct functioning of the application.

CAPTURING THE SCENE

- Virtual representations of real world scenes can be captured in real-time using SLAM technology. The combination of SLAM with depth sensors, allows live capturing of detailed models, either as volumes, depth images or polygon meshes. The output of an open source Kinect Fusion implementation applied to the scene Models can be constructed with monocular SLAM technology, stereo SLAM, or depth sensors.

- Alternatively, models can also be reconstructed using an online reconstruction service. While capturing the scene, the data is processed in the cloud and sent back to the user.

- A mesh approximation of the volume generated by an open source Kinect Fusion implementation. We reduce the number of voxels by using a grid filter that calculates the centroid of voxels within a cube having a side length of $1cm$. Then, we use a poison mesh algorithm using the filtered points to create the mesh representation. We chose the tree depth used for the poison reconstruction to find a trade off between reconstruction performance and accuracy of the created geometry. A tree depth of six allows for more frequent mesh updates, but the details of the geometry are smoothed.

- However, the general shape of the object is approximated well. A tree depth of seven creates a good approximation of the objects at the cost of a lower update frequency. The user to switch at runtime between these settings. Hence, a user may choose faster updates, when the real world scene is rearranged frequently, or a better mesh approximation with higher visual quality after the scene was arranged.

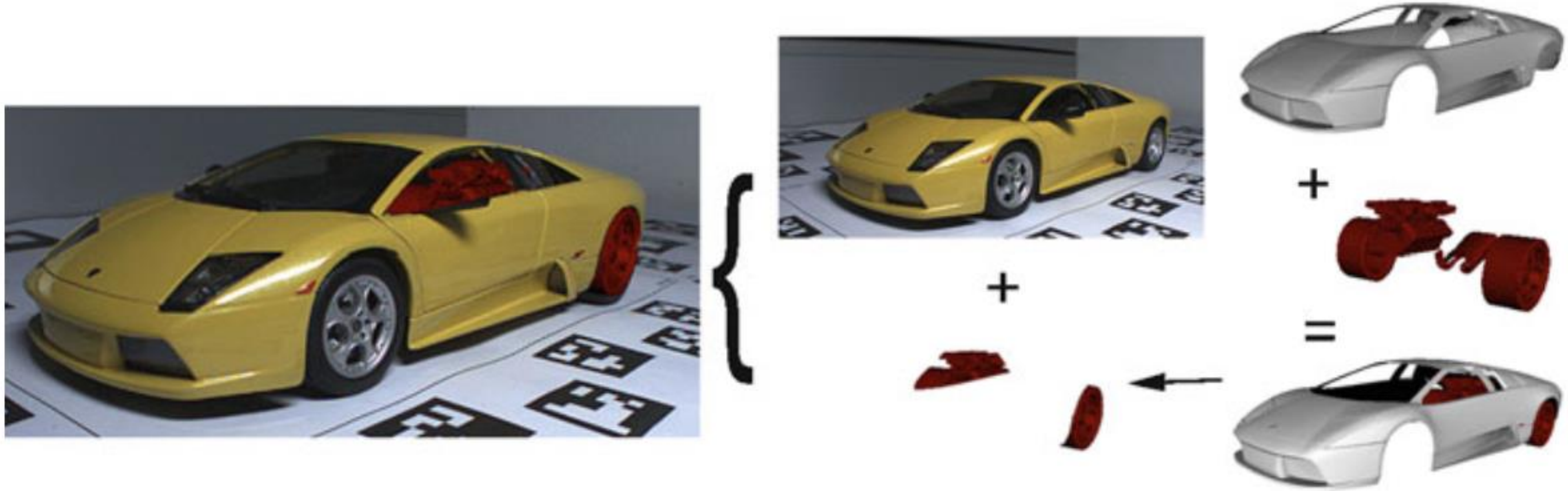
Scene Visualization

Visualizations in real world environments benefit from the visual interaction between real and virtual imagery. However, compared to traditional visualizations, a number of problems have to be solved in order to achieve effective visualizations within Augmented Reality (AR). This chapter provides an overview of techniques to handle the main obstacles in AR visualizations. It discusses spatial integration of virtual objects within real world environments, techniques to rearrange objects within mixed environments, and visualizations which adapt to its environmental context

Visualization Techniques for Augmented Reality

Augmented Reality (AR) applications enrich the real world environment with a certain amount of synthetic information, ideally just enough to overcome the limitations of the real world for a specific application. Azuma et al. Define three requirements of an Augmented Reality application. They claim that in addition to a mixture of real and virtual imagery, AR applications have to run in real time and virtual objects have to be aligned (registered) with real world structures. An example of a common AR system and the data which is acquired, computed, and presented. In order to register the virtual monster, the AR system derives tracking information from the video input. After rendering the registered 3D structure, its overlay allows to generate the impression of a virtual figure standing on a real world paper card.

This kind of visualization is a powerful tool for exploring real world structures along with additional contextual information. For example, by augmenting textual annotations, AR displays are able to provide semantics to real world objects or places. AR visualizations may also use the real world environment to provide



Data Integration

A simple overlay of hidden structure on top of the system's video feed can cause a number of cognitive problems, caused by the processes involved in creating the impression of depth. Understanding these causes allows to develop rendering techniques which successfully add and preserve such information in AR visualizations.

Depth Perception

Our cognitive system takes approximately 15–20 different psychological stimuli into account in order to perceive spatial relationships between 3D objects. These so called depth cues can be divided into monocular and binocular. While binocular depth cues require the use of two eyes, monocular cues appear even when one eye is closed.

Augmenting Pictorial Depth Cues

By rendering the virtual structure using a camera which uses parameters reflecting the characteristics of the real camera, the fusion of virtual and real world imagery will automatically provide pictorial depth cues which match to those present in the real world environment.

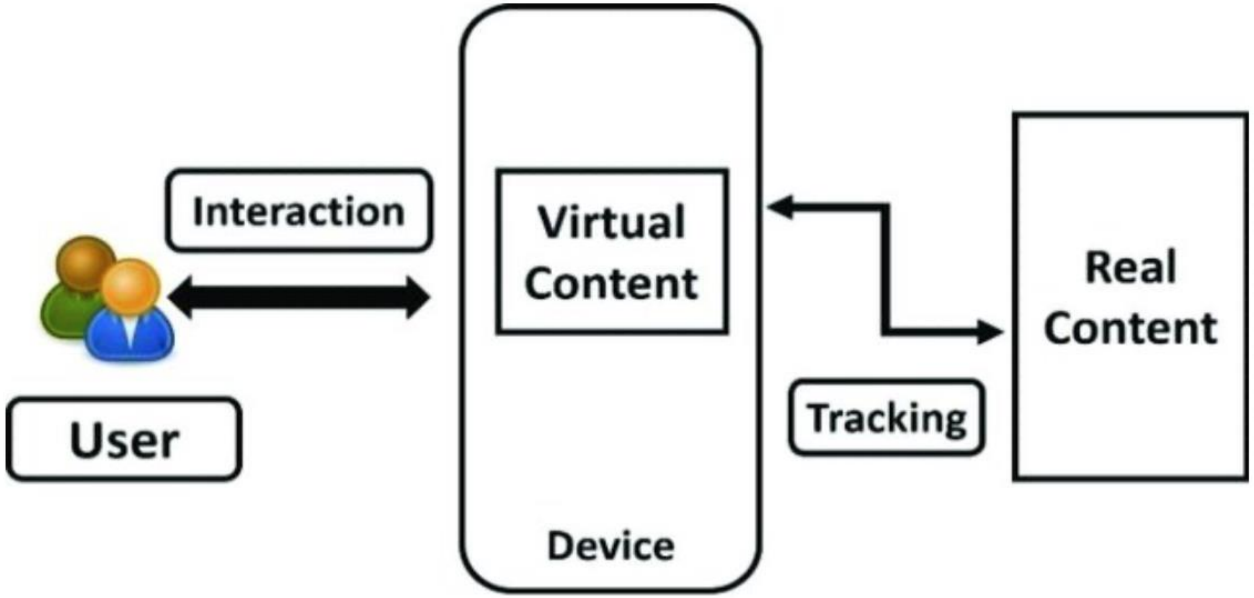
Augmented Reality architecture

AR can aid in visualizing building projects. Computer-generated images of a structure can be superimposed onto a real-life local view of a property before the physical building is constructed there; this was demonstrated publicly by Trimble Navigation in 2004. AR can also be employed within an architect's workspace, rendering animated 3D visualizations of their 2D drawings. Architecture sight-seeing can be enhanced with AR applications, allowing users viewing a building's exterior to virtually see through its walls, viewing its interior objects and layout.

With continual improvements to GPS accuracy, businesses are able to use augmented reality to visualize georeferenced models of construction sites, underground structures, cables and pipes using mobile devices. Augmented reality is applied to present new projects, to solve on-site construction challenges, and to enhance promotional materials. Examples include the Daqri Smart Helmet, an Android-powered hard hat used to create augmented reality for the industrial worker, including visual instructions, real-time alerts, and 3D mapping.

Components of Augmented Reality architecture

Example: The Christchurch earthquake, the University of Canterbury released City ViewAR, which enabled city planners and engineers to visualize buildings that had been destroyed. This not only provided planners with tools to reference the previous cityscape, but it also served as a reminder of the magnitude of the resulting devastation, as entire buildings had been demolished.



Components of Augmented Reality architecture

Augmented reality (AR) as a technological enabler is becoming very popular in different fields such as education, design, navigation and medicine. It combines virtual information with the real environment in the real-time performance [1] and enhances the user's perception such as vision, hearing, touch and smell with the combination of real world. Due to the development of the mobile devices, AR is growing rapidly and become more mature and robust. To ensure the acceptance and success of future AR system, it is vital to understand the end users' experience and requirements. AR Research priorities have shifted toward the design of effective and ease-of-use application with the growth of elderly mobile users, increasing trends in using AR system among older people have been observed. It becomes significant to design the suitable augmented reality application to bring benefits for older people in some of the areas. However, attempts to investigate how to develop the appropriate AR applications have been fairly limited especially those with any focus upon ageing population

An architecture of augmented reality

Augmented reality is growing rapidly and supports people in different fields. However, because of the broad meaning for this term, it is not exactly clear what is meant or intended when people use AR. Some of the sources use a very general meaning for it, whereas others mean something very specific and narrow. Putting different papers together, it gradually becomes clear to identify the whole map of AR architecture applied to any AR experience. This section introduces the six elements of augmented reality including user, interaction, device, virtual content, tracking and real content, in order to create the boundary of AR and ascertain the options available in the use of AR for particular activities. Hopefully, the AR architecture comprised by six elements could serve as to innovate and develop the actual AR application. The elements of AR have been followed by a brief summary of each meaning, features of each and detailed explanation including examples.

User

AR technology is to provide artificial stimuli to cause the users to believe that something is occurring in the virtual world. Take Tesco store finder created by Junaio AR browser as an example: the purpose of this application is to provide the users with the awareness of Tesco's location, opening time, website and further information. Users are playing an important role in what takes place in AR architecture. Normally, AR architecture could happen in the mind of one user. But sometimes it could be two or many. Construct 3D was a geometric construction AR system which aimed to promote students spatial capability due to the fact that they could view geometric entities in different sides. Two or many users could use an electric pen to modify all geometric entities, find the geometric relationship and work on the construction together. The concept of collaborative augmented reality is based upon two or many AR users. AR tennis is another application engaged with two users: two players could sit across the table from one another and hold their phone to view a virtual court overlaying the real world background and play the tennis game. Many AR researches focus on the normal users without particular group of people Some of research designed AR application for children with autism and cognitive disability.

Interaction

Interaction is composed of two components including inter- and action. Inter means the state between or among things. Action means there is an influence and something that has been done. Therefore, interaction can be simply speaking that one entity does something and the other entity responds in some way. In the user-based augmented reality (AR) system, the process of interaction in this research mainly concentrates on a trigger caused by users and the response of AR system which can occur between users and AR device or users and virtual content. The action of adjusting the AR device's physical position can be described as the interaction between *users and AR device interaction*. This action will end in a response of identifying the virtual target store on the device. Then, if people want to get more information about the particular Tesco store (e.g. opening time, distance or phone number), they need click a virtual icon that visually indicates the store's information. After that, some more overlaid information will be presented in another pop-up slide. The action of clicking the virtual Tesco bubble means *users and virtual content interaction*.

Device

The term of device means the carrier or object could acquire physical world information and provide the compelling augmentation. It could be the *mobile device, desktop computer, and big screen with projector* or etc. stated that there were three hardware functions in all AR device including sensors, processors and displays. Sensors recognize the state of the physical world which the AR system needs to be deployed. For example, a camera, one of the most popular AR sensors, could capture the physical world image and provide information to the AR users. GPS or other compass system aims to identify the location and orientation of user. A processor processes the sensors' information and generates the signals required to drive the display. Very often, the AR system will rely upon not just the processor on the device, but the processor on the server as well. A display will show the coexistence that users could sense the combination of physical world and virtual world. Based on these requirements of functions, the smart mobile phone or tablet seems to be the appropriate AR device compromising by a camera to capture, processors to process and a screen to display. The mobile device held by one hand could run different applications, which is moveable, easy to use and accessible from anywhere and anyplace.

Virtual content

Virtual content means the digital information presented by AR device, which plays the most important role in the AR architecture. The modality of virtual content could *be 3D animation, 2D image, text, website, audio information or even vibration*. AR users will not concern too much on devices, but will be attracted by different virtual content. Odom and Pierce state that participants often strongly express curiosity to what a digital device could provide but rarely if ever affection to the device itself. A key feature of virtual content is that the virtual information can be changed dynamically. Going back to the Tesco example, when users use the app and move around, the virtual content (Tesco bubble) will pop up automatically. An icon visually indicating a real-world store can be clicked and more Store information will be presented. This additional information like the opening time website or the video instruction replaces the previous virtual bubble. The content of virtual information has been changed easily and completely.

Real content

Real content is the real-world information directly presented by device without any rendering, which includes *geographic location, physical objects and real-world environment*. Taking Tesco app example again, AR devices not only display the virtual Tesco bubble (virtual content), but present the users' location information surrounding the real-time environment, which means real content. However, while users look through the AR device, the real content will be more or less hidden For example, Word lens AR translator generates the virtually translated words, which replace the real-world words. Users have to remove the AR device if they need to see the original words. They cannot see both the virtual and real content simultaneously. Obstruction of real content is the intrinsic risk of augmented reality.

Tracking

Tracking describes the way of generating virtual content based on the real content, comprising three different features: synchronicity, antecedent and partial one to one. Due to the changes of the real content, an AR virtual counterpart has to be updated synchronously. For example, Word lens is an AR translation application that scans and foreign text and displays the test translated in real time. Once the user changes his or her point of view to another word, the displayed translation on the device rapidly changes in the same time. If the process of generating virtual information is delayed for a long time, viewers are unable to obtain the useful information. The feature of antecedent means the real content (physical text) exists or happens before the virtual content (the translated digital word). If the virtual content is created before the real-world content, the virtual element is meaningless because it has no real world interpretation. Partial one to one describes another tracking feature of augmented reality. here is one and only one real content to correspond with the virtual content. However, there might be one or more than one piece of virtual information to correspond with the real-world content. That means the real content can be superimposed to different modality of virtual content.

Augmented Reality systems and functionality

Hardware

Hardware components for augmented reality are: a processor, display, sensors and input devices. Modern mobile computing devices like smartphones and tablet computers contain these elements, which often include a camera and Microelectromechanical systems (MEMS) sensors such as an accelerometer, GPS, and solid state compass, making them suitable AR platforms. There are two technologies used in augmented reality: diffractive waveguides and reflective waveguides.

Display

Various technologies are used in augmented reality rendering, including optical projection systems, monitors, handheld devices, and display systems, which are worn on the human body.

A head-mounted display (HMD) is a display device worn on the forehead, such as a harness or helmet-mounted. HMDs place images of both the physical world and virtual objects over the user's field of view. Modern HMDs often employ sensors for six degrees of freedom monitoring that allow the system to align virtual information to the physical world and adjust accordingly with the user's head movements.[25][26][27] HMDs can provide VR users with mobile and collaborative experiences.[28] Specific providers, such as uSens and Gestigon, include gesture controls for full virtual immersion

Eyeglasses

AR displays can be rendered on devices resembling eyeglasses. Versions include eyewear that employs cameras to intercept the real world view and re-display its augmented view through the eyepieces and devices in which the AR imagery is projected through or reflected off the surfaces of the eyewear lens pieces

HUD

A head-up display (HUD) is a transparent display that presents data without requiring users to look away from their usual viewpoints. A precursor technology to augmented reality, heads-up displays were first developed for pilots in the 1950s, projecting simple flight data into their line of sight, thereby enabling them to keep their "heads up" and not look down at the instruments. Near-eye augmented reality devices can be used as portable head-up displays as they can show data, information, and images while the user views the real world. Many definitions of augmented reality only define it as overlaying the information. This is basically what a head-up display does; however, practically speaking, augmented reality is expected to include registration and tracking between the superimposed perceptions, sensations, information, data, and images and some portion of the real world

Contact lenses

Contact lenses that display AR imaging are in development. These bionic contact lenses might contain the elements for display embedded into the lens including integrated circuitry, LEDs and an antenna for wireless communication. The first contact lens display was patented in 1999 by Steve Mann and was intended to work in combination with AR spectacles, but the project was abandoned then 11 years later in 2010–2011 Another version of contact lenses, in development for the U.S. military, is designed to function with AR spectacles, allowing soldiers to focus on close-to-the-eye AR images on the spectacles and distant real world objects at the same time

Virtual retinal display[edit]

A virtual retinal display (VRD) is a personal display device under development at the University of Washington's Human Interface Technology Laboratory under Dr. Thomas A. Furness III. With this technology, a display is scanned directly onto the retina of a viewer's eye. This results in bright images with high resolution and high contrast. The viewer sees what appears to be a conventional display floating in space.

EyeTap

- The EyeTap captures rays of light that would otherwise pass through the center of the lens of the wearer's eye, and substitutes synthetic computer-controlled light for each ray of real light.

- The Generation-4 Glass (Laser EyeTap) is similar to the VRD (i.e. it uses a computer-controlled laser light source) except that it also has infinite depth of focus and causes the eye itself to, in effect, function as both a camera and a display by way of exact alignment with the eye and resynthesis (in laser light) of rays of light entering the eye.

Handheld

A Handheld display employs a small display that fits in a user's hand. All handheld AR solutions to date opt for video see-through. Initially handheld AR employed fiducial markers, and later GPS units and MEMS sensors such as digital compasses and six degrees of freedom accelerometer–gyroscope. Today Simultaneous localization and mapping (SLAM) markerless trackers such as PTAM (parallel tracking and mapping) are starting to come into use. Handheld display AR promises to be the first commercial success for AR technologies. The two main advantages of handheld AR are the portable nature of handheld devices and the ubiquitous nature of camera phones. The disadvantages are the physical constraints of the user having to hold the handheld device out in front of them at all times, as well as the distorting effect of classically wide-angled mobile phone cameras when compared to the real world as viewed through the eye.

Spatial

Spatial augmented reality (SAR) augments real-world objects and scenes, without the use of special displays such as monitors, head-mounted displays or hand-held devices. SAR makes use of digital projectors to display graphical information onto physical objects. The key difference in SAR is that the display is separated from the users of the system. Since the displays are not associated with each user, SAR scales naturally up to groups of users, allowing for collocated collaboration between users.

Tracking

Modern mobile augmented-reality systems use one or more of the following motion tracking technologies: digital cameras and/or other optical sensors, accelerometers, GPS, gyroscopes, solid state compasses, Radio-frequency identification (RFID). These technologies offer varying levels of accuracy and precision. The most important is the position and orientation of the user's head. Tracking the user's hand(s) or a handheld input device can provide a 6DOF interaction technique

Networking

Mobile augmented reality applications are gaining popularity because of the wide adoption of mobile and especially wearable devices. However, they often rely on computationally intensive computer vision algorithms with extreme latency requirements. To compensate for the lack of computing power, offloading data processing to a distant machine is often desired. Computation offloading introduces new constraints in applications, especially in terms of latency and bandwidth. Although there are a plethora of real-time multimedia transport protocols, there is a need for support from network infrastructure as well.

Input devices

Techniques include speech recognition systems that translate a user's spoken words into computer instructions, and gesture recognition systems that interpret a user's body movements by visual detection or from sensors embedded in a peripheral device such as a wand, stylus, pointer, glove or other body wear. Products which are trying to serve as a controller of AR headsets include Wave by See bright Inc. and Nimble by Intugine Technologies.

Computer

The computer analyzes the sensed visual and other data to synthesize and position augmentations. Computers are responsible for the graphics that go with augmented reality. Augmented reality uses a computer-generated image which has a striking effect on the way the real world is shown. With the improvement of technology and computers, augmented reality is going to lead to a drastic change on ones perspective of the real world.

Projector

Projectors can also be used to display AR contents. The projector can throw a virtual object on a projection screen and the viewer can interact with this virtual object. Projection surfaces can be many objects such as walls or glass panes.

Software and algorithms

A key measure of AR systems is how realistically they integrate augmentations with the real world. The software must derive real world coordinates, independent of camera, and camera images. That process is called image registration, and uses different methods of computer vision, mostly related to video tracking. Many computer vision methods of augmented reality are inherited from visual odometry. An **augogram** is a computer generated image that is used to create AR. **Augography** is the science and software practice of making augograms for AR.

Development

The implementation of augmented reality in consumer products requires considering the design of the applications and the related constraints of the technology platform. Since AR systems rely heavily on the immersion of the user and the interaction between the user and the system, design can facilitate the adoption of virtuality. For most augmented reality systems, a similar design guideline can be followed.

Environmental/context design

Context Design focuses on the end-user's physical surrounding, spatial space, and accessibility that may play a role when using the AR system. Designers should be aware of the possible physical scenarios the end-user may be in such as:

- Public, in which the users use their whole body to interact with the software
- Personal, in which the user uses a smartphone in a public space
- Intimate, in which the user is sitting with a desktop and is not really moving
- Private, in which the user has on a wearable

Interaction design

Interaction design in augmented reality technology centers on the user's engagement with the end product to improve the overall user experience and enjoyment. The purpose of interaction design is to avoid alienating or confusing the user by organizing the information presented. Since user interaction relies on the user's input, designers must make system controls easier to understand and accessible. A common technique to improve usability for augmented reality applications is by discovering the frequently accessed areas in the device's touch display and design the application to match those areas of control.[91] It is also important to structure the user journey maps and the flow of information presented which reduce the system's overall cognitive load and greatly improves the learning curve of the application

Visual design

In general, visual design is the appearance of the developing application that engages the user. To improve the graphic interface elements and user interaction, developers may use visual cues to inform the user what elements of UI are designed to interact with and how to interact with them. Since navigating in an AR application may appear difficult and seem frustrating, visual cue design can make interactions seem more natural.

In some augmented reality applications that use a 2D device as an interactive surface, the 2D control environment does not translate well in 3D space making users hesitant to explore their surroundings. To solve this issue, designers should apply visual cues to assist and encourage users to explore their surroundings

Module II-Revision

Thank You