Unit - V

AR and MR

UNIT V - Augmented and Mixed Reality: Taxonomy - technology and features of augmented reality - Difference between AR and VR - Challenges with AR - AR systems and functionality - Augmented reality methods - visualization techniques for augmented reality - wireless displays in educational augmented reality applications - mobile projection interfaces - marker-less tracking for augmented reality - enhancing interactivity in AR environments - evaluating AR systems.

What is AR?

Augmented Reality (AR) is a live, direct or indirect view of a physical, real-world environment whose elements are augmented (or supplemented) by computer-generated sensory input such as sound, video, graphics or GPS data.

AR exists on top of our own world it provides as much freedom as you are given within your normal life. AR utilizes your existing reality and adds to it utilizing a device of some sort.

Mobile and tablets are the most popular mediums of AR now, through the camera, the apps put an overlay of digital content into the environment. Custom headsets are also being used.

Popular AR examples : Pokemon Go and Snapchat's new AR bitmojis.

What is VR?

- **Virtual Reality (VR)** is an immersive experience also called a computer-simulated reality.
- It refers to computer technologies using reality headsets to generate the realistic sounds, images and other sensations that replicate a real environment or create an imaginary world.
- VR is a way to immerse users in an entirely virtual world. A true VR environment will engage all five senses (taste, sight, smell, touch, sound), but it is important to say that this is not always possible.
- Gaming industry uses this technology into more practical applications.
- The market and the industry are still excited about this tech trend and further progress is expected in the near future

What is MR?

- **Mixed Reality (MR),** sometimes referred to as hybrid reality, is the merging of real and virtual worlds to produce new environments and visualizations where physical and digital objects co-exist and interact in real time.
- It means placing new imagery within a real space in such a way that the new imagery is able to interact, to an extent, with what is real in the physical world we know.
- The key characteristic of MR is that the synthetic content and the real-world content are able to react to each other in real time.

What is XR?

Real-and-virtual combined environments and human-machine interactions generated by computer technology and wearables.

- Extended Reality includes all its descriptive forms like the Augmented Reality (AR), Virtual Reality (VR), Mixed Reality (MR).
- In other words, XR can be defined as an umbrella, which brings all three Reality (AR, VR, MR) together under one term, leading to less public confusion. Extended reality provides a wide variety and vast number of levels in the Virtuality of partially sensor inputs to Immersive Virtuality.

	VR	AR/MR
PC	 VIVE	 CINOPTICS
Standalone	 oculus	 magic leap
Smartphone	 SAMSUNG Gear VR	 IKEA

Application & Challenges

Applications:

- Education
- Science
- Business and manufacturing
- Medicine
- Public safety and military
- Art
- Advertising
- Entertainment

https://www.youtube.com/watch?v=2ZMg0mfEw-k

Milgram's Reality-Virtuality Continuum

Mixed Reality (MR)

Real Environment → Augmented Reality (AR)

Augmented Virtuality (AV) ← Virtual Environment

Milgram coined the term "Augmented Virtuality" to identify systems which are mostly synthetic with some real world imagery added such as texture mapping video onto virtual objects.

Key aspects of augmented reality:

■ The physical world is augmented by *digital* information superimposed on a view of the physical world.

■ The information is displayed in registration with the physical world.

■ The information displayed is dependent on the location of the real world and the physical perspective of the person in the physical world.

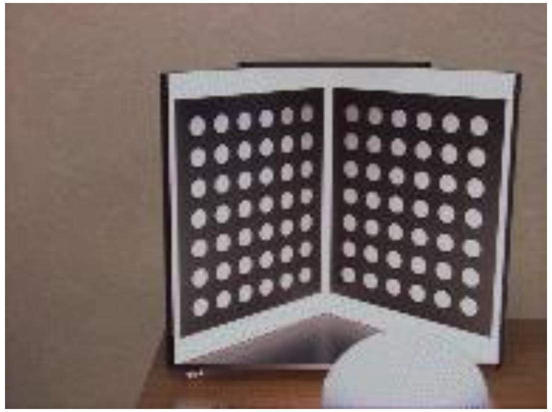
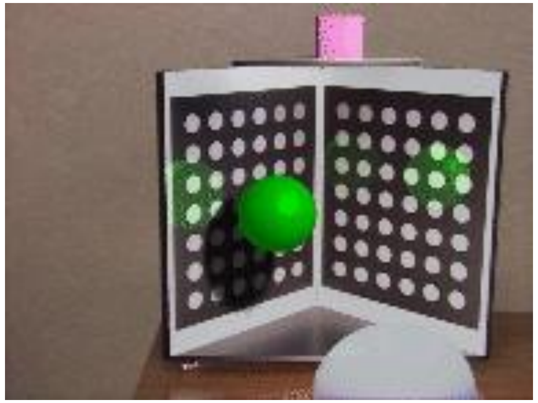
■ The augmented reality experience is interactive, that is, a person can sense the information and make changes to that information if desired. The level of interactivity can range from simply changing the physical perspective (e.g., seeing it from a different point of view) to manipulating and even creating new information.

Characteristics that define augmented reality:

1. Combines real and virtual
2. Interactive in real time
3. Registered in 3D

1. Combining the Real and Virtual Worlds

- Precise models
- Locations and optical properties of the viewer (or camera) and the display
- Calibration of all devices
- To combine all local coordinate systems centered on the devices and the objects in the scene in a global coordinate system
- Register models of all 3D objects of interest with their counterparts in the scene
- Track the objects over time when the user moves and interacts with the scene

2. Interactive

- See, hear, smell, touch, and taste the physical world.
- The digital information is added to, or *superimposed on* the physical world.
- Different technologies can be used to display the digital information, but a key aspect is that they don’t fully occlude the physical world. For example, if augmented reality is experienced with a display headset, the physical world can be seen either because the display is made in such a way that you can see through the display (optical see-through) or because a video camera captures the real world from your point of view and displays that view onto your headset (video see-through). In the case of audio information, headphones or earpieces can be utilized that don’t block out physical world sounds.

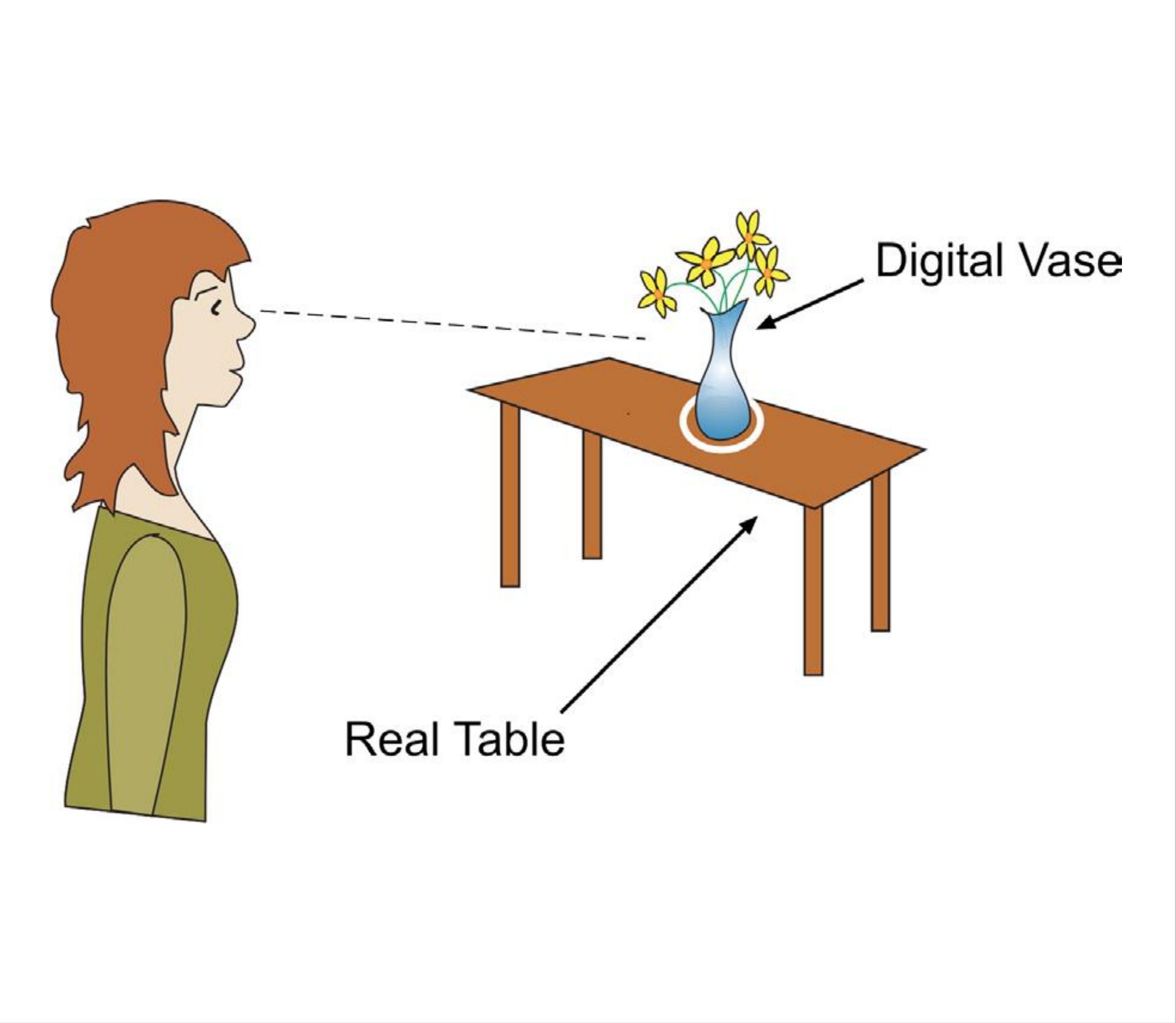
Two basic modes in which the digital world and the physical world can be merged.

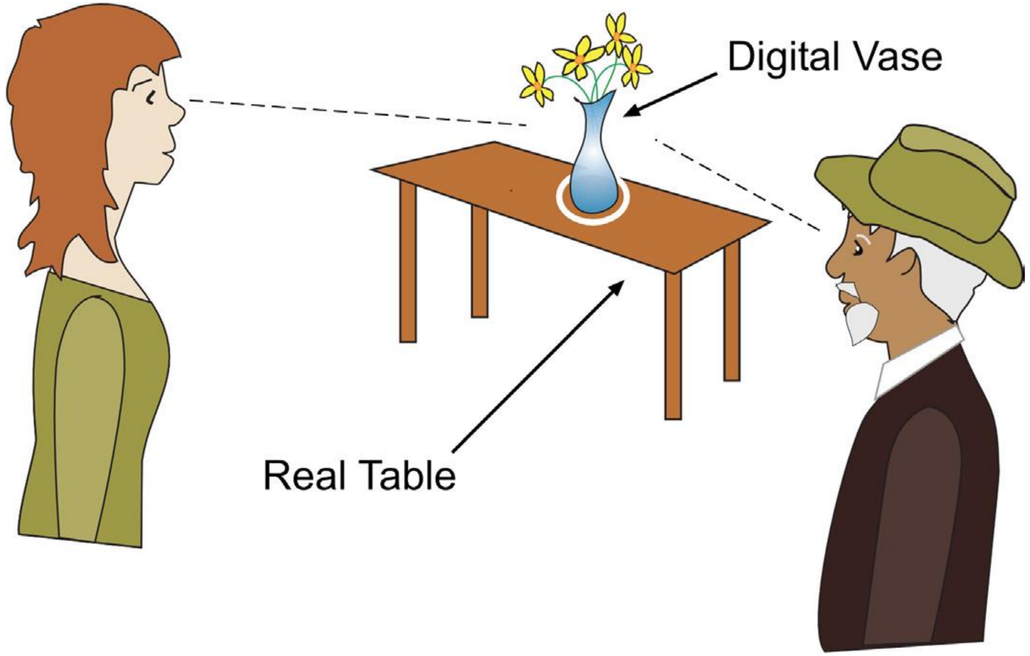
1. Gather information from the physical world, generate the digital information in the computer, and meld those two worlds together in a computer to be displayed.

2. Simply project the augmentations onto the physical world by means of projection devices (both stationary and mobile projection displays).

3. Registration

- The information has a physical space or location in the real world just like a physical counterpart to the digital information would have.

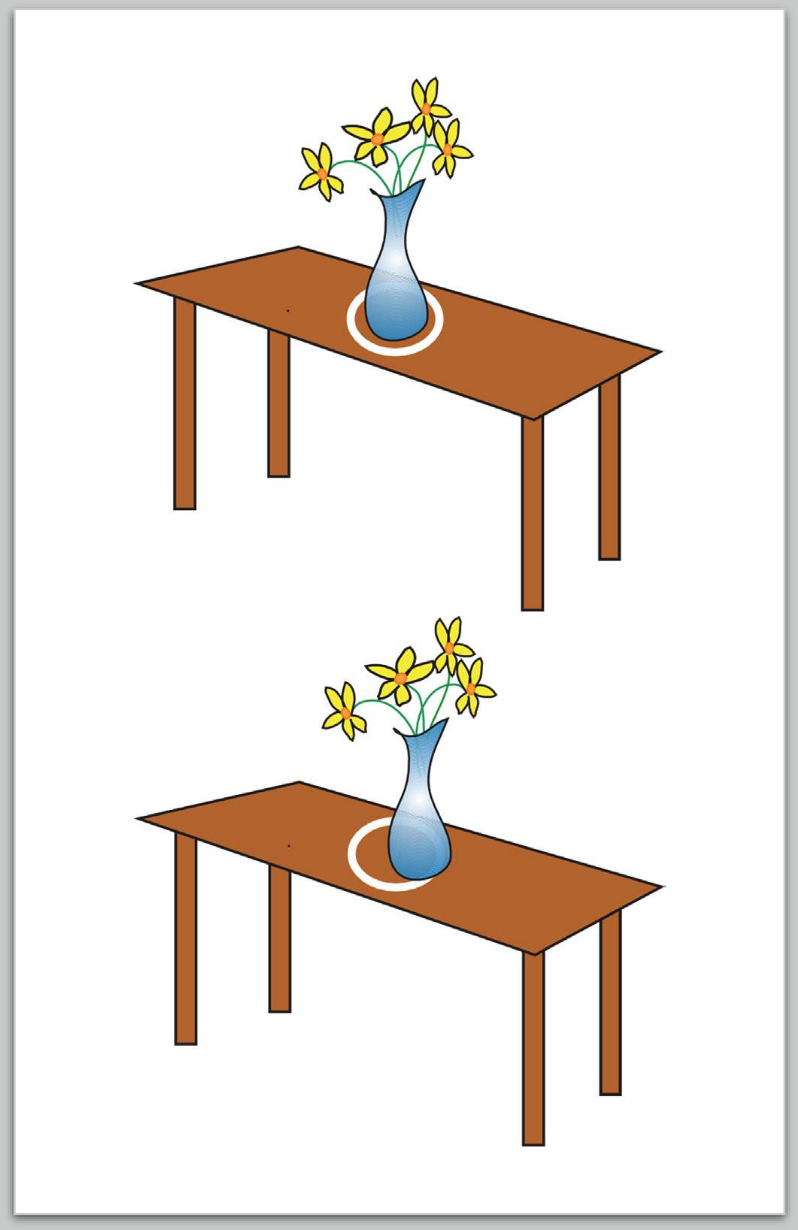




Two people are looking at a digital representation of a vase on a real table. Each person sees the vase from his or her own perspective. Of course, in augmented reality anything is possible, so it could be made such that each person sees entirely different things. In the conventional mode, however, where two people are interacting in the same world, each person sees the vase on the table from his or her own point of view, just like they would see a physical vase on a table from their own point of view.

Registration with the real world

- Close registration - digital object should be placed in the physical world to very close tolerances. The tightness of the tolerances depends on the application.
- Temporal registration - more difficult to achieve due to the inherent time lags involved with processing the information. Because the view of an object depends on the participant's physical point of view, the object must be re-rendered every time the viewer changes position even a tiny bit. If there is a lag in the system, and the viewer changes his or her perspective very rapidly, the lag in the computing might provide a noticeable lag in the scene where the participant sees the scene "catch up" with his or her change in point of view. Such lags can be somewhat overcome in systems that provide a video feed of the physical world by delaying the view of the physical world by the same amount as the lag in the digital representation, but latency will still be involved.



- Spatial registration can be either
  - *absolute* (a very specific place on the globe) or
  - *relative* (with respect to some entity).
- An example of an application that uses absolute spatial registration would be an application that allows you to see an as of yet not constructed building in place at the actual address in a specific city. An example of an application with relative spatial registration would be an application that lets you see the internals of some object. For example, an application that lets you hold a calculator in your hand but lets you see the circuit board and components inside it would use relative spatial registration. The registration is done with respect to that object rather than an absolute place on earth, that is, no matter where you take the object, the internals are in registration with that object correctly, but it is not tied to any specific place on the planet.

Realistic Merging

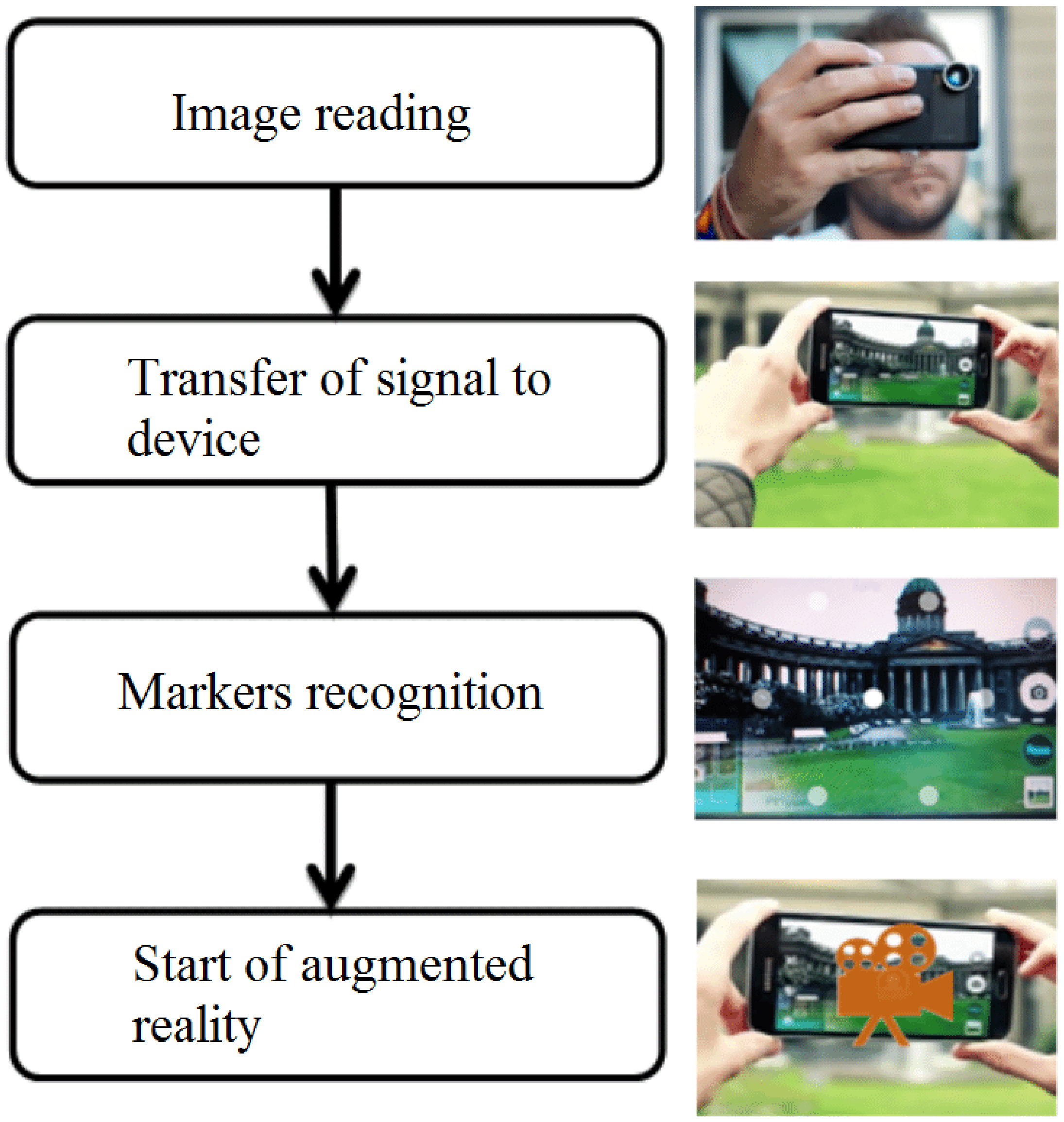
Requires:

- Objects to behave in physically plausible manners when manipulated
- Occlusion
- Collision detection
- Shadows

Relationship between Augmented Reality and Other Media, Technologies, and Ideas

	Classification	Fully Synthetic Virtual World	Fully Real Virtual World	Absolute Spatial Registration Critical	Relative Spatial Registration Critical	Real-Time Interactivity Critical
Augmented reality	Medium (supported by technology)	No	No	Maybe	Yes	Yes
Virtual reality	Medium (supported by technology)	Yes	No	No	Yes	Yes
Telepresence	Medium (supported by technology)	No	Yes	Maybe	Yes	Yes
Global positioning system	Technology	N/A	N/A	Yes	Yes	No
Geographic information system	Technology	N/A	N/A	No	Yes	No
Fabrication	Technology	No	Yes	No	No	No
Cyberspace	Descriptive idea	No	Maybe	No	No	Maybe
Mixed reality	Descriptive idea	Maybe	Maybe	Maybe	Maybe	Maybe
Virtuality	Descriptive idea	Maybe	Maybe	Maybe	Maybe	Maybe
Metaverse	Descriptive idea	Maybe	Maybe	Maybe	Maybe	Maybe

Working of AR



Augmented reality application consists:

1. Augmented reality application - the computer program that orchestrates and controls different aspects of the augmented reality experience.

   An AR browser might do something as simple as making it appear that a specific 3D object is placed on a specific fiducial marker placed in the real world. It is important to draw a distinction between the AR application and the *content* used within the application. When this distinction is executed skillfully, the same AR application can be used in many different contexts. The AR application also interacts with the various sensors, devices, and displays used in the experience. In actual practice, many of these lower level tasks are handled in AR libraries that are used by many different AR applications.

2. **Content** - Content is key to any AR application. Content includes all objects, ideas, stories, sensory stimuli, and “laws of nature” for the experience. The laws of nature govern what actions take place during the experience. This may include computational simulations, game rules, or any other aspects of the content that are under computer control.

3. **Interaction** - Every AR experience must be interactive in one way or another. One of the most typical ways the world is interactive is that it allows the participant to view/perceive the world from different physical points of view. Beyond this basic interaction, participants may interact with the experience by pressing buttons, by making physical gestures, by speaking commands, or by any number of different actions.

4. **Technology** - Every AR experience does involve technology. Some require much more sophisticated technology than others, but all have a base level of technology. At a minimum, an AR experience requires some sort of sensor to gather information about the real world, some form of computation to integrate the virtual elements of the experience with the real world, and some mechanism to display the virtual elements of the experience.

5. Physical world - Every AR experience takes place in the physical world. The physical world is a key part of the AR experience. The physical world may or may not be an exact place (although it can be), but in some cases a generic space is used to represent the physical world at large. That is, in some cases the experience must take place where the experience indicates.

For example, if an AR experience is to decorate the (real) Eiffel Tower like a Christmas tree, then you must be near the Eiffel Tower. In other cases, the experience can take place virtually anywhere. One could conceive of an application in which one decorates a (virtual) Eiffel Tower where the (virtual) Eiffel Tower is in a football field in the United States instead of in France.

6. Participant(s) - All of the magic of an AR experience takes place in the mind of one or more participants. Indeed, the role of AR technology is to provide artificial stimuli to cause the participant(s) to believe that something is occurring that really is not. Things that don’t really exist in the physical realm have characteristics that cause them to be perceived as though they do. The participant(s) has an active role in what takes place in the AR experience. All of their motions, actions, and activities affect how the system responds.

When there are multiple human participants, the stakes are even higher in terms of complexity of interaction. Of course, there might be participants that other participants believe are real but are actually digital creatures—mere augmentations to the physical world.

- Adoption of AR try-on technology has massively increased

https://econsultancy.com/14-examples-augmented-reality-brand-marketing-experiences/

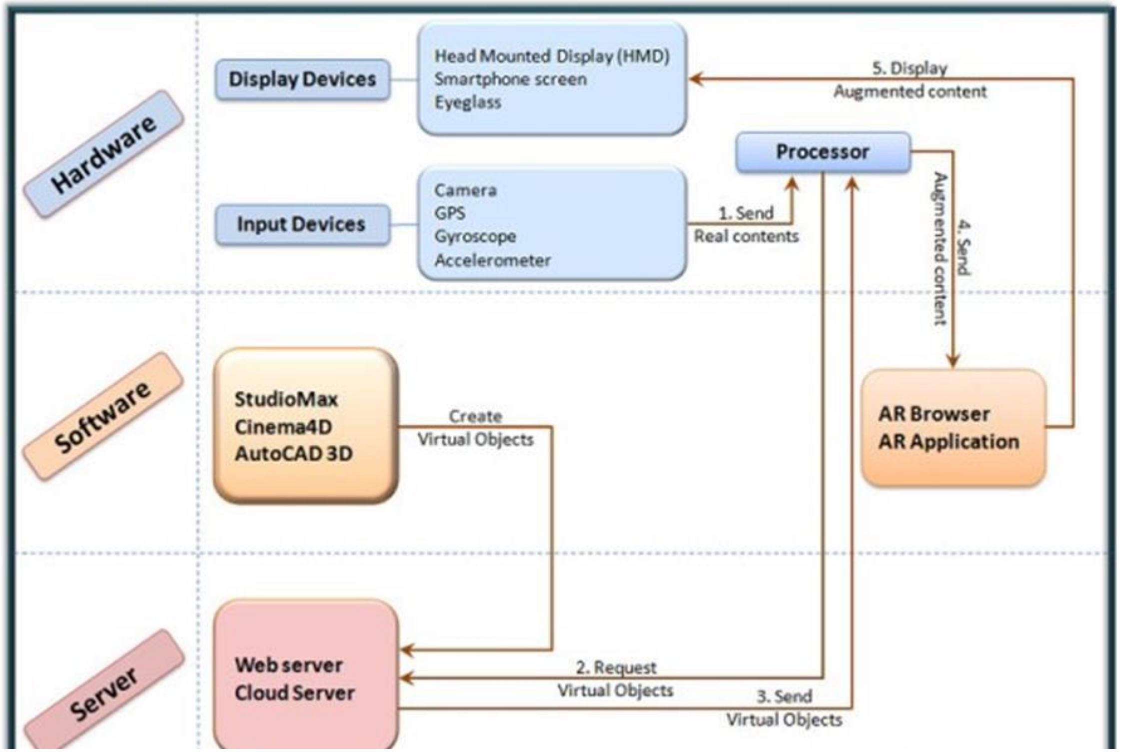
E.g.,

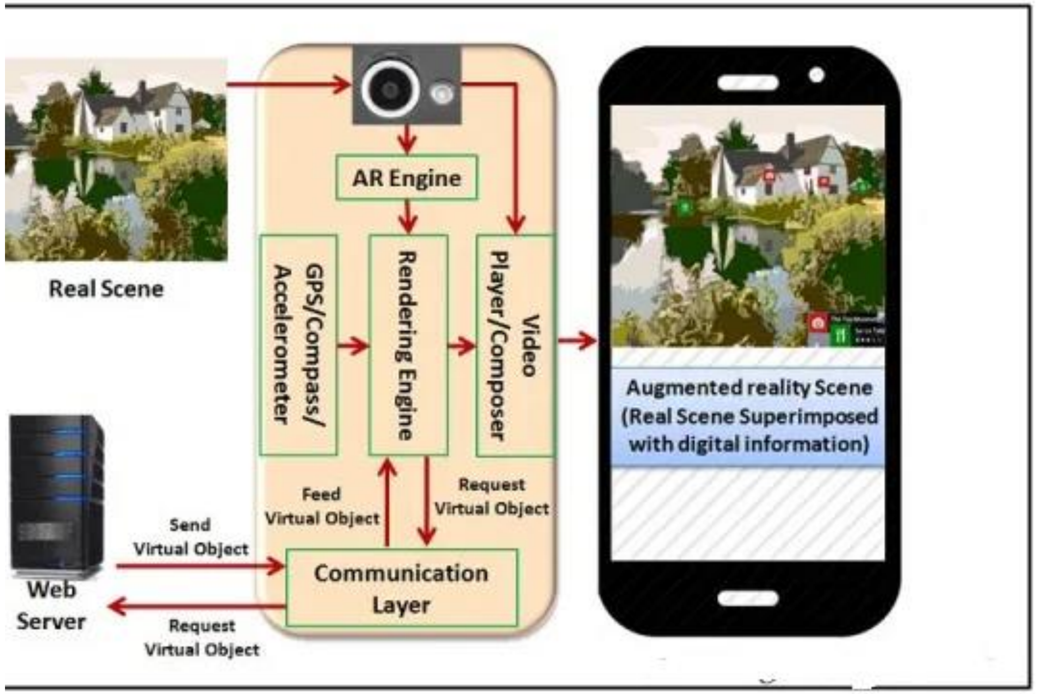
https://youtu.be/kHiI-oQvB5c

Flipkart

Lenskart

Lakme try-on





Components of AR

Example

Primary steps in AR application

1. The application needs to determine the current state of the physical world and determine the current state of the virtual world.

2. The application needs to display the virtual world in registration with the real world in a manner that will cause the participant(s) to sense the virtual world elements as part of his or her physical world and then return to step 1 to move on to the next time step.

Hardware components :

1. *Sensor(s)* to determine the state of the physical world where the application is deployed

2. *A processor* to evaluate the sensor data, to implement the "laws of nature" and other rules of the virtual world, and to generate the signals required to drive the display

3. *A display* suitable for creating the impression that the virtual world and the real world are coexistent and to impinge on the participant's senses such that he or she senses the combination physical world and virtual world.

4. *Input devices* such as buttons, keyboards, and other actuators. However, these are actually sensors in which the participant takes an active role in setting their values. Other sensors are more passive in nature, in that the participant doesn't take an active role in determining their values.

I - Sensors : to respond correctly to the physical world, an augmented reality application must have information about the real world in real time.

- Three primary categories of sensors are used in AR systems:

1. *Sensors used for tracking*-AR depends on being spatially registered, there must be some mechanism to determine information about the position of the participant, the real world, and any AR devices. P*osition* includes both *location* and *orientation*. (To determine position requires information about six degrees of freedom of the entity being tracked)

    - Camera

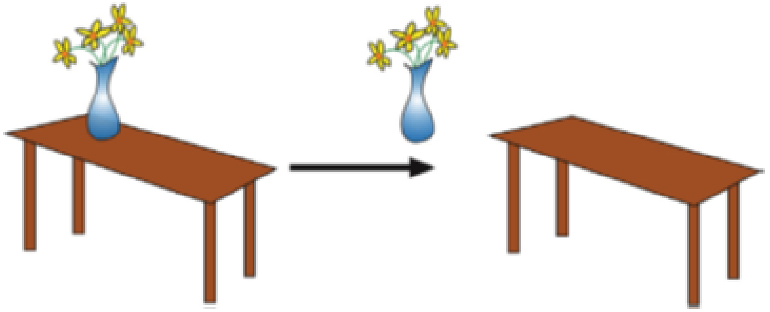
    - GPS

    - Gyroscopes, Accelerometers, and Compasses

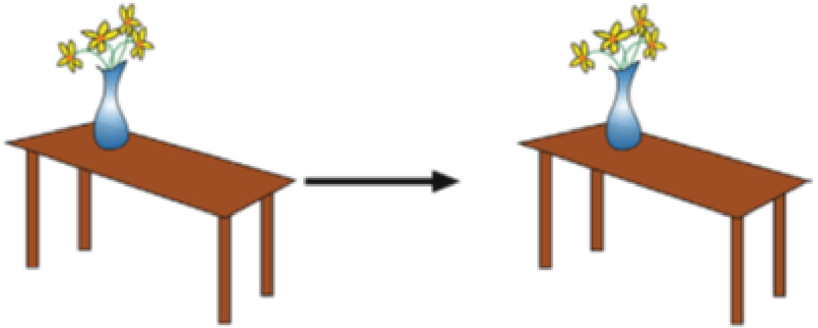
2. Sensors for gathering environmental information
3. Sensors for gathering user input

1. Tracking - **Camera**

- Many current augmented reality applications use techniques from *computer vision* to determine the participant's location and perspective with respect to the real world.
- The location and orientation may be
  - •*absolute* in nature (the location and orientation are very specific places regardless of the location and orientation of anything else (in other words, a specific place in the world)),
  - •*relative* (they are determined with respect to the location and orientation of something else).



Absolute location the vase remains in place even if the table were moved



Relative location is if the location is determined with respect to a table in the environment

- In order to enable computer vision, the sensor required is a camera. The camera “sees” the real world and, based on what it “sees,” can determine where it (the camera) is located and how it is oriented with respect to the scene.
- In order to carry out computer vision, software is needed to analyze the images collected by the camera in order to determine what the camera “sees.” Based on that information the software calculates where the camera *must* be in order to see that view.
- There must be cues in the environment that the camera can use as landmarks to aid in determining location and orientation with respect to those landmarks. The landmarks can be natural features in the environment or can be placed artificially into the environment.
- In order to make the computer vision problem simpler, many AR applications utilize landmarks that are placed artificially into the environment that are images that the application can easily recognize. Images used specifically for this purpose are called *fiducial markers*, and sometimes *fiducial symbols*.

- When the system “sees” a fiducial marker it can determine where the camera is with respect to that symbol and how it is oriented with respect to the symbol. It is apparent that either the symbol can move with respect to the camera and/or the camera can move with respect to the symbol. Regardless, the computer vision software must determine the relative location and orientation of the camera compared to the symbol
- In addition to providing information about the relative position of the camera with respect to the marker, markers can also provide additional information to the AR application. For example, a certain graphical object can be associated with a particular symbol. When this is done, that object is displayed from the point of view that it should be displayed from based on the relative position of the camera to the marker.
- Fiducial marker typically conveyed two information to the AR system.

    1. to designate what computer graphics object to display,
    2. what point of view should be displayed.

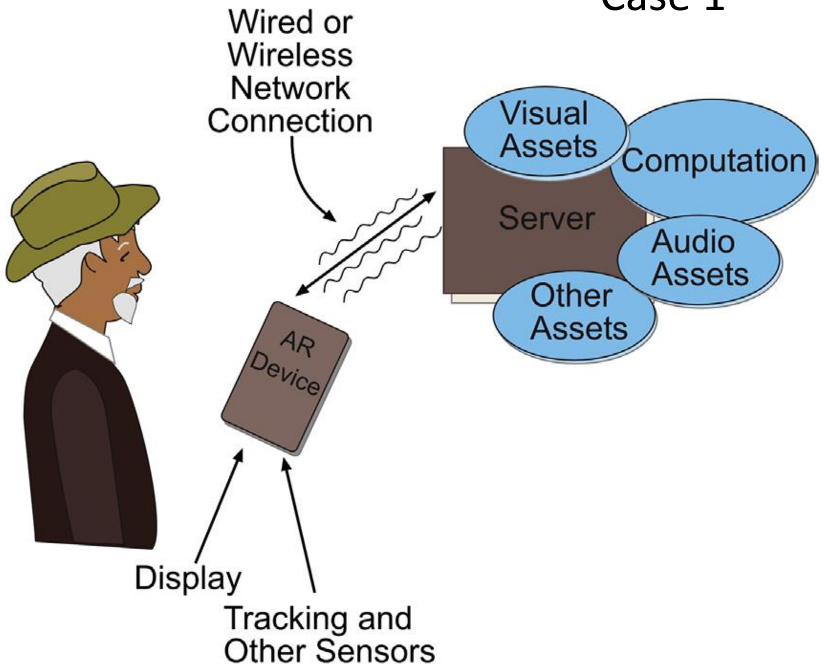
Fiducial markers

- Physical entities that are placed in a scene.
- more fiducial markers in the scene at any one place and time. This provides a solution if many different virtual entities need to be present simultaneously in a small amount of physical space that need to be able to be manipulated independently from each other
- may be static or mobile.
- Usage:

  Case 1: Encode information in the fiducial marker such as the URL of an object to be displayed. This way, the object would not be required to be loaded onto the AR system a priori and could be fetched from a network-based server when needed

  Case 2: Embed more information into the marker, people have utilized QR codes as fiducial markers. These codes can represent a great deal of information, are easy to recognize as markers, and are asymmetrical in nature

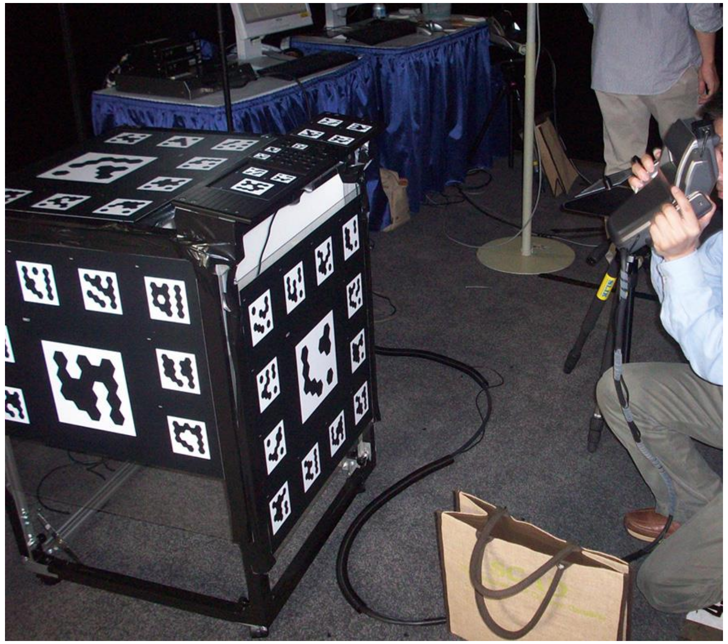
Case 1



A server-based AR system offloads some of the burden of an AR system to a server that may be located remotely from the participant and connected via a wired or wireless network

The participant is using a handheld device to interact with the system, but is taking advantage of the greater computational power and storage of a remote server.
The disadvantage is that the participant must be able to connect to a network to use the system

Case 2

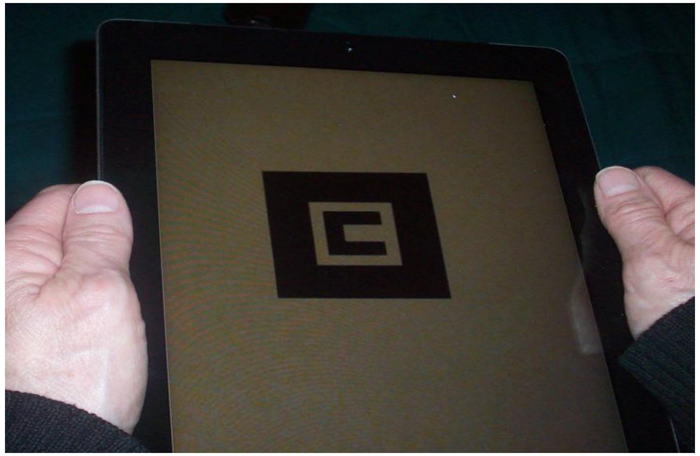


Multiple fiducial markers present in a single view of the vision system. Each marker unique from the others and able to be discerned from each other

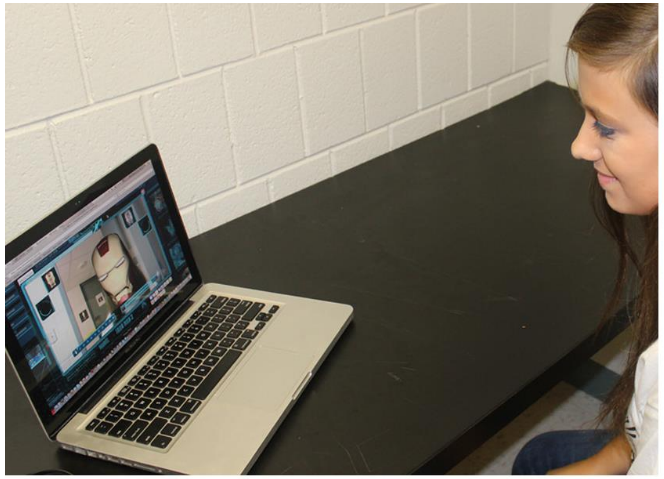
Case 3: Attached to physical objects—if the physical object moves, the marker moves with it

When an object with a fiducial marker is viewed through a smart tablet, one can see a representation of an object registered with that fiducial marker

Case 4: Fiducial markers are usually physical objects, such as pieces of paper with ink printed on them, but there is nothing to prevent them from being displayed electronically such as on a laptop screen, an iPad, or a smartphone. On these types of devices the markers can be created dynamically such that the marker itself can change based on the situation, or change over time, etc. This is different from the idea of changing what objects are associated with the marker, but both of these methods can achieve a similar result.



Case 5: Real-world objects as fiducials in augmented reality applications. Some natural and man-made objects provide the same types of traits as specifically made abstract markers.
For example, the human face can be recognized and because it is also asymmetrical, it is possible to use a person's face in the role of a fiducial marker. This has been done to great effect by using a face as a marker in an AR application that superimposes a virtual mask, a hat, or sunglasses onto a participant.

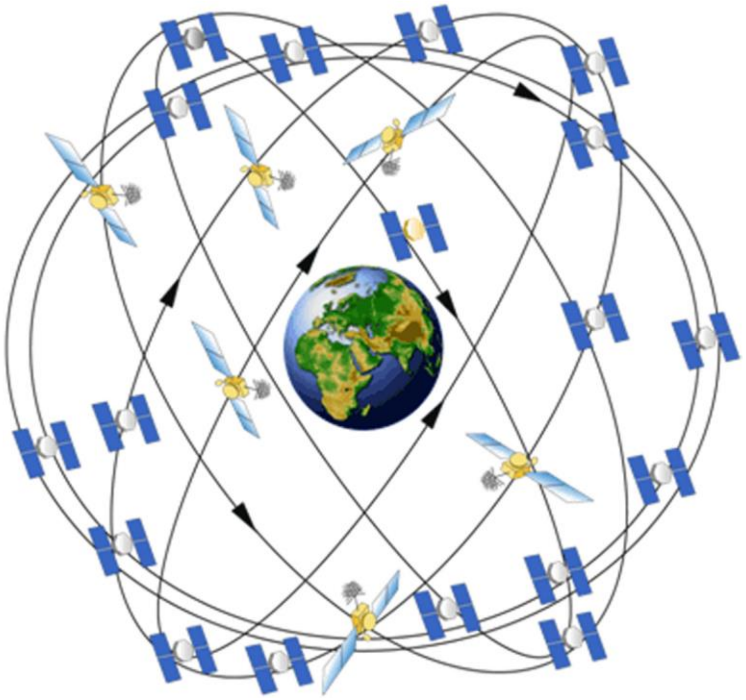


(Natural feature Tracking – NFT : The idea of using natural features to aid in visual tracking for augmented reality is developing rapidly)

Skylines are very recognizable and seldom change. Hence, they can potentially be used as fiducial markers.

Tracking - GPS

- GPS is a navigation system that utilizes a network of 24 satellites in outer space.
- The receiver can determine its location in *X* and *Y* if it can receive 3 satellites by measuring the amount of time it takes for the GPS signal to travel from the satellite to the receiver. By comparing the amount of time the signal takes from several different satellites, the location of the receiver can be computed to within several meters.
- If there are 4 or more satellite signals available, the altitude of the receiver can also be computed.
- AR systems can take advantage of the location information provided by a GPS receiver to gain information about its location in *X*, *Y*, and potentially *Z*.
- In general, a GPS is not that helpful in determining the orientation of the receiver in yaw, pitch, and roll. This limits the role of a GPS in AR systems in general, but the GPS information can still be used to great advantage—if it is available.
- GPS is not suitable as the only sensing mechanism for AR applications that require information about point of view and/or very high-resolution information about the location of the device



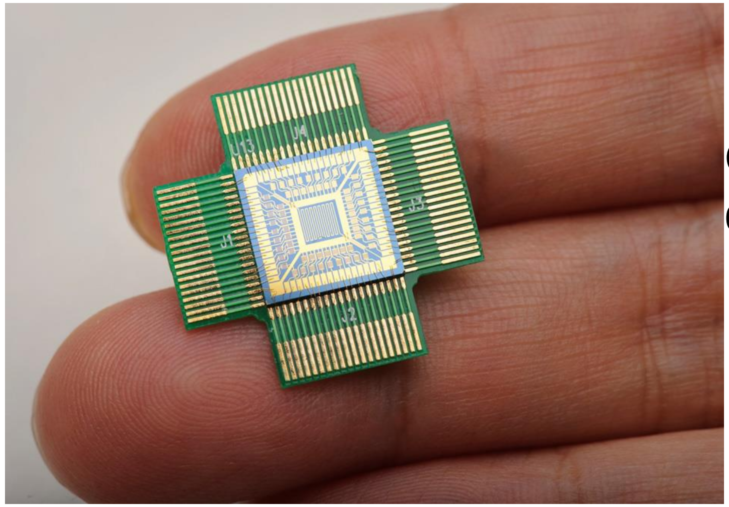
Constellation of 24 satellites makes it possible for a GPS receiver to determine its location on earth.

Tracking - **Gyroscopes, Accelerometers, and Compasses**

- The goal is to get information about the physical world and to use that information to inform the application. Some of the more common sensors that are used, particularly in mobile AR applications because many handheld devices are equipped with them, include gyroscopes, accelerometers, and compasses.
- ***Gyroscopes*** report values related to orientation, that is, they can provide information including yaw, pitch, and roll. Gyroscopes don't provide location information. They can be useful in measuring the orientation of a handheld computer or interaction device.
- ***Accelerometers*** do exactly what their name implies. They report acceleration. They can be used to determine the direction something is moving and changes in movement, and they do not depend on the presence of GPS signals. However, using accelerometers for a navigation system is hazardous because each number from the accelerometer is dependent on the number that came before it. Hence, any errors are amplified over time. Thus it is important to have some type of error correction or another sensor that can provide a reality check for the accelerometers. On the positive side, accelerometers are inexpensive and can provide data useful in AR systems
- ***Compasses*** provide information regarding what direction they are pointing. Just like an old-fashioned compass, an electronic compass can report whether you are pointed north, south, east, west, or anywhere in between.

2. Sensors for Gathering Environmental Information

- There are a number of other sensors that are used less commonly in AR systems but can provide information about the physical world at the time and place that the AR application is being used.
- For example, temperature sensors, humidity sensors, and other sensors can be used to overlay atmospheric information on a scene in an AR application. An AR application that uses these types of sensors could be used to make a handheld “magic lens” from a smartphone that makes it possible to “see” where the hot spots are on a surface that you are looking at.
- Other sensors can measure most any environmental condition, such as pH (how acidic or basic something is), voltage, radio frequency information, and many more specific aspects of the physical world



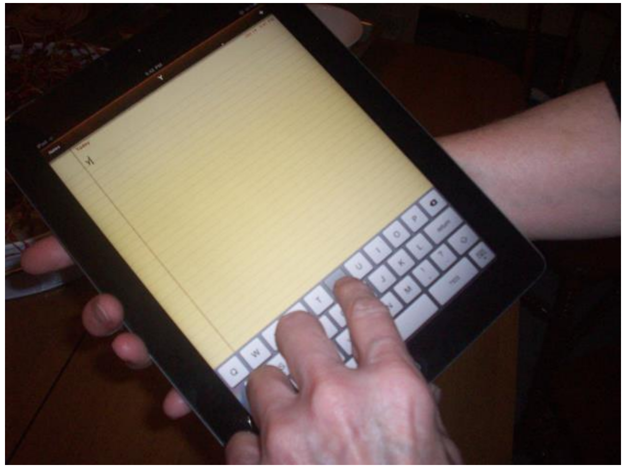
Chemical sensor from NASA designed for Cell phones

Types of sensor

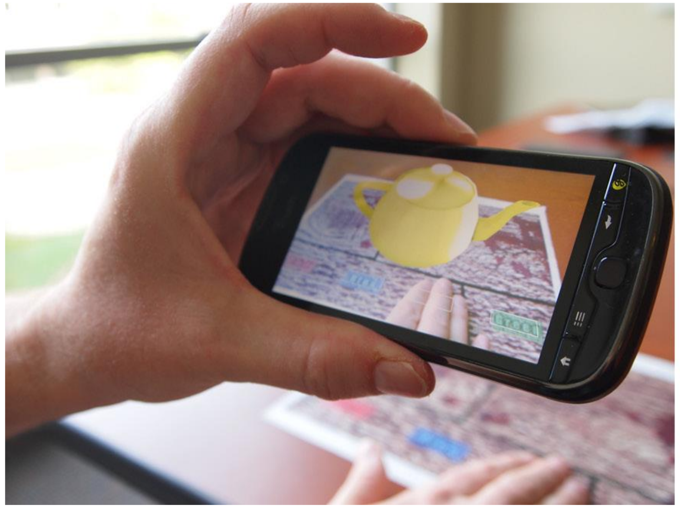
- Passive - sensors do their job without the participant consciously taking action with them. Certainly as the participant moves, the sensors report that move and update the display appropriately, but it is not a conscious action on the part of the participant to engage with a sensor actively.
- Active - sensors are useful when a participant is initiating the experience and at any point where the participant is allowed to make a conscious decision about how to proceed with the experience.

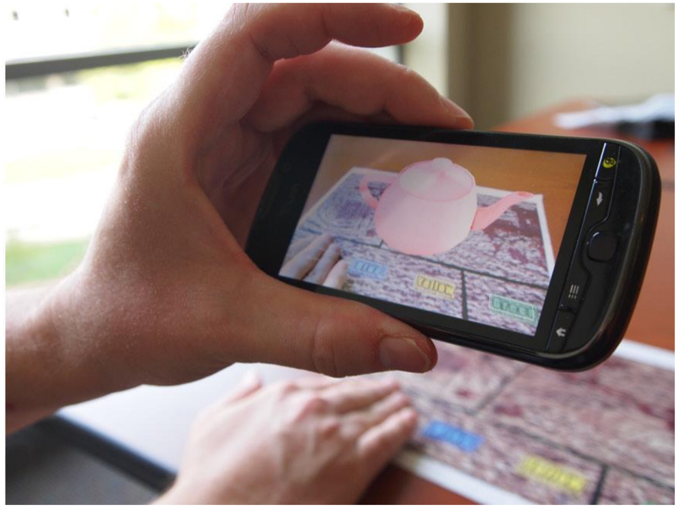
3. Sensors for Gathering User Input

- Sensors that gather user input would include buttons, touchscreens, keyboards, and other typical user interface devices.
- Most portable devices such as smartphones and tablets have a set of sensors in the form of buttons and keyboards (whether real or virtual) that an AR application designer can make use of to provide the participant with a way to directly interact with and control an application.
- These sensors are useful when a participant is initiating the experience and at any point where the participant is allowed to make a conscious decision about how to proceed with the experience. They are also useful in other ways that the participant might use—for tasks such as pointing to a place on a map on a touchscreen or enlarging the view on their tablet



- Cameras can also be used as a sensor to gather user input if a gesture recognition system is provided. In this case, the camera tracks the participant’s hands or fingers and interprets their movements as a direct input from the user.





Note : anything can be used as a “device” to provide input from the participant

II – Processor : Coordinates and analyzes sensor inputs, stores and retrieves data, carries out the tasks of the AR application program, and generates the appropriate signals to display

- Heart of any augmented reality system.
- Every augmented reality system includes a computer of some sort.
- Computing systems for augmented reality can range in complexity from simple handheld devices such as smartphones and tablets to laptops, desktop computers, and workstation class machines all the way through powerful distributed systems. In some cases, a handheld computer is in communication with a powerful server that might be located at a distance.
- **Requirement :**
- Enough computational ability to do the tasks it needs to do in *real time* (any action is made, such as a button press or a change in orientation of a handheld device or change in point of view, the system must respond with an updated display of the combination of the physical world and the virtual world with no apparent lags or hesitation)
- Latency

- Scene must be updated smoothly and at a rate that the participant in the experience perceives as a constant stream of information.
- To draw an analogy, a typical movie in the cinema is played at a rate of 24 frames per second, which is a sufficient frame rate for the viewer to perceive as smooth motion. If the frame rate were reduced to about 10 frames per second, the viewer would perceive the movie as a series of individual pictures rather than as a single motion picture.
- AR applications must sustain a frame rate of at least 15—preferably more— frames per second for the participant to perceive the display as continuous. Displays that are simulating the feel of a solid object must be updated about 1000 times per second or else the object will feel “mushy.”
- In augmented reality, it is particularly important that the rate of display of the virtual entities matches or is close to the rate that the physical world is displayed or the participant will perceive the virtual world lagging the physical world. In cases where the physical world is mediated by the computer, the displays can be synchronized, but in cases where the physical world is perceived directly, any lag in the creation and display of the virtual components becomes obvious.

III – Display Device - A display is the component that causes an appropriate signal to impinge on a participant's senses

Displays:

1. *Visual display (vision)* shows visual imagery to the participant. E.g. computer monitor
2. *Audio display (hear)* causes the sounds made by the system to be audible to the participant. E.g., headphones, loudspeaker
3. Olfactory signals (smell),
4. Gustation (taste), and
5. Haptics (touch).

Display Categories based on attached to the participant in some way or are not attached in any way to the participant:

- Computer monitor is not connected directly to the participant.
- Head-mounted display is worn on the head in the form of a helmet or glasses. The smartphone and smart tablet each provide a visual display that is held in the hand.

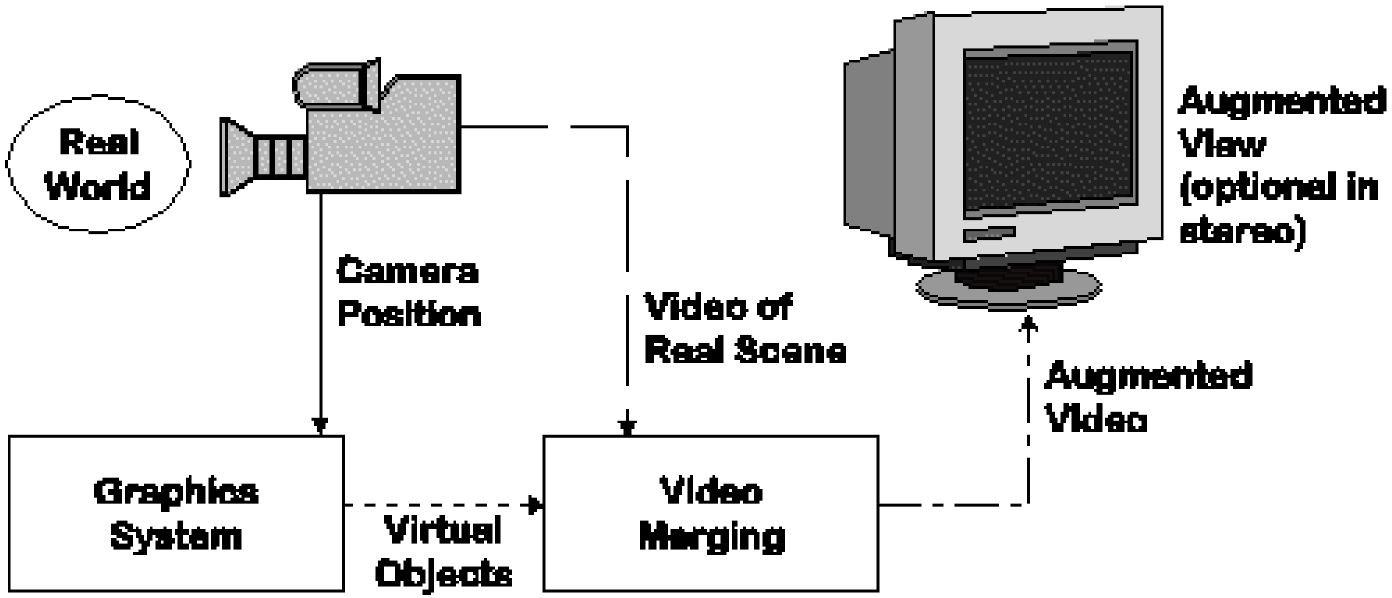
Display Devices

- Monitor Based
  - Laptops
  - Cell phones
  - Projectors

- Head Mounted Displays
  - Video see-through
  - Optical see-through

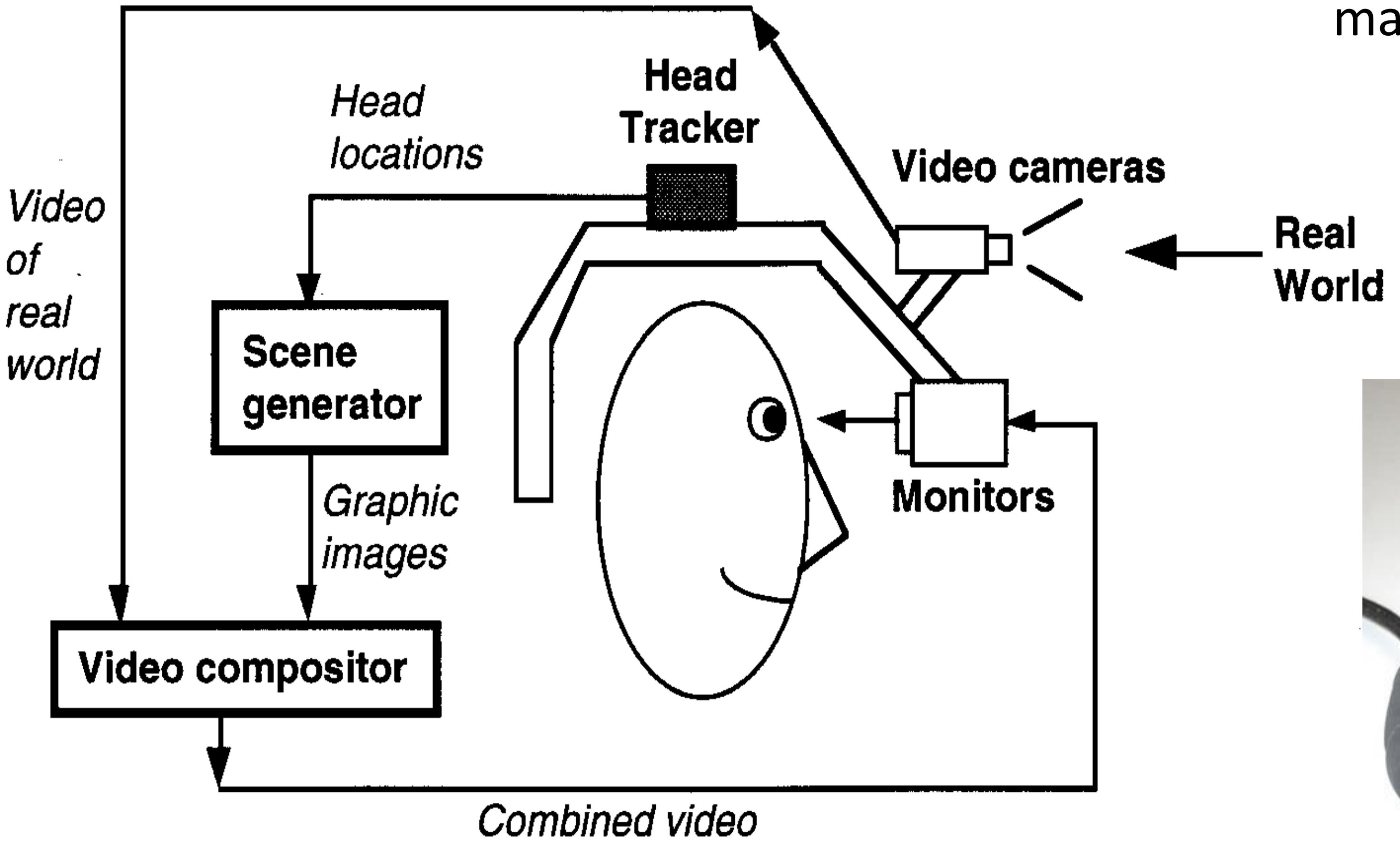
Monitor Based Augmented Reality

- Simplest available
- Desktop/ laptop/cell phone as a window through which see AR world.
- Sunglasses



Advantage:
Consumer-level equipment
Most practical
A lot of current research aimed it

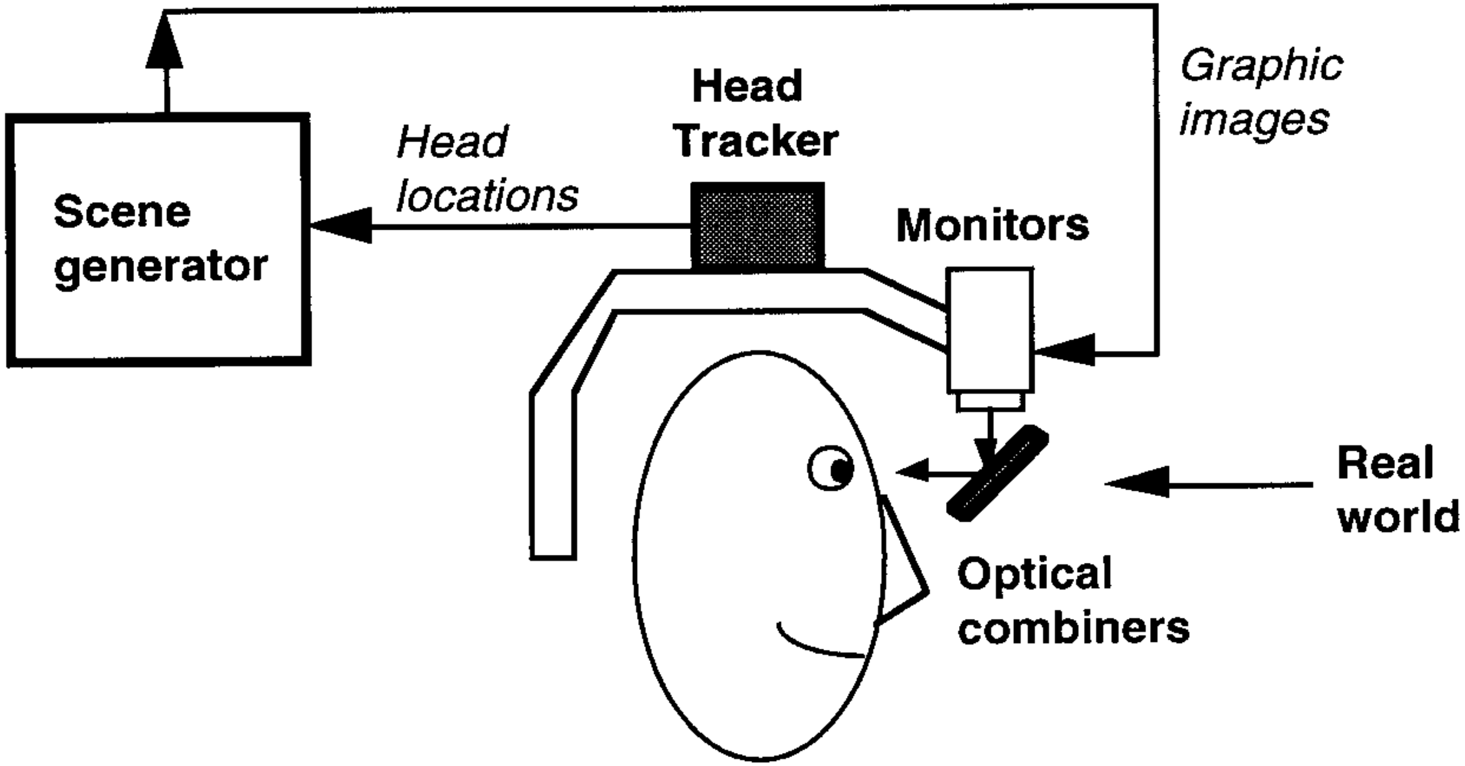
Video see-through HMD



Advantage:

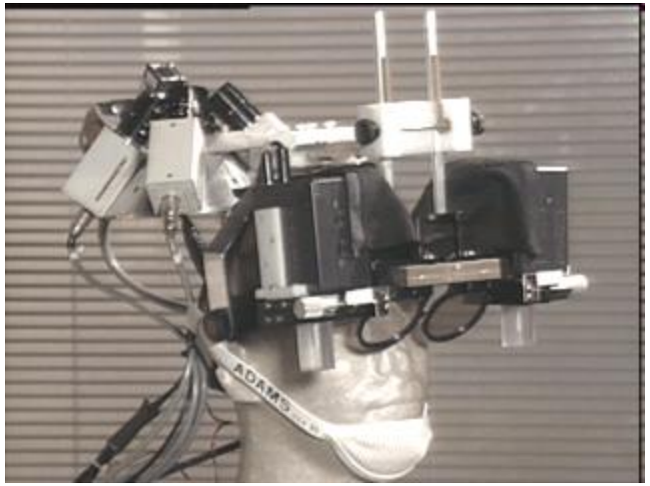
- Flexibility in composition strategies
- Real and virtual view delays can be matched

Optical see-through HMD



Advantage:

- Simple
- Resolution
- No eye offset



IV – Input Devices

- Input Devices - Buttons, keyboards, and other actuators.
- These are actually sensors in which the participant takes an active role in setting their values.
- Other sensors are more passive in nature, in that the participant doesn't take an active role in determining their values.

Hardware Components

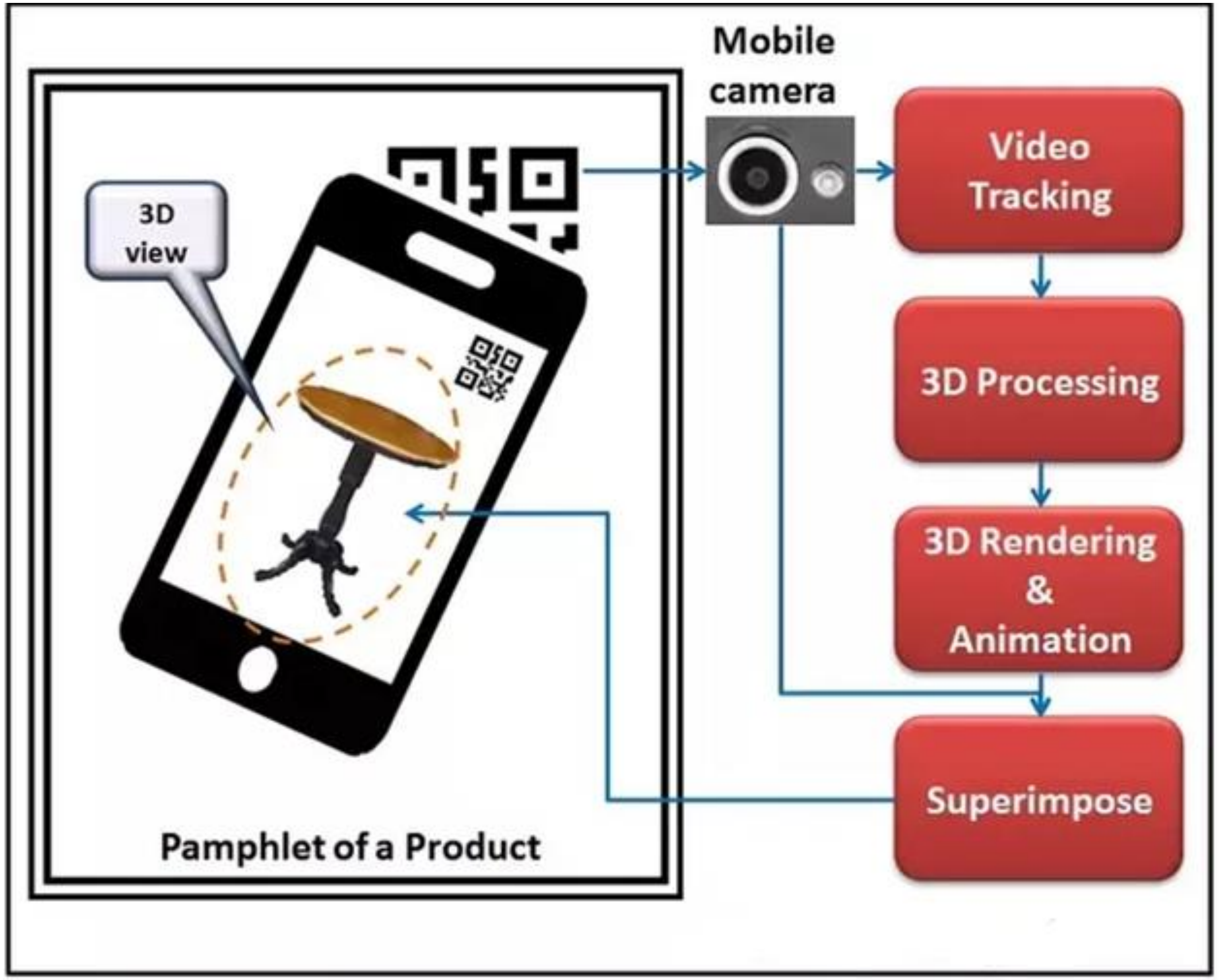
(Sensors, Processors, Displays)

1. Sensors

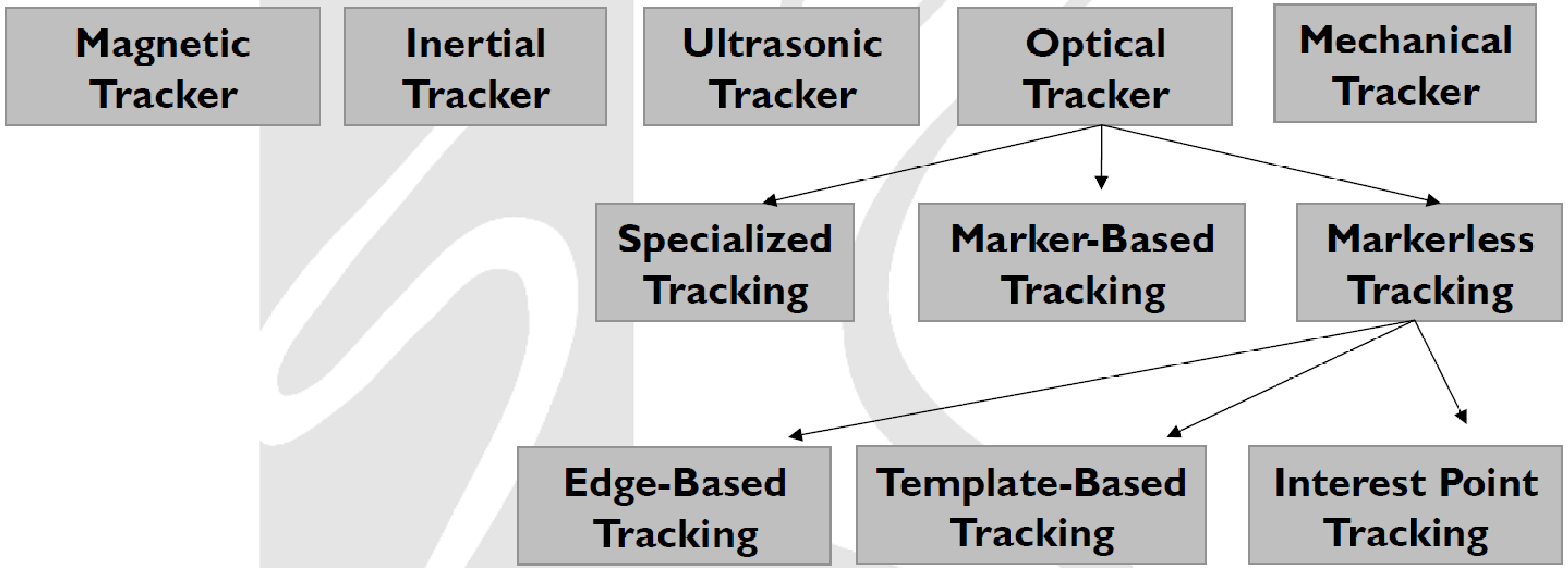
- Sensors provide information about the real world to the AR application.
    - Provide information about the location and orientation of the participant to the AR application. (Tracking)
    - Provide information such as temperature, pH, lightness / darkness, or any other types of information about the environment to the AR application (Gather environment information)
    - Gather user input

AR Tracking

Primary function of sensors is to provide information about the real world to enable the application to determine the location and orientation of different things in the real world



Tracking Types



Optical Tracking

- The camera gathers light through a lens and provides a signal that represents an image of what the camera “sees.” That image is then analyzed to determine the desired tracking information.
- The most common cameras used in AR systems are web cameras (webcams), smartphone and smart tablet cameras, and special-purpose cameras
- Most familiar cameras that operate in the visible light range, it is also possible, and sometimes desirable, to use cameras that operate in other frequency ranges, such as infrared or ultraviolet
- Optical tracking systems can be implemented
    - by placing cameras in the environment and having them “watch” the entities being tracked
    - attaching them to the tracked entities and “watching” the environment

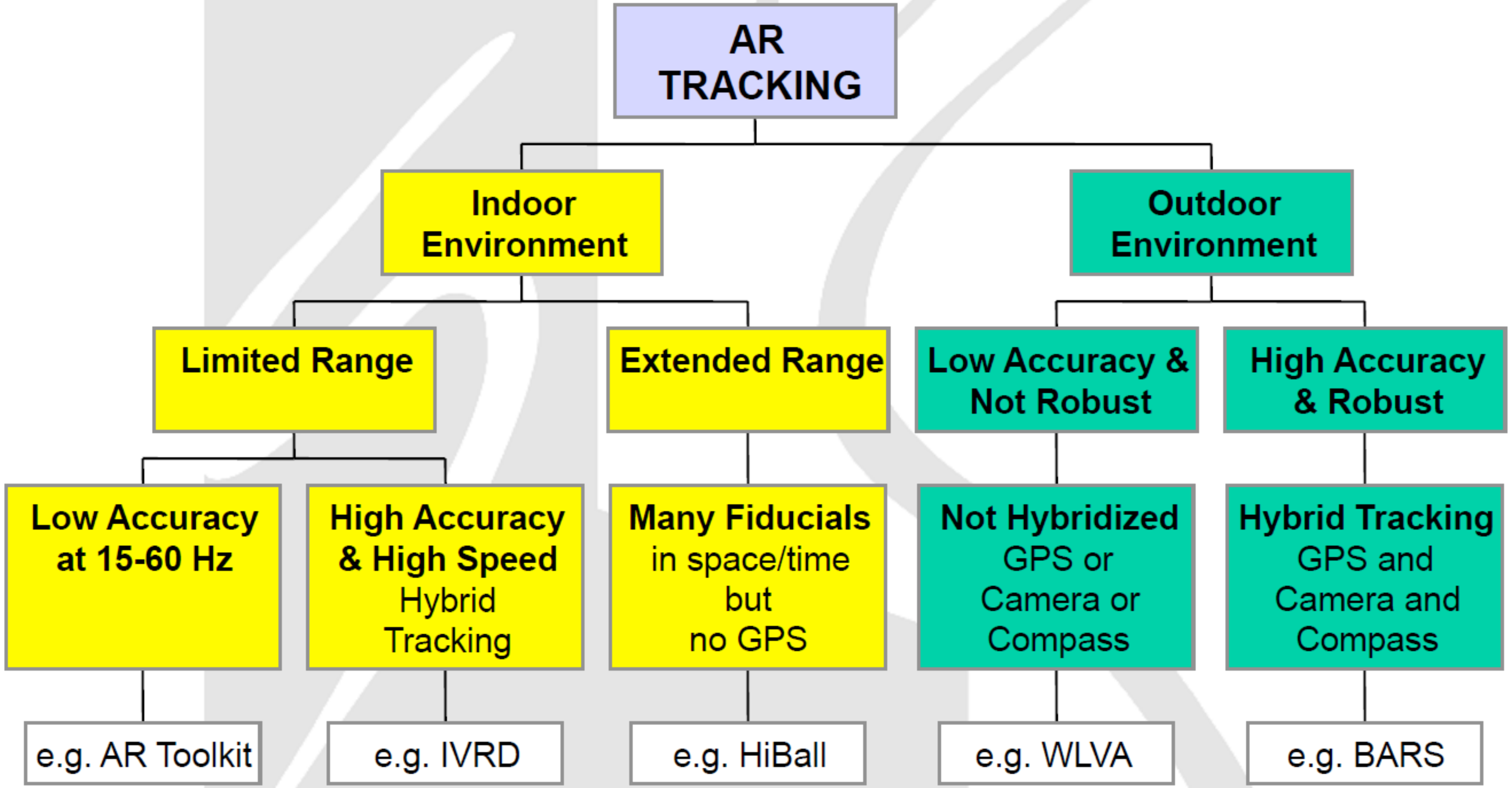
Infrared camera

Sensor(s) on the participant that watches the environment from the perspective of the participant.
Here, Single sensor on the participant's head. (There could be multiple cameras on the participant)

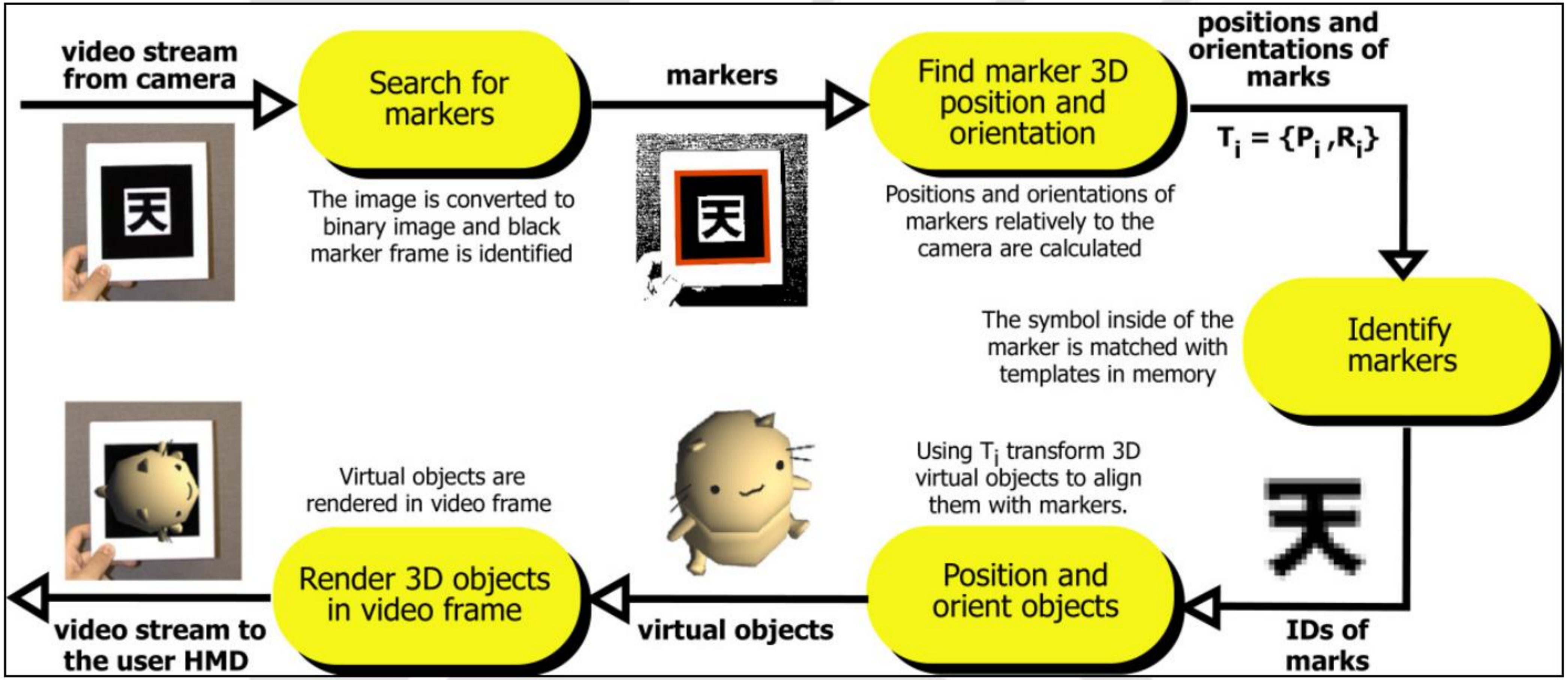
Cameras in the environment watching what is happening in the environment and watching the participant(s).
Here, multiple cameras in the environment (although there could be only one)

- Optical tracking is also used for motion capture systems & used to capture the performance of complex motions of multiple entities.
- This technique is typically used to achieve realistic motions for animated characters, but the same idea can be used for tracking multiple entities in an AR application.
- Typically, multiple cameras are placed in an environment and the character or objects being tracked are adorned with markers or other objects that are easy for the computer vision algorithms to identify

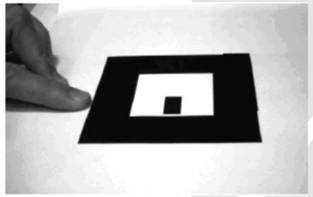
AR Tracking Taxonomy



Marker Based Tracking: ARToolKit

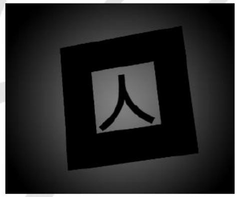


Tracking challenges in ARToolKit

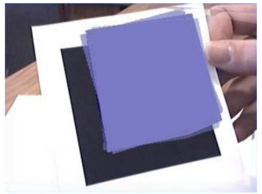


Occlusion
(image by M. Fiala)

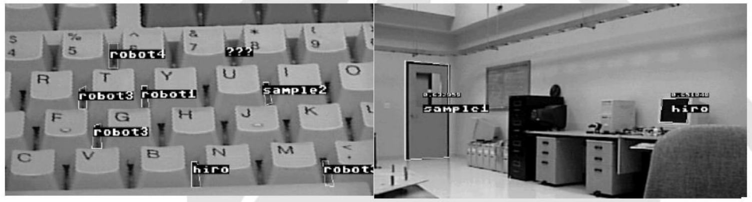
Unfocused camera, motion blur



Dark/unevenly lit scene, vignetting



Jittering
(Photoshop illustration)



False positives and inter-marker confusion

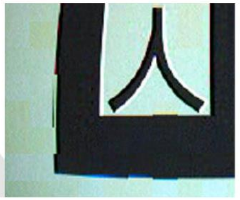


Image noise
(e.g. poor lens, block coding /

Limitations of ARToolKit

- Partial occlusions cause tracking failure
- Affected by lighting and shadows
- Tracking range depends on marker size
- Performance depends on number of markers
  - cf artTag, ARToolKitPlus
- Pose accuracy depends on distance to marker
- Pose accuracy depends on angle to marker

- # Advantages – Optical Tracking

- Doesn't require wires or anything else to be attached to the object being tracked. (There is no need to have a physical connection between the object being tracked and the real world)
- Allows numerous entities to be tracked simultaneously.
- Inexpensive (doesn't require any alterations to the real world other than, in some cases, adding some markers to the world for the system to watch)
- Cameras are already available in many devices used for augmented reality applications, such as smartphones

- # Disadvantages

- Enough ambient light to "see" the real world in enough detail to provide suitable images to analyze by the computer vision software
- Needs clear "line of sight" between the camera and the real world and the entities being tracked. (If there are multiple objects being tracked that are in motion, if one occludes the other from view of the tracker it can create problems tracking them individually. In some cases, cameras that operate outside the visible light spectrum can be used to advantage in applications that require optical tracking in low-light environments)
- Environment is suitable for tracking (If the environment is completely the same color with no further information available, the system will not have adequate information available to determine anything about the location and orientation of anything in that environment) - Fiducial symbols can be introduced in the environment to overcome this shortcoming. Of course always possible or desirable to place fiducial symbols in the environment that the AR application will be in.
- Latency into the system. (It takes time to acquire an image, deliver that image to the processor, and then analyze that image to determine the desired information)

Acoustical Tracking

- *Microphones are used a*s sensors in acoustical tracking systems.
- Microphone(s) can be attached to the object being tracked or can be placed in the environment. With acoustical tracking systems, there needs to be a source of acoustical information to be sensed by the microphones.
- Generally, *ultrasound* is used. Ultrasound is sound that is higher in frequency than the human hearing system can perceive. Ultrasonic tracking systems work by having the object being tracked emit sound, and an array of microphones in the environment to capture that sound. Based on the timing and amplitude of sound sensed in each microphone, the location of the source of the sound can be computed. Different objects can emit different frequencies of sound in order to track multiple objects in the same space.

Ultrasonic sensor

- Advantage
- Not affected by lighting conditions (function in the dark, as well as in bright sunlight)
- Disadvantages
- Cannot be used in environments that are noisy with audio signals in the same frequency range that the system uses.
- Each object to be tracked must have a sound source attached to it. (This means it can only be used in environments in which it is known a priori what objects will be tracked, and also that it is feasible to equip those objects with a sound source)
- Limited range

Electromagnetic Tracking

- Electromagnetic tracking systems are able to track six degrees of freedom.
- There is a transmitter with three orthogonal antennas and corresponding receiver attached to the entity that you wish to track (also has three orthogonal antennas). The sensor is the receiving unit.
- The transmitter emits a signal, sequentially through each of its antennas. The signal acquired by each of the antennas is then analyzed, and the level of signal reported by each antenna can be used to compute the location and orientation of the receiver

Transmitting unit of an electromagnetic tracking system.
The transmitter is in a fixed position, and the receiving units (small objects about the size of a die) are attached to the object being tracked.

- Advantages
- Very precise and accurate.
- Not dependent on ambient light levels (used in completely dark environments, very bright environments, or anything in between
- Disadvantages
- Sensitive to metal in the environment that they are in.
- Limited in range (receivers need to be within a few meters of the nearest transmitter)
- Costly. Lower cost electromagnetic trackers couple the receiving sensors (with the antennas) to a larger receiver by a wire overcome by using a wireless system to communicate with the receiver, but then there needs to be a (potentially) bulky battery pack with each sensor.

Mechanical Tracking

- Mechanical tracking operates by attaching linkages to the object that user wish to track. Those linkages have sensors at each of the joints that report the angle between the linkages. Often this is done by placing a variable resistor (potentiometer) at the joint and reading the voltage there. As the angle of the linkage changes, the amount of resistance in the potentiometer changes and a corresponding change in voltage (that you can measure) occurs. The voltage can then be used to determine the angle between linkages. This information, in combination with the angles between all other linkages in the system, can be used to compute the location and pose of the object

- Advantages
- Very fast and precise. (Used when the area of tracking is very compact and the need for accurate tracking)
- Integrate force feedback with an AR application (force feedback provides another level of “realness” to the participant)
- Disadvantages
- Work well in applications where the mechanical components can be hidden from view.
- Quite costly
- Limited range (need to be connected physically to the mechanical tracking system)

Depth Sensors

- Result from a depth sensor is a measurement of how far an object is from the depth sensor. This information can be used in conjunction with other tracking technologies to provide information about the location of an object.

- # Advantages

- Relatively inexpensive.
- Do not require anything to be attached to the objects being tracked. (There is no need to know ahead of time what objects need to be tracked)

- # Disadvantages

- Provide depth information only. (don’t provide location or orientation information directly. Hence, for use as a tracking system, they must be used along with data from other types of sensors. They also don’t report what object they are “seeing.” They merely report the distance they are from whatever object is nearest to them)
- Require a line of sight (Not enable to track multiple objects that occlude each other)

Multiple Sensors

Some tracking schemes require more than one type of sensor for tracking. Each sensor can contribute to the overall goal of tracking. For example, *accelerometers* can be used to obtain information about relative motion but don’t provide any information about an exact location.

Xbox Kinect uses a combination of a depth sensor, optical tracking, and an array of microphones (acoustical tracking) to achieve a remarkable system that tracks the actions of a participant with no encumbering devices (the participant is not required to wear, hold, or do anything in particular for the system to be able to determine their location and actions)

Microsoft Kinect uses multiple sensors

- ## Advantages

- Use each sensor to its best advantage.
- Combining sensors can create a solution to meet the needs of specific application. (For example, the combination of sensors in the Kinect allows the participant to engage with the game system without requiring the participant to wear or hold anything special.)

- ## Disadvantages

- Complicated when need to integrate them all to provide a unified signal to report the information required by the system.
- More than one sensor costs more than using a single sensor if everything else is equal.

Other Sensors

- Other sensors are sometimes used as part of a tracking system.

  E.g., *GPS receiver* - provides a gross level of location information.

  C*ompass* - determine basic information about the orientation of the system.

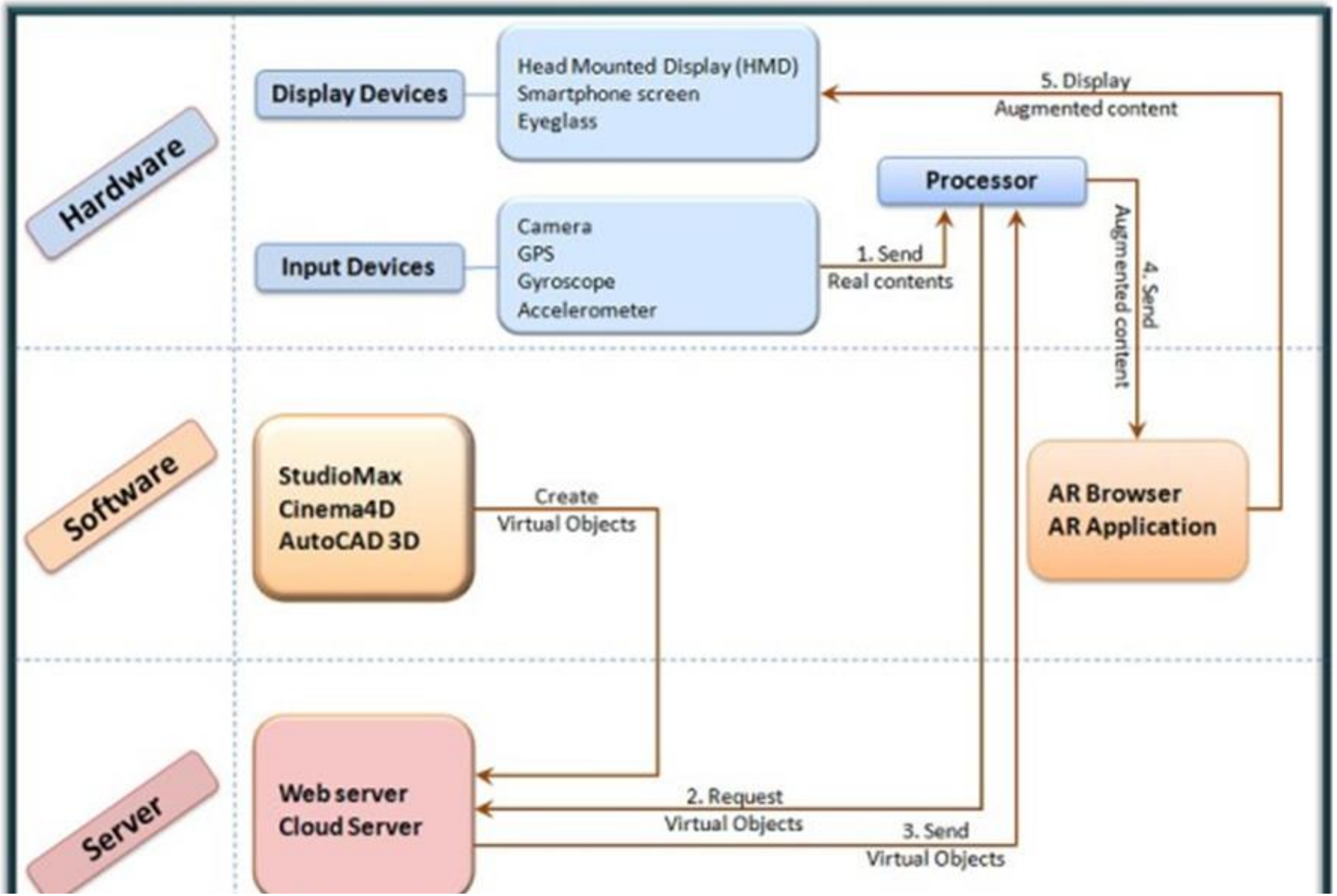
  G*yroscopes* - determine data regarding relative orientation such as leaning, turning, and twisting

- Many sensors are becoming fairly ubiquitously available.

E.g., *iPhone 4* - contains a multitouch screen, a camera, a microphone, an accelerometer, a gyroscope, and a depth sensor (typically used to disable the touchscreen when the phone is placed near the ear). Typical smartphone provides a number of sensors that can be used for tracking or other purposes in augmented reality applications.

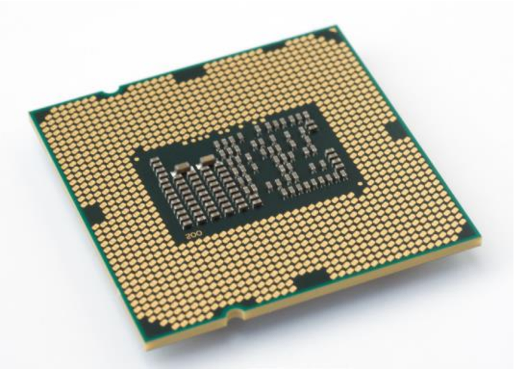
- Sensors are used to obtain information about the real world so that the AR application can use that information

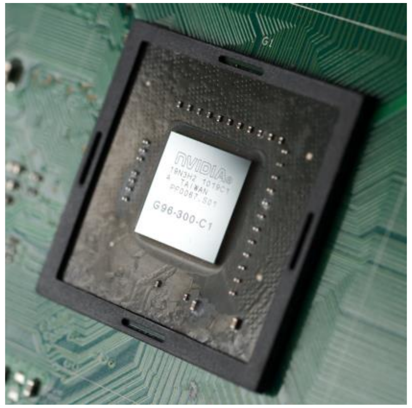
E.g., if an AR application needs to respond differently depending on the temperature in the real world at that place and time, a *temperature sensor* would be used.


Hardware
Display Devices
Head Mounted Display (HMD)
Smartphone screen
Eyeglass
5. Display
Augmented content
Processor
Input Devices
Camera
GPS
Gyroscope
Accelerometer
1. Send
Real contents
4. Send
Augmented content
Software
StudioMax
Cinema4D
AutoCAD 3D
Create
Virtual Objects
AR Browser
AR Application
Server
Web server
Cloud Server
2. Request
Virtual Objects
3. Send
Virtual Objects

2. Processors

- The core roles of the processor are to receive the signals from the sensors, execute the instructions from the application program based on the sensor information, and create the signals that drive the display(s) of the system.
- The processing system in AR applications consists primarily of
  - one or more general-purpose microprocessors as the central processing unit (CPU) and





  - one or more special-purpose graphics processing units (GPUs) - executing 3D graphics computations

Processor System Architectures

Application run on

1. handheld system such as a smartphone
2. handheld system connected to remote server(s)
3. desktop/laptop computer
4. desktop/laptop computer connected to remote server(s)
5. as a web application
6. a cloud with a thin client
7. Other combinations of local and remote systems

Handheld system

- Modern handheld devices - *smartphones and smart tablets (c*omputational power available.)
- Computational capability to execute the tasks required for basic tracking, interfacing with sensors, etc. But, they are not sufficiently powerful to compute complex simulations or other processor-intensive tasks.
- Every piece of software involved with the application, as well as all data (including graphical objects), must be resident on the handheld device.
- Entire application operates on a handheld device without requiring being connected to any network or other servers.

Limitations:

- Weak in the area of memory and storage
- Limit the number of virtual objects that can be displayed and the complexity of the objects, and limit the all-around utility of the system.
- Application needs to run in a remote location without a network

Handheld system connected to Server

- To handle data in handheld devices, it is helpful to couple the handheld device to a server (via a wired or wireless connection).
- With a server connected to the handheld device(s), limitations on numbers, size, and complexity of objects available via the handheld can be curtailed.
- Still be hard limits in terms of what can be resident on the handheld at any one time (i.e., having a server does not solve the problem if a single object is too large to handle on the handheld without serious trickery in the software environment), but by using clever caching and data swapping techniques, it is possible for the application to handle more objects than could be maintained directly on the handheld device.
- Typically the server would handle such tasks as content management, collecting logging data on the use of the application, and fiducial marker management.
- In general, this architecture is appropriate for applications where the portability of handheld devices is essential and where a network connection to the server is available

Laptop or Desktop Computer

- Modern laptop and desktop computers provide sufficient computational resources to be able to run typical augmented reality applications.
- The one case where such machines might prove inadequate is if the application program is especially computationally intensive. The most likely scenario for that is if the program must deliver results from high-fidelity simulations (perhaps they utilize complex physics or chemistry) that are typically run on supercomputers or other computers that are more powerful than typical laptop or desktop systems.
- Another scenario that might not be possible to execute on typical laptop or desktop systems would be if the program requires more memory than can be provided by the systems.
- For example, if the application requires a very large network graph to be constructed and analyzed, there may be insufficient resources available for such requirements. The downside to using a laptop or desktop system is its limited portability. Granted, laptops can be carried easily, but for applications that really require total mobility, they are cumbersome and awkward. This is not a limitation for all application scenarios though. In a case where there is a single point in space where the application is run, these types of systems can be fully adequate. For example, in a classroom, a museum, or an arcade where there is a kiosk where the application is run, it is perfectly reasonable and appropriate to use a laptop or desktop computer "behind the scenes."
- Used for basic for AR

Laptop or Desktop Computer with Server

- Some scenarios where it is advantageous for the laptop or desktop system to be connected to a server.
- One of those scenarios is where many graphical (or auditory, haptic, etc.) objects must be managed or where it is unknown a priori what objects will be required by the system. The server system can maintain a library of objects and software for managing those objects. A server system can also help overcome the limitations of laptop or desktop, such as if a computationally expensive simulation must be run. For example, a supercomputer could serve as a computational server to do the massive number crunching required by such simulations. The requirement for augmented reality applications to respond in real time, however, means that there must be virtually no latency in the time that it takes the supercomputer to respond. In order to be useful in this scenario, the supercomputer must be dedicated to the application (not running other programs) and must be connected by a high-speed, very low latency network.

Application Run as a Web Application

- Web app is a tool. The participant uses to interact with the application through web browser. The browser may or may not need a special plug-in to run the application.
- It is very convenient to be able to access the application from any computer that has a web browser (plus any required sensors). This scenario is the closest implementation to a “cross-platform” augmented reality application.

Limitations:

- In order to work effectively, the web browser needs to have some way of accessing the sensors that are either within or connected to the device.
- Requires the device to be connected to the World Wide Web. This may or may not be possible in some remote locations.
- Latency. (optimize the application for what tasks are handled on the server side vs what tasks are handled on the application side (client side)). Objects loaded from the server to the client as needed and new objects can be loaded (and no longer needed objects discarded) on the client.

Application Run on a Cloud with a Thin Client

- This scenario is virtually identical to running as a web application with the exception that instead of using a browser as the mode of interaction with the application, a special purpose application runs on the client side to interact with the application on the cloud.
- The client in this scenario is a lightweight "thin" client, (it is only the minimal software that is needed to connect to the cloud.) The bulk of the application runs on the cloud.
- Of course, latency is a serious consideration in this scenario. Most clouds are not designed to provide an option for consistent low-latency computing. It remains to be seen whether applications that require very low latency (such as AR applications) will be successful as a cloud service.
- Any client/server model is a trade-off between the capabilities offered by the remote system and the latency involved in using it. No matter how powerful the remote resources are, if the round-trip latency between client and server is too great, it will outweigh the advantage of the additional resources. Hence, a client-server AR application requires careful design, testing, and tuning for a given combination of processor, network, and server.

Other Combinations of Local and Remote Processors

- Of course, all of the afore mentioned architectures can be combined, mixed, and matched in many different ways.
- The important thing is to find the best possible combination for the particular application.

Choosing the most appropriate architecture for Application

- Largely an issue of optimizing a number of trade-offs in a way that best suit the needs. The *trade-offs* include decisions such as
  - portability vs power,
  - portability vs the need for network access,
  - cross-platform performance vs the need for network access.
- In general, the best solution where the need is true portability is a standalone application on a handheld device such as a smartphone or smart tablet. However, if the application requires more computation or memory than the device supports, it then becomes a question of whether it is better to use a larger system or to connect the portable device to a remote server via a network.
- There are many advantages to using a web-based application if the specific application can be delivered this way, and if there is a connection to the web available.

<u>Note:</u> Most AR applications are rather hardware specific. That is, they are developed for a specific platform. (Need for more platform independence for augmented reality.)

	Handheld	Handheld with Server	Desktop/ Laptop	Desktop/ Laptop with Server	Web Based	Cloud Based	Combination
Portable	Yes	Yes	No	No	No	No	Maybe
Portable with network	Yes	Yes	No	No	Yes	Yes	Maybe
Sufficient computation alone	Maybe	Yes	Yes	Yes	Yes	Yes	Maybe
Requires network	No	Yes	No	Yes	Yes	Yes	Maybe
Cross-platform solution	No	No	No	No	Maybe	Maybe	Maybe
Latency important	Maybe	Yes	No	Yes	Yes	Yes	Maybe

Processor Specifications

1. Number of processors - the basic processing required for carrying out "augmented reality" tasks, such as integrating inputs from sensors, executing computer vision algorithms (if used), and carrying out basic instructions from the AR application, can be handled by virtually any modern-day processor. The need for more processing capability comes if the AR application requires any heavy-duty computation, such as intense simulations or network analysis. In general, for most AR applications, there is not a need to concern oneself with the number of processors unless you are doing something particularly computationally complex

2. Processor speed - the processor to be as fast as is available at a reasonable cost. The need for higher speeds is usually only warranted when doing computationally complex tasks. Graphics computations take considerable processor speed. However, graphics computations are usually handled by the graphics processing unit.

3. Available memory - latency is detrimental to AR applications, it is best to have as much data stored in main memory as possible as opposed to being available only from slower disk storage. Memory is the primary concern with handheld devices such as smartphones and smart tablets. The limited memory on such devices limits the number and complexity of digital objects on the device. If there is need to have more data available in memory on a smartphone or tablet, then it becomes essential to have a scheme for swapping data in and out of main memory to and from a server. If using a GPU, the GPU needs its own memory available for graphics computations.

4. Available storage (disk space) - It is slower to access and retrieve data from a disk than it is to store and retrieve from memory. Consequently, it is important to decide what, if anything, can be relegated to disk storage. Often, with handheld devices, the disk storage is done on a remote server

5. Graphics accelerator(s) - Graphics accelerators are hardware that is optimized for doing the computations for 3D computer graphics.

- In desktop systems, the graphics accelerator is usually in the form of a *graphics card*, but sometimes the graphics accelerator is part of the motherboard. Modern graphics accelerators have phenomenal performance.
- Smartphones and smart tablets have them built in.
- In general, the graphics accelerator drives the display system(s). Thus, it is important that the accelerator supports the types of operations and functions needed. For example, some graphics accelerators can support multiple display devices. This can be handy if you need to have more display real estate available than a single display can accommodate. Thus, if you are using a computer monitor as your visual display, with an appropriate graphics card you can use two displays that act together as though they were one single larger display or you can use the displays in different configurations. For example, if you need to drive a CAVE-like display, you can drive each of the different “walls” from different displays.
- Graphics accelerator in the system is whether you plan to display stereoscopic imagery. In order to display stereoscopic imagery you need to compute different images for each of the participant’s eyes and then ensure that each image is (somehow) displayed to the correct eye. For certain schemes for stereoscopic imagery you need special capabilities in the graphics accelerator.

6. Network bandwidth - Network bandwidth is the rate at which data can pass in a network. It is expressed in bits per second. The network bandwidth is not as critical in AR systems as latency, but bandwidth becomes important if you need to pass large amounts of information, such as many high-complexity computer graphics models in a very short amount of time. This is especially important in scenarios where the AR application is loading data from a remote server.

7. Network latency - Latency is the bane of all AR applications. Latency is the delay between the time something *should* happen and when it actually does. Latency is present in all aspects of an AR application (because it takes time to read sensors, compute what needs to be done with those, and then to display the results), but in situations where a network is involved, network latency is typically the most significant source of problems. It takes time for information to pass over a network. Simply increasing the bandwidth of the network doesn't solve the latency problem. If there are more data than the bandwidth can support, then the bandwidth does need to be increased, but even transferring one bit of information over a high bandwidth network still has latency

3. Displays

- The display is the device that provides the signals that our senses perceive. Displays provide signals to our eyes, our ears, our sense of touch, and our nose, and perhaps provide a sensation of taste. Additionally, some displays provide stimuli designed to cause other sensations,
    - Visual displays
    - Audio displays
    - Haptic displays
    - Other sensory displays
    - Stereo displays (stereoscopic and stereophonic)

Visual displays

- ***Role of visual displays :*** to create signals of light that impinge our eyes that we perceive as visual imagery. Eg., desktop computer monitor.
- Visual displays used for AR applications include:
    - Stationary visual displays
    - Visual displays that move with the participant's head
    - Visual displays that move with the participant's hand or other parts of his or her body

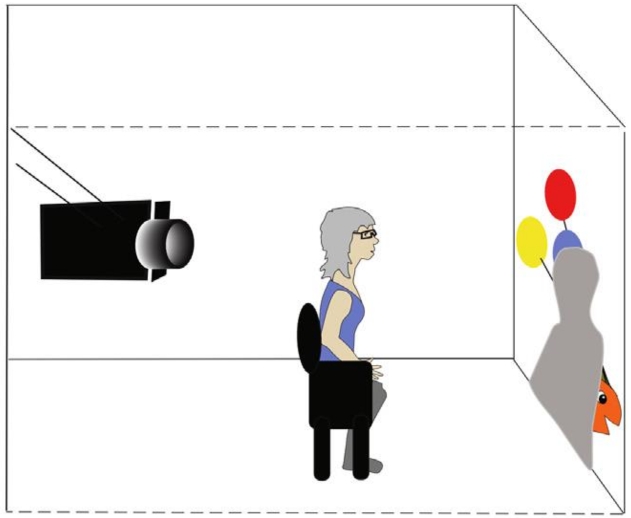
1.Stationary visual displays: Used in museums, schools, libraries, public buildings, etc

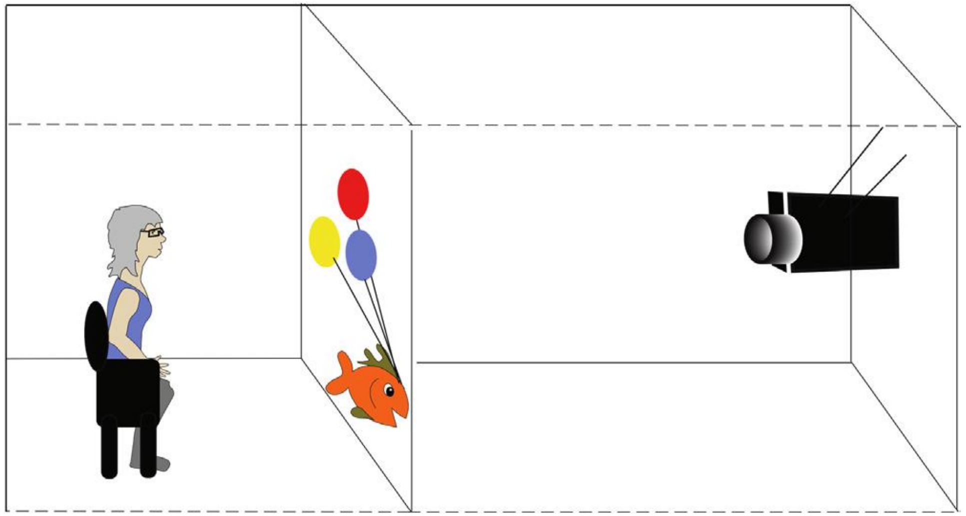
a) Kiosks are small, temporary booths placed in areas with high foot traffic that are used by businesses to reach their customers in a more simple and informal manner. Kiosks are primarily used for marketing purposes and can be staffed by individuals or self-service.

Uses a stationary AR display.
The participant holds up and can see an enhanced image of the box, as well as the box, on the display. The participant can manipulate the box to see the model from different perspectives.

b) Projection-based augmented reality - the participant sees light that is reflected on a surface.

Front projection -
the projector is facing
the front of the screen





Rear projection –
the projector is on the
opposite side of the screen

- Advantage of rear projection

  - Solve the participant occluding the image by standing between the projector and the screen.

- Disadvantage of rear projection

  - ❖ requires space behind the screen for the projector, and enough space to accommodate the distance the projector needs to have between itself and the screen (hence they are not typically portable).
  - ❖ more specialized item required
  - ❖ in the absence of a screen one can even front project directly on a wall or other surface. Not possible in rear
  - ❖ carefully consider the placement of other hardware, such as speakers, such that they don’t cast shadows on the screen

Projected AR systems may use

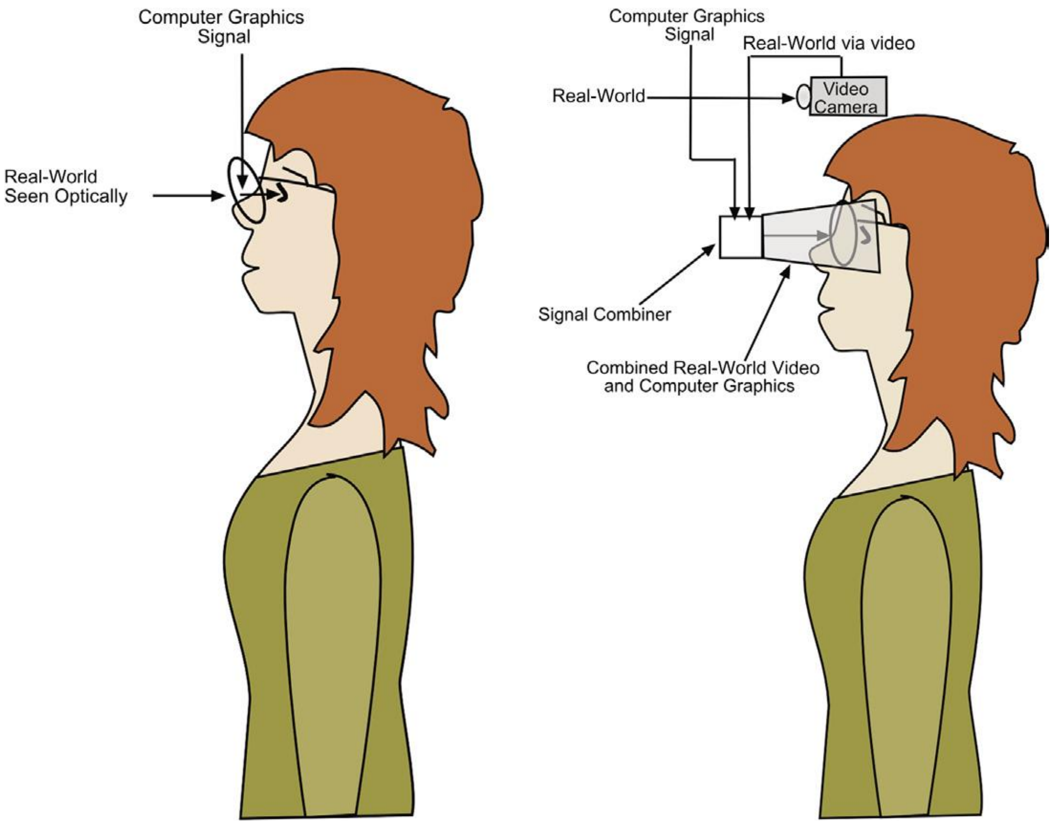
- one projector and one screen,
- multiple projectors on one screen (tiling the images together on the screen to achieve a single high-resolution image),
- multiple projectors with multiple screens (one or more projector per screen).
- no “screens” per se at all, but rather the imagery is displayed directly on the physical world.

Note:

- multiple projectors increase *brightness*, *resolution*, *apparent resolution*, *field of view*, and *field of regard*.

2. **Visual Displays That Move with the Participant's Head -** use displays that are mobile and move with the participant's head and are sometimes referred to as *head-mounted displays* (HMDs).

There are two basic types of see-through, head-mounted displays (or glasses). The first diagram illustrates the *optical see-through type*. In this type, the participant views the (real) world directly or perhaps through an optical lens. The computer graphics signal is then fed to the participant's eye(s) in conjunction with the optical signal, often by the use of a see-through mirror. The *video see-through* method illustrated in the second diagram generally precludes the participant from seeing the (real) world directly and instead provides a video camera (or potentially two video cameras for stereoscopic viewing) that is aligned as closely as possible with the participant's eye(s). This provides a video signal of the real world that can be combined with the video signal from the computer graphics system, which can then be displayed to the participant's eye(s).



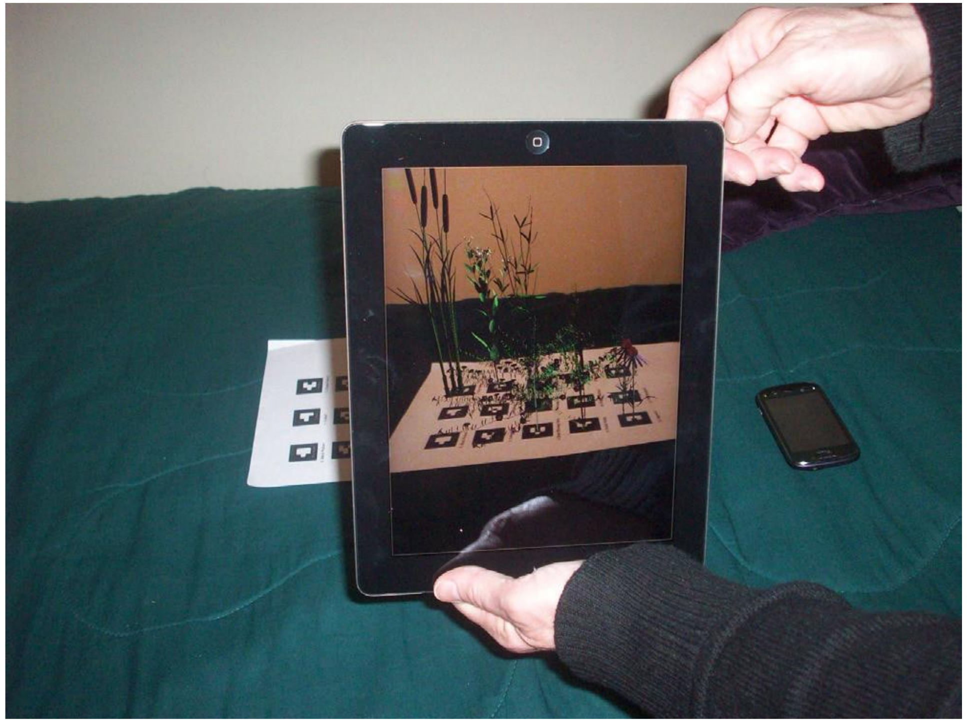
Head worn - types of mobile displays worn on the head take the form of a helmet or glasses

Eye worn - possible to have a visual display embedded into a contact lens (Future)

Stationary - display that moves with the participant's head, but the overall display is stationary.
E.g. , stationary binoculars

The view that a participant sees when looking through the digital binocular station "comes to life" in this museum application

3. **Visual Displays That Move with the Participant's Hand or Other Parts of His or Her Body -** smartphones (such as the iPhone and Android phones) and smart tablets (such as the iPad and similar devices)



Smartphones and smart tablets can be used as a "magic lens" to allow the participant to see things invisible without it.
Here, a sheet of fiducial markers shows a garden of different plants when viewed through an iPad.

- *Mobile / Portable projectors* - project onto surfaces in the real world that are present naturally (such as walls, the ground, and trees) but they can also project onto (real) objects that a participant provides specifically for the purpose of serving as a surface on which to project. For example, if the participant is wearing a projector on his or her head and is holding a surface in his or her hand covered with retroreflective material (retroreflective material reflects only in the direction that the light is coming from), the end result is that the participant can see a display on the surface (made of cardboard, plastic, or whatever), but other people can't see what is on that tablet.
- *Head-mounted projector* - displays onto a surface that is transparent, but in front of the participant's eyes. In this way, the participant sees the projected image but can also see through the surface and see the real world.
- *Projection-based systems* - projectors are usually mounted in the environment and focused to project on different entities

Audio displays - displays produce audio signals that our ears can sense

- Types:
  - Stationary audio displays
  - Audio displays that move with the participant’s head
  - Audio displays that move with the participant’s hand or other parts of his or her body

1. Stationary audio displays - Speakers

Speakers

Headphones, earbuds are examples of a mobile audio display system

Sound comes from some location. The dichotomy in AR is whether the sound appears to come from a fixed point in the world as the participant moves about or whether the sound appears to come from a fixed point with respect to the participant as he or she moves about. In general, stationary speakers are conducive to sound appearing to come from a fixed space with respect to the world, whereas with mobile displays, such as headphones, the natural mode is for sound to appear to come from a fixed space with respect to the participant. In both cases, both modes are possible, but to cross over to the less natural mode of display requires tracking the participant (which is already happening in AR systems) and then doing computations to cause the sound to seem to come from the desired location.

2. Audio displays that move with the participant's head- Headphones, Earbuds

- Headphones are worn *over* the ear(s), whereas earbuds are worn *inside* the ear(s)

3. Audio Displays That Move with the Participant's Hand or Other Parts of His or Her Body - Smartphones, smart tablets

- Smartphones, smart tablets, and other handheld devices often have an audio display built in. They often also have a jack available to allow you to plug in a pair of headphones or earbuds. The key aspect about handheld audio displays is that they can be either private or not private depending on the participant's wishes.

Haptic displays - sense of *touch*

Components:

- Skin sensations (taction) - include things such as temperature, texture, and pain
- Forces (kinesthetics) - sensed by how our body responds to them

Types:

1. Passive - Whenever you have real-world physical contact with something you are engaging with your haptic system. Digital binoculars provide information to your body when you use it.
2. Active - provide active display for both forces and skin sensations. There are readily available transducers that can heat and cool rapidly based on an electrical signal. Consequently, participants could wear sensors of this nature on the tips of their fingers, and when one of their fingers intersects a virtual object they could feel the temperature of that object

Note: To convey force feedback to participants

- AR application that makes use of force feedback to see an object (virtual) on a real table.
- A device
- PHANTOM Omni – used to “feel” the object

Move the pen around to probe into a space.

If the device was in registration with an AR visual display, you could see the virtual object on the table and feel it by touching it with the pen. You could gain a sense of the shape of the object in much the same way as you could gain information about the shape of a real object by probing it with a pen.

*__Note: S__*martphones have a vibrator built in. Hence, one could use the internal vibrator to indicate different things.
E.g. To indicate when the device has intersected with a virtual object. Thus, participants get an active haptic sensation regarding their activity in the virtual aspect of the world.

Other sensory displays

- Smell (Olfaction)
- Taste (gustation)
- Vestibular senses,

Note : Not fully explored in augmented reality .

Software

Software - AR

- The software involved with creating and using an augmented reality application is divided into four categories:
  - I) Software involved directly in the AR application
  - II) Software used to create the AR application
  - III) Software used to create the content for the AR application
  - Iv) Other software related to AR
- Another way to conceptualize the software components for augmented reality systems is:
  - Low-level programming libraries (e.g., tracking software)
  - Rendering and application building libraries
  - Plug-in software for existing applications
  - Standalone applications (e.g., content building, complete AR authoring)
  - Software to create the content for the AR application

- The key components include the following:
  - Tracking library (to recognize the fiducial markers or features)
  - Model loading and animation (to provide the AR elements of the game)
  - Game logic/game engine (i.e., something that provides the actual “game” of the application)
  - Rendering software to provide the images (visual, audio, etc.) to the display(s). Note that many game engines include rendering software as part of the game engine.

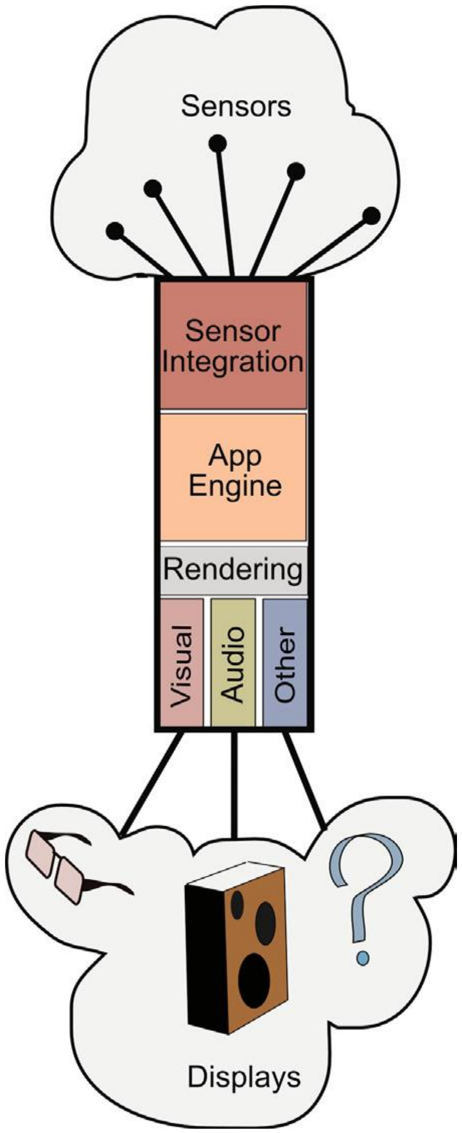
I) Software involved directly in the AR application

- The software involved directly in the AR application includes the following functional components:
  - Environmental acquisition (sensors)
  - Sensor integration
  - Application engine
  - Rendering software (visual, audio, etc.)

  These components are often bundled together in an AR library.

- openFrameworks (http://www.openframeworks.cc/), that make it easier for developers to combine different library modules, including AR library modules. openFrameworks, for example, includes AR tracking libraries such as ARToolkit

  https://www.lenskart.com/compare-looks



Relationship between the core components (hardware and software) of an augmented reality system.

The state of the environment is assessed by the sensors—the results of those are combined and fed to the main AR application. Based on the state of the total environment, the system renders the information as signals that are then fed to the various displays. This process repeats constantly, and the displays are updated anytime any piece of information changes. This might include the participant changing his or her position slightly or some other aspect of the environment or AR application changing

i) Environmental Acquisition

- The software that supports environmental acquisition - interfaces between the AR system and the sensors used to gather information about the state of the real world.
- For example, there is software required to take information from a camera and make it available as an image that the rest of the software can make use of. This is handled differently in different AR libraries that are available. It is important that the system is able to gather information from the camera very quickly.
- Likewise, in the case of gathering audio information from the real world through the use of a microphone, how well and how fast the transition is made from the analog signal from the microphone to the digital signals needed by the AR system can affect the overall quality of the AR experience.

ii) Sensor Integration

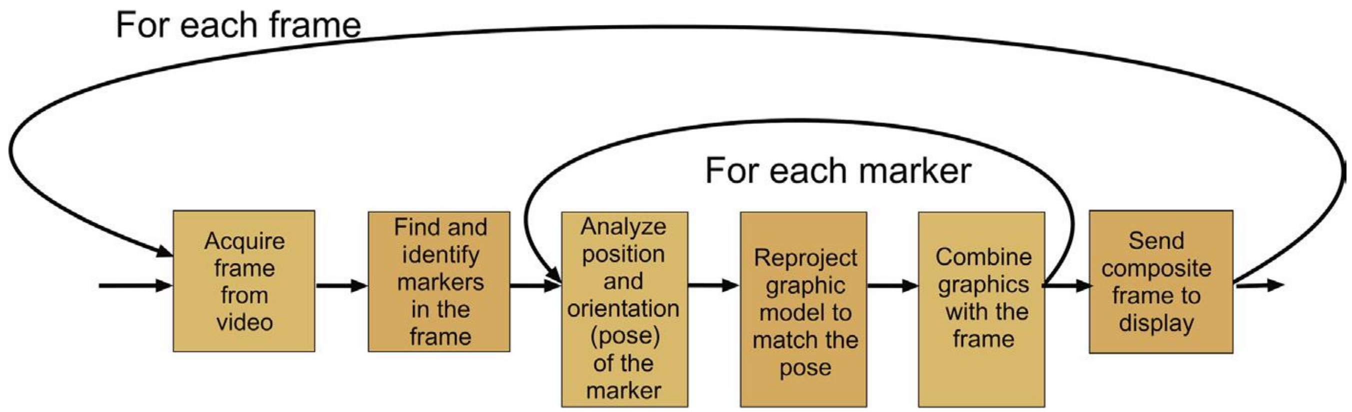
- The signals from the sensors are processed and integrated to make them suitable for the AR application.
- That is, the low level sensor signals are converted into a piece of information that is required for the AR application. The most obvious instance of this is the use of camera data to provide tracking information to the AR application. The camera images alone don't tell the application directly any information about where the camera must be in order to provide that image. The information must be (potentially) combined with another sensor, such as GPS, to tell the application where the camera must be; also, the imagery the camera is providing must be processed to determine the camera location and pose in the environment.
- During sensor integration : combine the signals in meaningful ways. For example, sometimes it is necessary to use the signals from multiple sensors to accomplish a single task. Sometimes tracking can be enhanced by combining the vision results with data from accelerometers, GPS, a compass, etc. The combination of multiple sensors can be advantageous to determining the camera location and pose accurately and speedily.

- once the camera image is acquired, the computer vision algorithm (software) “looks” through the image to find any fiducial markers and/or natural feature tracking (NFT) information.
- Provides information to the system about where the camera is with respect to that marker or other data.
- It determines this information through a complex set of processes that “look” at the size of the marker (to determine distance), how the marker is oriented with respect to the camera, distortions present in the image (it uses these distortions to determine the orientation of the marker with respect to the camera), and other entities in the scene and their relationship to the marker. This process must occur continuously (at least 15 times per second) because the application needs to “know” whether the camera has moved with respect to the marker. If so, it needs to recalculate where it is and how it is posed with respect to the marker.
- Note: in practice, the vision algorithm runs continuously whether there is any movement or not in the scene. The application updates at each time step whether anything has changed or not

iii) Application Engine

- The application engine is the core structure and framework for the AR application that the participants will interact with. The application engine gathers the inputs from the sensor integration component(s) and the participant(s) and generates the information that will be given to the renderer(s) to generate the signals for the display device(s). It also provides a simulation loop supporting user interaction.
- Component for content management - elements involved with the AR experience, or it might be a very involve component might be a very simple table that keeps track of the digital assets, such as objects, sounds, and other content section that keeps track of what assets are currently available on a mobile device

Note: In some AR applications, the role of the AR application engine could be handled by a computer game engine.



Basic flow of an augmented reality application

iv) Rendering :

Visual, audio , etc…

Augmented Reality Libraries

- Available commercially, as well as for free download from various sources.
- Each library brings with it a set of capabilities, and a philosophy for developing AR applications. Hence it is important to pick your AR library wisely, taking into account the capabilities the library offers, the needs of your application, the target deployment platform, the level of support available, the development methodology and cost of the library.

Cross-Platform Support

- AR libraries offer the current best hope for an application that can be deployed on different hardware and software platforms

II) Software Used to Create the AR Application

- Tools can be used to develop AR applications. There are, however, some environments that are especially conducive to developing AR applications.
- For example,

i) the free AndAR AR library (for Android devices) exists as a plug-in for the (free) Eclipse software development kit (SDK). Eclipse was written primarily in Java with plug-in support for other languages. Hence, it is friendly for development using the Java language.

ii) Apple's Xcode development (for Apple iOS) environment and need a connection to the Apple developer site. Applications developed in Xcode cannot run on Android.

iii) Develop for other platforms such as Sony PlayStations and Nintendo devices, you will need to be a licensed developer for them.

III) Software used to Create Content for AR Application

a) Software for creating and editing three-dimensional graphics

b) Software for creating and editing two-dimensional graphics

c) Software for creating and editing sound

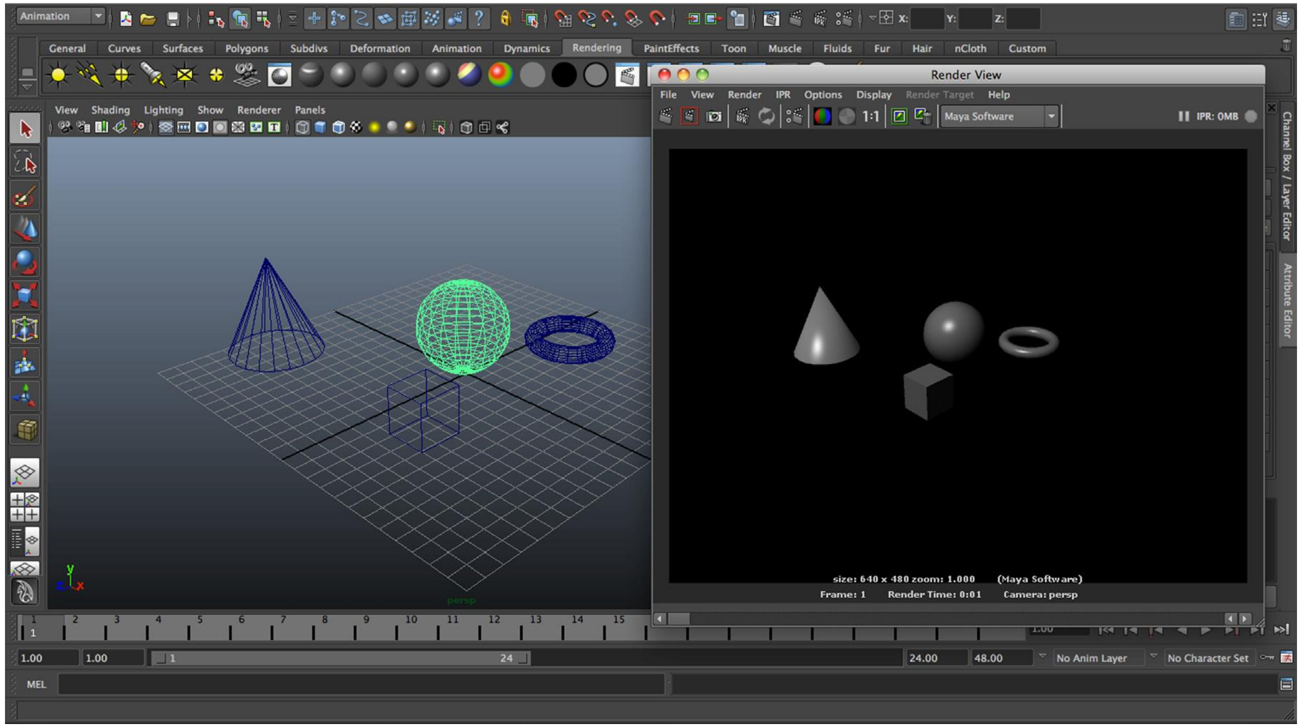
a) Software for creating and editing three-dimensional graphics

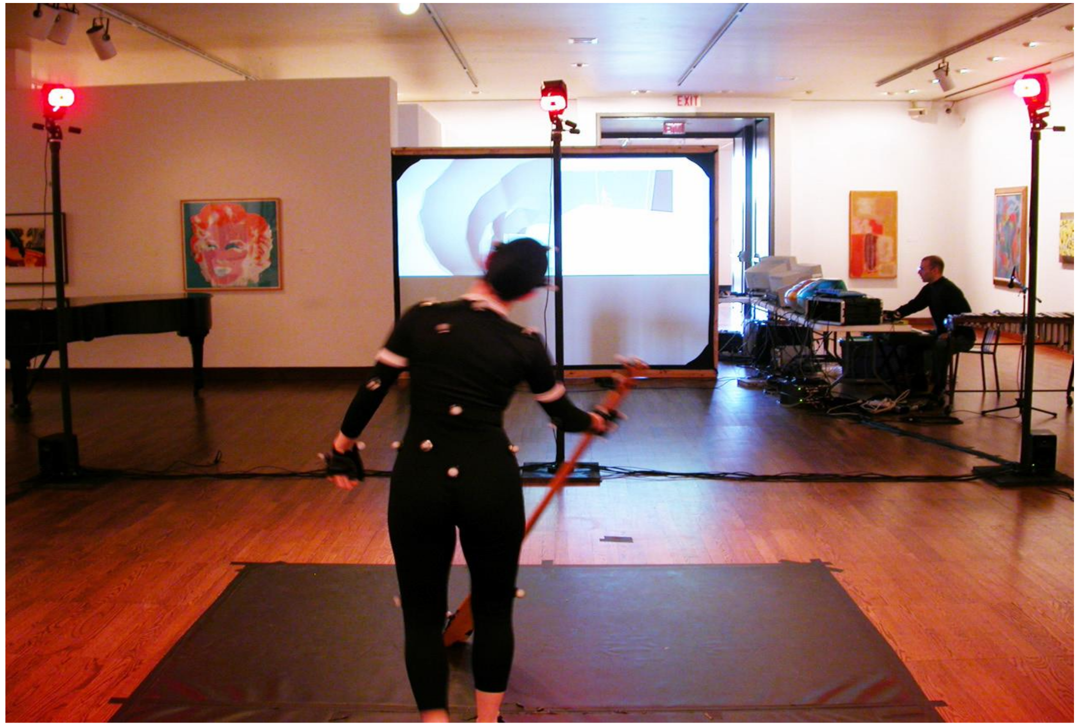
- Two basic categories of software for creating and editing 3D graphics.

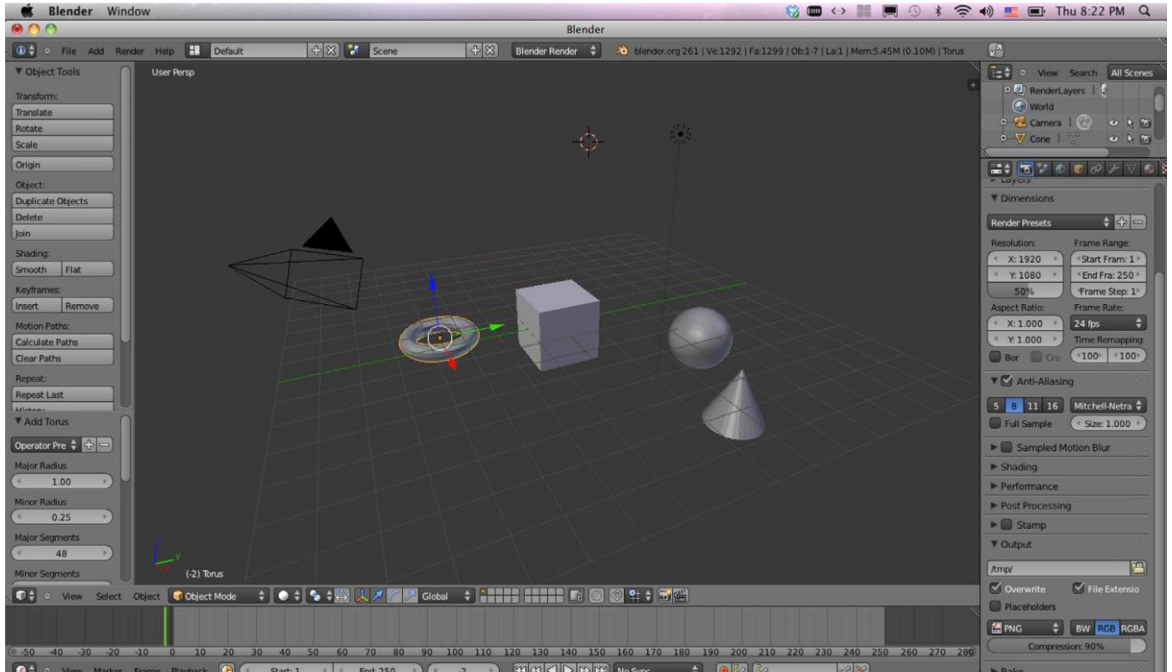
i) used to create computer graphics from scratch : - 3D modeling packages (Maya, 3d Max, SketchUp, Anim8or, Art of Illusion) used to create 3D computer graphics. These packages may be used to create augmented reality content as well. The main criterion is being able to generate the types of files that AR library requires. Because most of these types of packages can export and/or convert between many different graphics file formats.

ii) software that helps to "import" real-world objects through some type of a three dimensional scanning process or by analyzing photographs of 3D objects and creating graphical models of the object in the photos - popular free modeling tools is called Blender

- 3D scanner
- Motion capture systems work by "observing" a real-world entity and recording its motion over time, which can then be used to drive computer graphics objects in a similar way. One of the most common ways of doing motion capture is for an actor to wear a suit that has retroreflective markers on it at various critical locations on his or her body. Cameras then record the locations of those markers over time, and software turns those paths into a description of motion that can then be applied to computer objects. This is done most typically when there is the need to create complicated motions that are difficult to achieve through programming or directly via an animation software package


Render View
Maya Software
size: 640 x 480 zoom: 1.000 (Maya Software)
Frame: 1 Render Time: 0:01 Camera: persp


EXIT


Blender
User Persp
Object Tools
Dimensions
Anti-Aliasing
Output


VIVID 9i

b) Software for creating and editing two-dimensional graphics

- Types of software used for creating and modifying two-dimensional graphics:

i) vector graphics programs - used mostly when creating imagery from nothing. That is, vector graphics programs are a tool often used by illustrators.

ii) Raster graphics programs - allow to modify the pixels in a two-dimensional image. Hence, they are often used for retouching photos, for modifying images, and for processing color maps or anything that is done to all or part of an image. They are not typically used for creating an image from nothing.

C) Software for creating and editing sound

Primary ways that sounds are created for AR applications are as follows:

■ Record sounds from the real world that can be played back by the AR application at the appropriate time(s)

■ Precompute (synthesize) the sounds using software outside of the AR application for later playback in the AR application

■ Edit sounds from other sources using editing software

Sound is represented as a stream of numbers.

- A monophonic (single channel) sound - represented as a single stream of numbers
- A stereophonic sound (two channels) - requires two streams of numbers
- Multichannel sound - represented by the same number of streams of numbers as there are channels of sound

A digital recording is a direct translation from a real-world sound captured by a sensor (most commonly a microphone) and passed through an analog-to-digital converter to change the analog voltage produced by the sensor into a number, which is then stored in a list for later playback.

- General-purpose tools typically used for developing any other type of interactive applications and utilized when developing augmented reality applications. Some of these tools include:
  - Simulators and debuggers
  - Tools to manage multimedia assets (e.g., in a database)
  - For network applications, tools for creating, managing, and delivering web content
  - Packaging and license management technology, for both the application and the multimedia content

Steps for applying AR to a problem:

1. Identify the problem
2. Determine if there are other solutions to the problem
3. Determine the affordances of AR that will aid with the problem
4. Design AR application
5. Implement AR application
6. Test AR application
7. Evaluate results of AR application with respect to the problem
8. Modify design and application
9. Test modified application
10. Loop iteratively to appropriate step

Interaction

- The content, the interaction, the sensors, the displays, and all other components of AR application must work together to communicate the desired content to the participant, and the participant's actions must communicate to the AR application their intent and goals.
- Interaction can be defined as a mutual influence of one thing on another. That is, one entity does something, and the other entity responds in some way.
- Interactions can be between:
    - participant and AR application
    - participant and another participant(s) via the AR application
    - virtual world and real world
    - participant and virtual world
    - participant and real world

some of the actions in the real world

- We *press* a button.
- We *flip* a switch.
- We *speak* to others.
- We *listen* to a bird sing.
- We *kick* a ball.
- We *choose* what we want to take.
- We *select* between options.
- We *travel* to the store.
- We *issue* a command.

- Interactions in the virtual world can be boiled down to three primary categories:

1. Manipulation
2. Navigation
3. Communication

1. Manipulation - how we interact with things in the virtual world

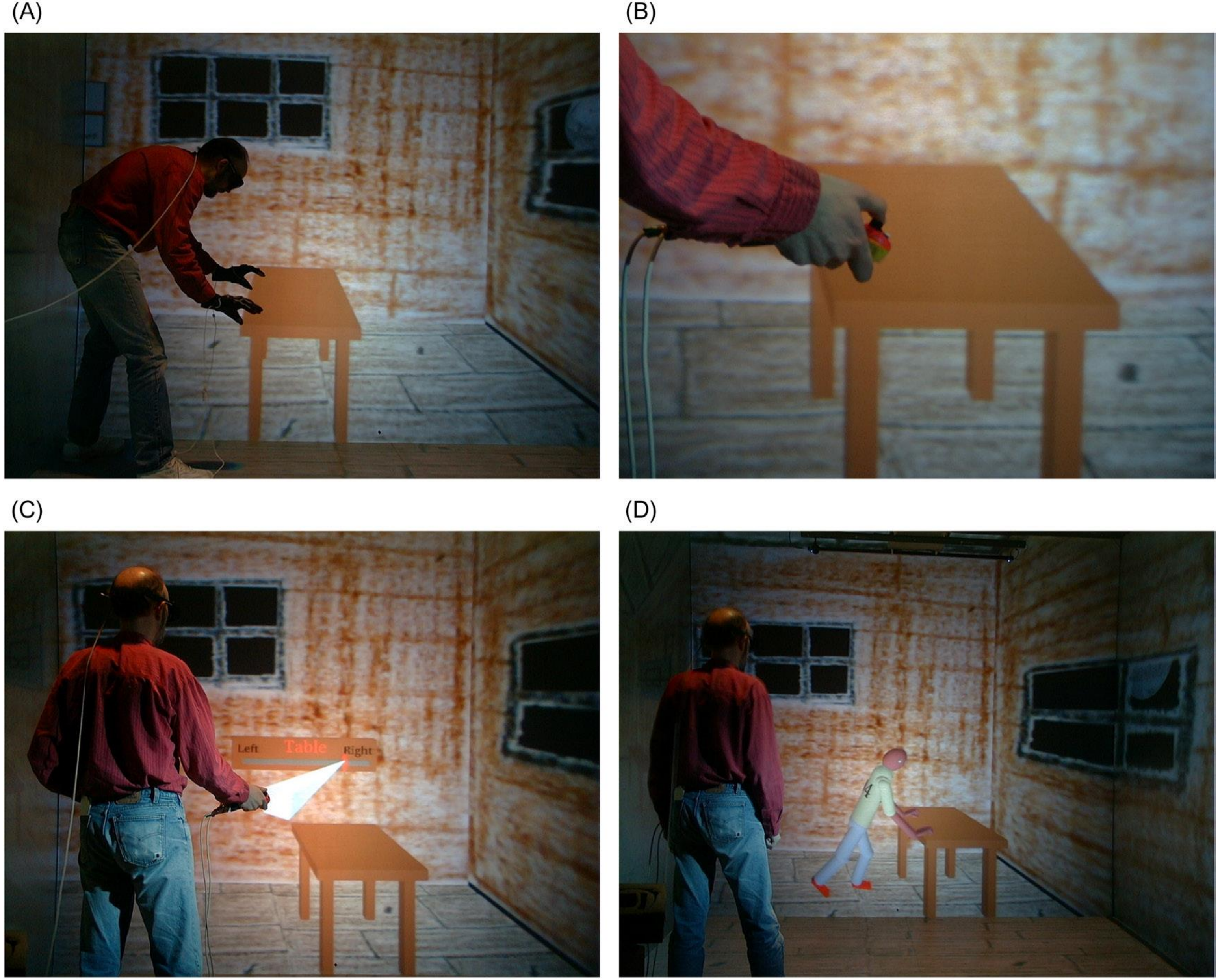
- Manipulate something (virtual) in an AR experience, make a
  - *selection* (i.e., identify what it is he wants to take action on) and
  - perform an *action*.
- In the real world, if we see three objects and want a specific one of them, we reach out and grab the object we want. If we are to emulate that in AR, if there are three physical objects, and they are within our reach, we have the option of doing the same thing. However, if the objects are virtual, we must somehow indicate to the system which object we are interested in and then must indicate we want to take it.
- Ways

1. Direct user control—in which participants manipulate the virtual world in a manner that is directly analogous to how they do it in the physical world. (VR)

2. Physical control—where participants use a physical device that they can hold and touch, such as a tablet computer or a physical push button. (VR)

3. Virtual control—where participants interact with virtual versions of a physical control, such as a (virtual) button, that they can press in the virtual world. (VR)

4. Agent control—where participants issue commands to some agent in the virtual world to carry out the action on behalf of the participants.



Different levels of physicality involved in different types of control systems in VR and AR systems.
(A) Direct user control, (B) physical control, (C) virtual control, and (D) agent control.

2.Navigation - how we move through the world

- Two components:

  (1) *Travel* - actual physical act of locomotion through the world

  (2) *Wayfinding* - related to how we know where we are in the world, and what we need to do to get to the place we want to go

- For example, if I want to go to the store, and my chosen method of locomotion is walking, I need to figure out where I am and the path that I am going to take to get to the store. As I walk toward the store I must adjust my course, make turns, and so on to cause myself to end at the store.
- Combined navigator and locomotor. If I am driving to a place that I don't know how to get to, a passenger in the car may serve as the *navigator* and read a map and tell me where to make turns, even though I am the person actually manipulating the steering wheel.

- *Methods that can be used as navigation aids in an AR application (or) to use AR as a navigation aid in the real world include:*
  - Provide a virtual map, including showing your current location on that map
  - Provide a virtual assistant that walks (or hovers) beside you to show you the way
  - Provide indicators of a path to follow, such as providing a (virtual) line or (virtual) footprints on the ground for you to follow
  - Issue spoken directions to you, such as “turn right at the next opportunity” or “head east for another 100 feet”
  - Provide (virtual) landmarks in the environment that can help you know where you are and how to find your way
  - Allow you to drop (virtual) “bread crumbs” that you can use to find your way back to where you came from
  - Provide you with (virtual) binoculars so that you can see long distances, and perhaps through objects to help you find your way

3. Communication (Multi-persons)

- The participants can communicate with each other in any of the typical ways that they can communicate in the real world. However, there are more ways that they can potentially communicate within an AR application. If real-world communication mechanisms are masked, then all communication is mediated through the AR application. Hence, if the software was good enough, the AR application could do real-time language translation between participants.
- If participants are too far apart to hear each other, there could be an audio channel in the AR application to allow the participants to hear each other, but one could also do more “AR-ish” things such as display large font text over each of the participants so that everyone can see what they are saying.
- Each person might see a map or be able to summon a map that shows the location of the other participants. Each person is tracked, thus making it possible to display their location on a map or otherwise.
- *Note:* One might be able to see the other participants in a totally realistic way, with the magic of AR, each participant could potentially see an altered representation of the others in the AR experience

Basic points of view of participants in augmented reality applications

- **Subjective Point of View -** most commonly used point of view is the one in which participants see the real and augmented world as though they are seeing it from their own eyes. Also termed as *subjective view augmented reality or first person augmented reality*.
- Participants don’t see whatever it is they are interacting with directly in front of them as they would see it from the first person
- **Objective Point of View -** participants see themselves in the environment and interacting with the environment. Also terms as objective view AR or *second person* AR.
- Looking at a projected (or some other type of large monitor) display and seeing themselves interacting with the virtual world

Implemented

Subjective view applications tend to be implemented with head-mounted or handheld displays, whereas objective view applications tend to be implemented with projection or some type of stationary display, such as a computer monitor or large screen.

The image shown here is actually of the display screen on which the participants see themselves in the augmented world. In order to see the dinosaurs, the participants must look at the screen. Participants stood near a fiducial marker and could see themselves along with the augmentations in their world

Displayed on the large screen. The participants must look at the screen to see themselves interacting in the augmented world

Views in augmented reality

Mobile AR

Mobile AR - hardware required to implement an AR application is something that you take with you wherever you go

- . Types of AR

 i) Portable augmented reality - uses technology that you can move from place to place. A desk-side computer with a monitor is somewhat portable in that it can be moved from one place to another relatively easily. A laptop computer is even more portable.

 ii) Mobile augmented reality - Device fits in pocket and is easy to operate wherever you are, even if you are walking or otherwise engaged. E.g., Smart phone and tablet

- *Note: Handheld gaming consoles and e-readers are easy to carry around. They may or may not provide the technological support for AR*

Different architectures - AR application

1. Application run on handheld system such as smartphone
2. Application run on handheld system connected to remote server(s)
3. ~~Application run on desktop/laptop computer~~
4. ~~Application run on desktop/laptop computer connected to remote server(s)~~
5. Application run as a web application
6. Application run on a cloud with a thin client
7. Other combinations of local and remote systems

Creating on AR application

Create AR App - Demo

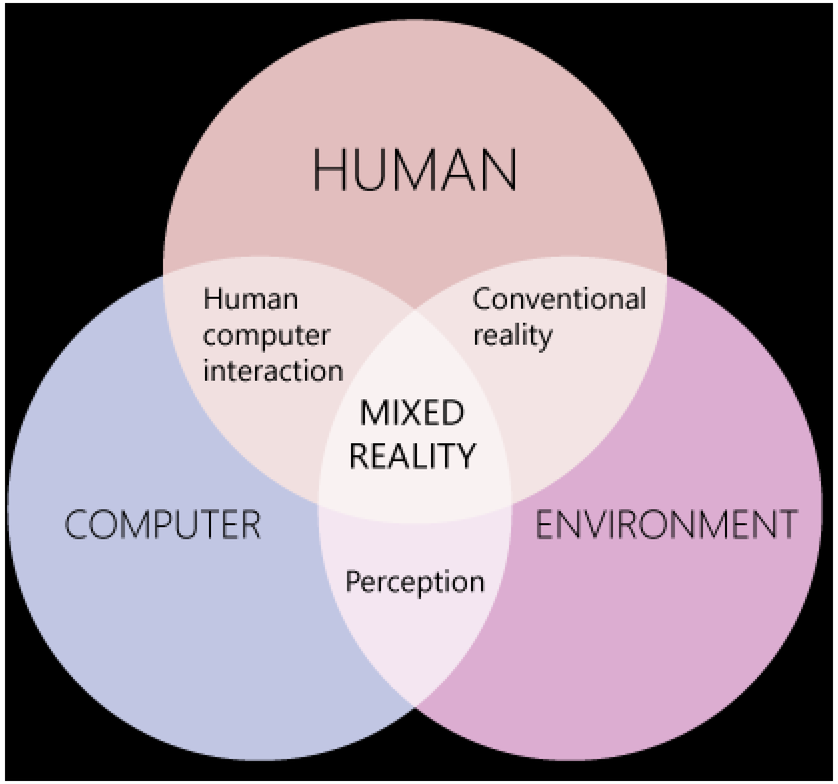
- https://www.youtube.com/watch?v=MtiUx_szKbI
- https://www.youtube.com/watch?v=l8abAMxF5iE
- https://www.youtube.com/watch?v=Ml2UakwRxjk

Mixed Reality

- Mixed Reality is a blend of physical and digital worlds, unlocking natural and intuitive 3D human, computer, and environment interactions.
- This new reality is based on advancements in computer vision, graphical processing, display technologies, input systems, and cloud computing.

Relationship

- The relationship between humans and computers has continued to evolve by the means of input methods. A new discipline emerged, which is known as human-computer interaction or HCI. Human input can now include keyboards, mice, touch, ink, voice, and Kinect skeletal tracking.
- Advancements in sensors and processing power are creating new computer perceptions of environments based on advanced input methods.



A combination of the three essential elements sets the stage for creating true Mixed Reality experiences:

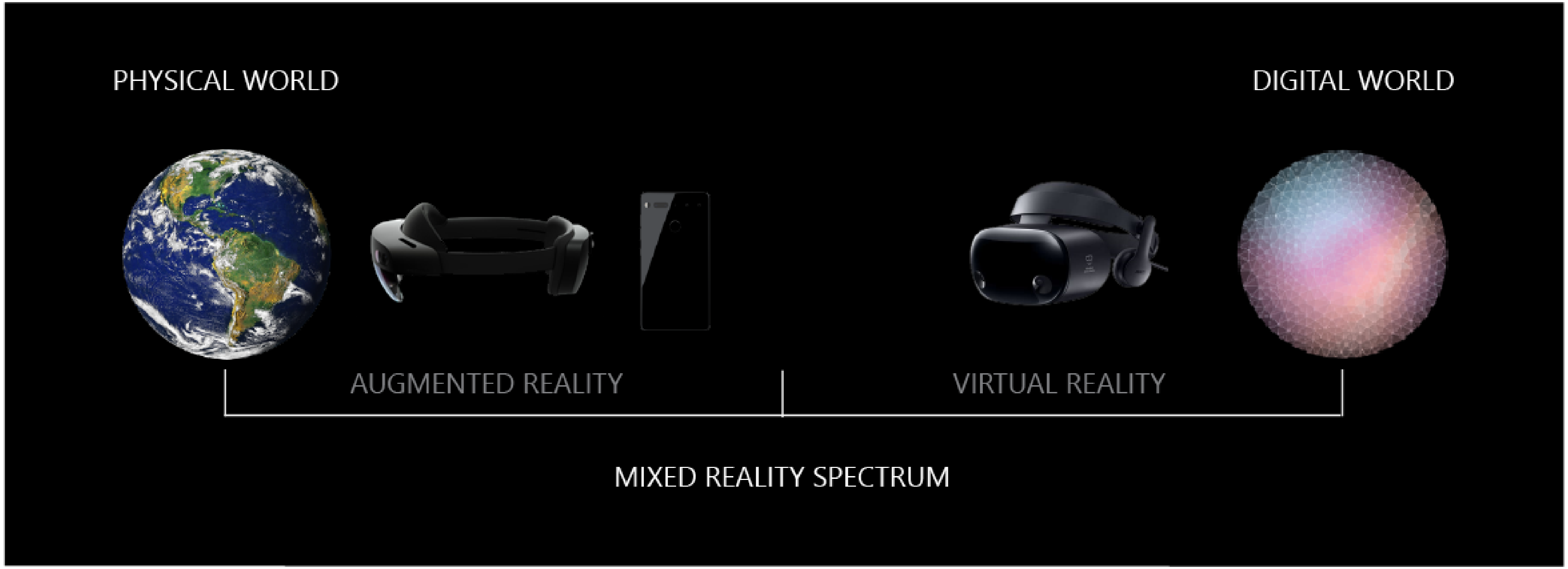
- Computer processing powered by the cloud
- Advanced input methods
- Environmental perceptions

Environmental inputs capture

- Person's body position in the physical world (tracking)
- Objects, surfaces, and boundaries (spatial mapping and scene understanding)
- Ambient lighting and sound
- Object recognition
- Physical locations

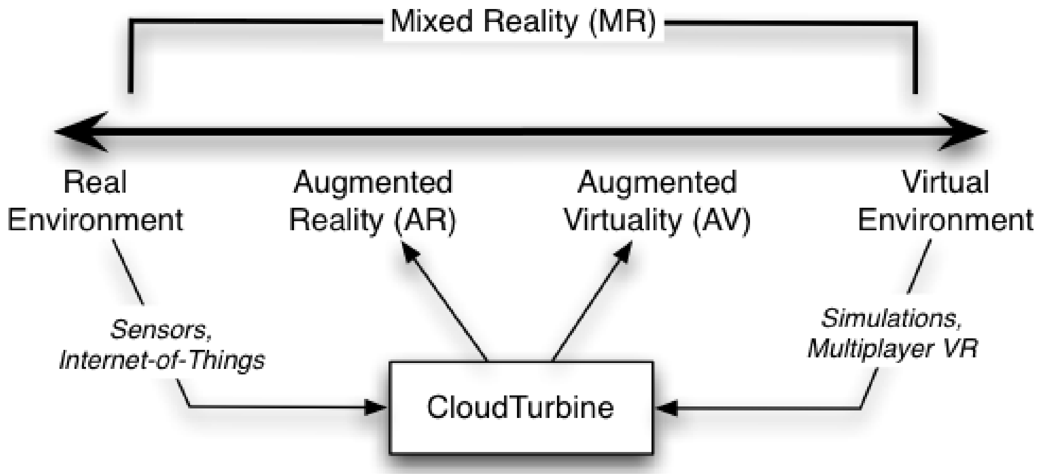
MR includes

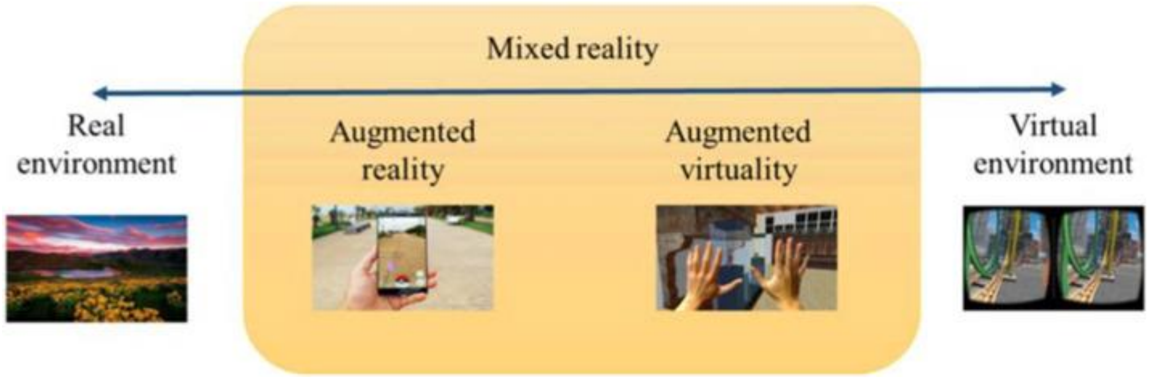
1. Environmental understanding: spatial mapping and anchors.
2. Human understanding: hand-tracking, eye-tracking, and speech input.
3. Spatial sound.
4. Locations and positioning in both physical and virtual spaces.
5. Collaboration on 3D assets in mixed reality spaces.

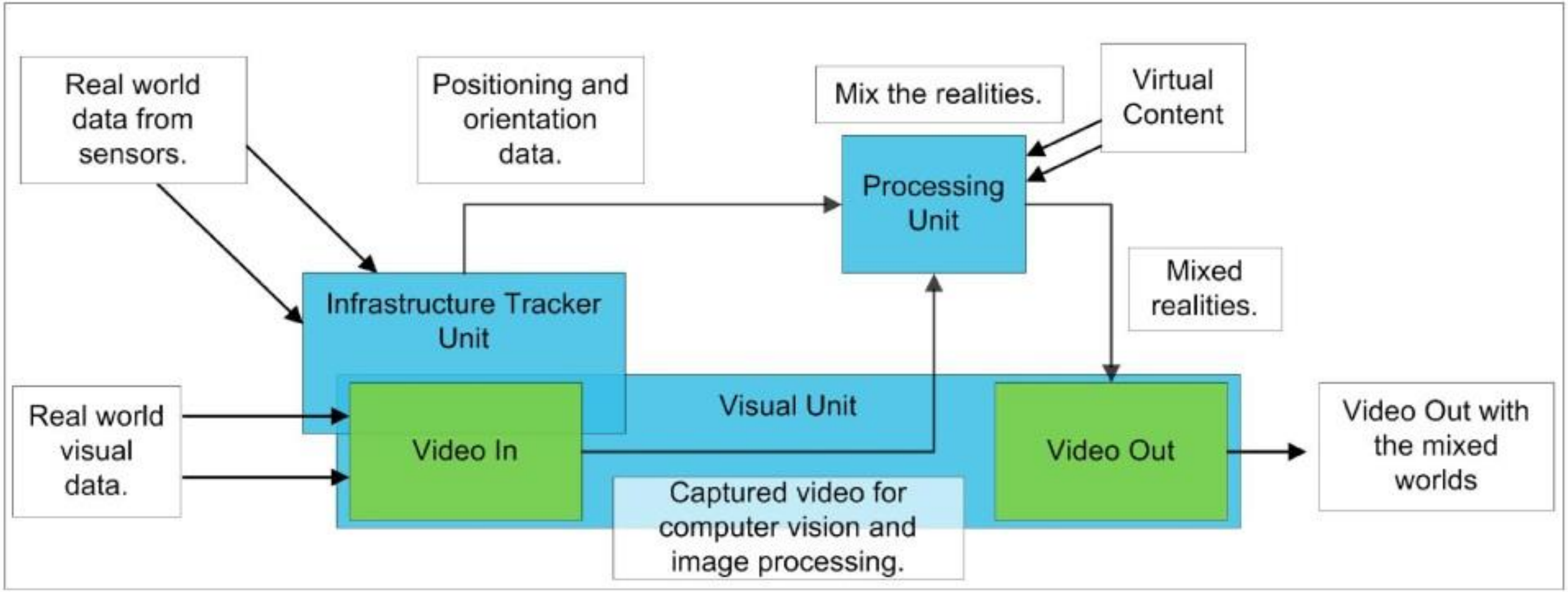


- **Physical reality -** Users remain present in their physical reality, and aren't made to believe they have left that reality.
- **Digital reality -** Users experience a digital reality and are unaware of the physical reality around them.
- **Mixed Reality -** These experiences blend the real world and the digital world.

Mixed Reality







Real world data from sensors.
Positioning and orientation data.
Mix the realities.
Virtual Content
Processing Unit
Infrastructure Tracker Unit
Mixed realities.
Real world visual data.
Video In
Visual Unit
Video Out
Video Out with the mixed worlds
Captured video for computer vision and image processing.

Device

There are two main types of devices that deliver Windows Mixed Reality experiences:

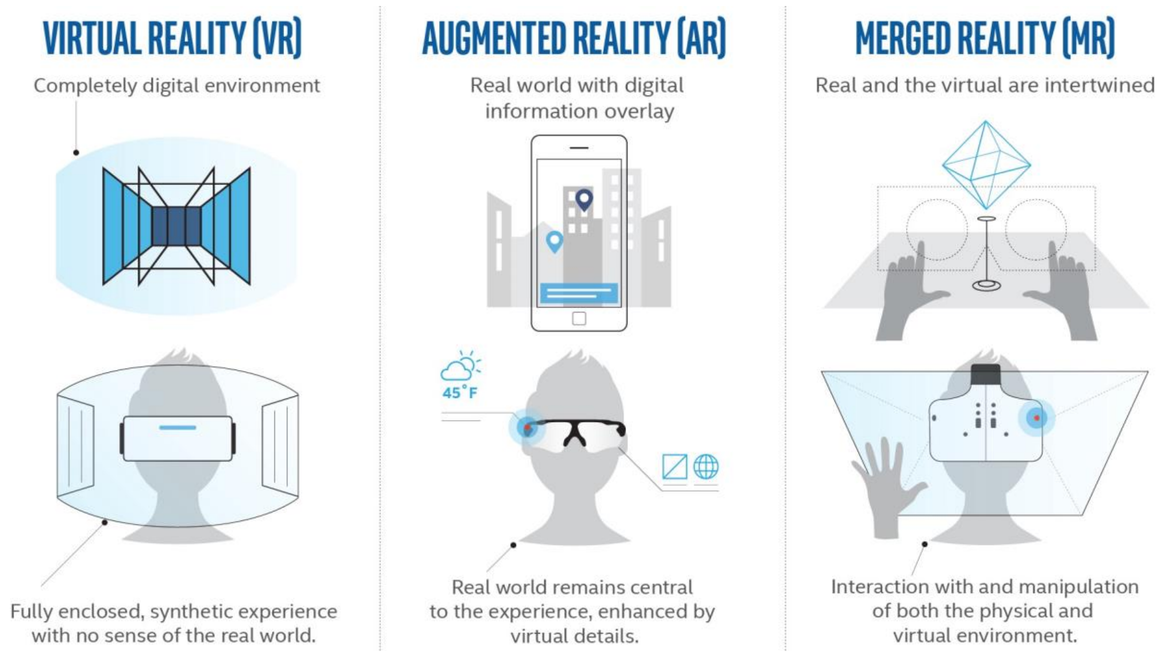
1. **Holographic devices** are characterized by the device's ability to place digital content in the real world as if it were there.
2. **Immersive devices** are characterized by the device's ability to create a sense of "presence"--hiding the physical world, and replacing it with a digital experience.

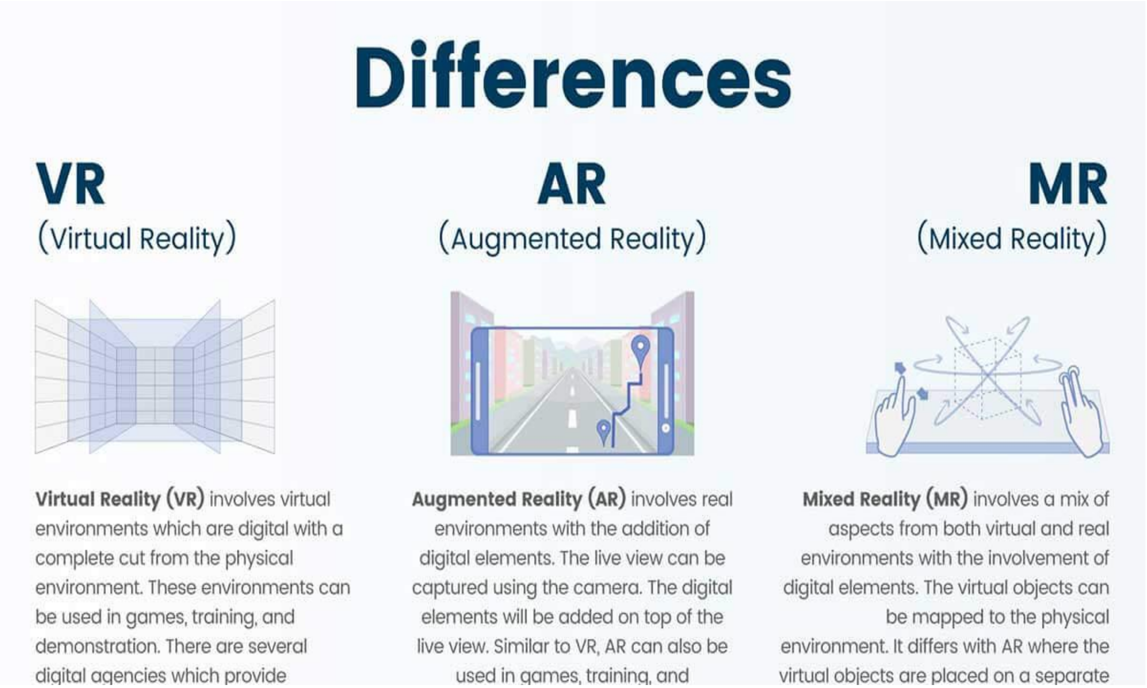
Characteristic	Holographic devices	Immersive devices
Example device	Microsoft HoloLens 	Samsung HMD Odyssey+ 
Display	See-through display. Allows user to see the physical environment while wearing the headset.	Opaque display. Blocks out the physical environment while wearing the headset.
Movement	Full six-degrees-of-freedom movement, both rotation and translation.	Full six-degrees-of-freedom movement, both rotation and translation.

MR

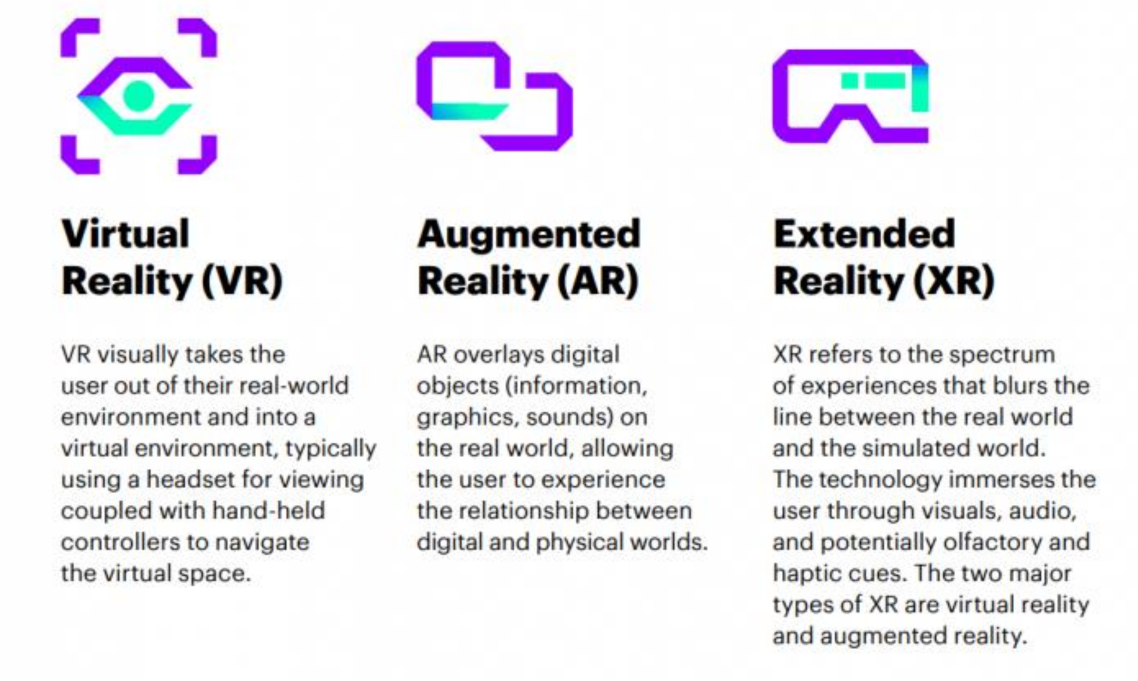
https://hololens.reality.news/news/hololens-assists-live-surgery-0178887/

VR, AR, MR


VIRTUAL REALITY (VR)
Completely digital environment
Fully enclosed, synthetic experience with no sense of the real world.
AUGMENTED REALITY (AR)
Real world with digital information overlay
45˚F
Real world remains central to the experience, enhanced by virtual details.
MERGED REALITY (MR)
Real and the virtual are intertwined
Interaction with and manipulation of both the physical and virtual environment.


Differences
VR
(Virtual Reality)
AR
(Augmented Reality)
MR
(Mixed Reality)
Virtual Reality (VR) involves virtual environments which are digital with a complete cut from the physical environment. These environments can be used in games, training, and demonstration. There are several digital agencies which provide
Augmented Reality (AR) involves real environments with the addition of digital elements. The live view can be captured using the camera. The digital elements will be added on top of the live view. Similar to VR, AR can also be used in games, training, and
Mixed Reality (MR) involves a mix of aspects from both virtual and real environments with the involvement of digital elements. The virtual objects can be mapped to the physical environment. It differs with AR where the virtual objects are placed on a separate

VR, AR, XR


Virtual Reality (VR)
VR visually takes the user out of their real-world environment and into a virtual environment, typically using a headset for viewing coupled with hand-held controllers to navigate the virtual space.
Augmented Reality (AR)
AR overlays digital objects (information, graphics, sounds) on the real world, allowing the user to experience the relationship between digital and physical worlds.
Extended Reality (XR)
XR refers to the spectrum of experiences that blurs the line between the real world and the simulated world. The technology immerses the user through visuals, audio, and potentially olfactory and haptic cues. The two major types of XR are virtual reality and augmented reality.


Solving the key XR technology challenges ahead
Display
Displaying richer visual content, and switching seamlessly between fully and partially virtual worlds
Common illumination
Making virtual objects in augmented worlds indistinguishable from real objects within the same view
Motion tracking
Intelligent, completely on-device tracking for intuitive head, hands, and eye interactions
Power and thermal
All day battery life, years of recharging, and compatible with sleek, thin, light passively cooled devices with no fans
Connectivity
The next level of ubiquitous, wireless connectivity for anywhere usage at fiber-optic speeds

Types of Reality

- **Augmented Reality (AR)** is an overlay of computer generated content on the real world that can superficially interact with the environment in real-time. With AR, there is no occlusion between CG content and the real-world.
- **Virtual Reality (VR)** encompasses all immersive experiences. These could be created using purely real-world content (360 Video), purely synthetic content (Computer Generated), or a hybrid of both.
- **Mixed Reality (MR)** is an overlay of synthetic content that is anchored to and interacts with objects in the real world—in real time. Mixed Reality experiences exhibit occlusion, in that the computer-generated objects are visibly obscured by objects in the physical environment.
- **Extended Reality (XR)** refers to all real-and-virtual environments generated by computer technology and wearables. The 'X' in XR is a variable that can stand for any letter.

Reference

- Alan B. Craig, "Understanding Augmented Reality, Concepts and Applications", Morgan Kaufmann, 2013.
- Images and videos used in this presentation are referred from the Internet source

Course Contents

- UNIT I Introduction of Virtual Reality: Fundamental Concept and Components of Virtual Reality - Primary Features and Present Development on Virtual Reality - Multiple Models of Input and Output Interface in Virtual Reality: Input - Tracker - Sensor - Digital Glove - Movement Capture - Video-based Input - 3D Menus & 3D Scanner – Output - Visual /Auditory / Haptic Devices.
- UNIT II Visual Computation in Virtual Reality: Fundamentals of Computer Graphics - Software and Hardware Technology on Stereoscopic Display - Advanced Techniques in CG: Management of Large Scale Environments & Real Time Rendering.
- UNIT III Interactive Techniques in Virtual Reality: Body Track - Hand Gesture - 3D Menus - Object Grasp. Development Tools and Frameworks in Virtual Reality: Frameworks of Software Development Tools in VR. X3D Standard; Vega - MultiGen - Virtools.

Course Contents – Contd…

- UNIT IV Application of VR in Digital Entertainment: VR Technology in Film & TV Production - VR Technology in Physical Exercises and Games - Demonstration of Digital Entertainment by VR.
- UNIT V Augmented and Mixed Reality: Taxonomy - technology and features of augmented reality - difference between AR and VR - Challenges with AR - AR systems and functionality - Augmented reality methods - visualization techniques for augmented reality - wireless displays in educational augmented reality applications - mobile projection interfaces - marker-less tracking for augmented reality - enhancing interactivity in AR environments - evaluating AR systems.

Course Objectives

- To know basic concepts of virtual reality
- To understand visual computation in computer graphics
- To understand interaction between system and computer
- To know application of VR in Digital Entertainment
- To know basic concepts of augmented reality

Course Outcomes

- Provide opportunity to explore the research issues in Augmented Reality and Virtual Reality (AR & VR)
- Know the basic concept and framework of virtual reality
- Understand fundamentals of computer graphics
- Know the computer-human interaction
- Develop simulator for real time application using AR & VR

Books

Text Books

1. Burdea, G. C., P. Coffet., “Virtual Reality Technology”, Second Edition, Wiley-IEEE Press, 2003/2006
2. Alan B. Craig, “Understanding Augmented Reality, Concepts and Applications”, Morgan Kaufmann, 2013.

Reference Books

1. Alan Craig, William Sherman, Jeffrey Will, “Developing Virtual Reality Applications, Foundations of Effective Design”, Morgan Kaufmann, 2009.
2. Hearn, Donald, Baker, M. Pauline , “Computer Graphics”.

Thank You

ALL THE BEST….