Augmented and Mixed Reality

What is Augmented Reality????

- **Augmented reality (AR)** is an **interactive experience** that seamlessly blends the **real world** with **computer-generated content**.
- AR combines elements of the physical environment with digital information. These elements/modalities can engage multiple senses, including **visual**, **auditory**, **haptic**, and even **olfactory** sensations.
- Unlike **virtual reality (VR)**, which completely replaces the real world with a simulated one, AR enhances our perception of the existing environment by overlaying digital content onto it.

How Augmented Reality works????

AR systems incorporate **three key features**:

1. **Real and Virtual Worlds**: AR seamlessly merges real-world surroundings with virtual objects or information.
2. **Real-Time Interaction**: Users can interact with the overlaid content in real time.
3. **Accurate 3D Registration**: Virtual objects align precisely with their real-world counterparts.

Contd...

- The sensory information can be **constructive** (adding to the natural environment) or **destructive** (masking parts of the real world).
- Imagine seeing digital annotations on a physical object, like step-by-step instructions projected onto a car engine during repairs.

Augmented Reality Used for……

- Augmented reality either makes visual changes to a natural environment or enhances that environment by adding new information.
- It can be used for various purposes, including gaming, product visualization, marketing campaigns, architecture and home design, education, and industrial manufacturing.

Basic Augmented Reality process

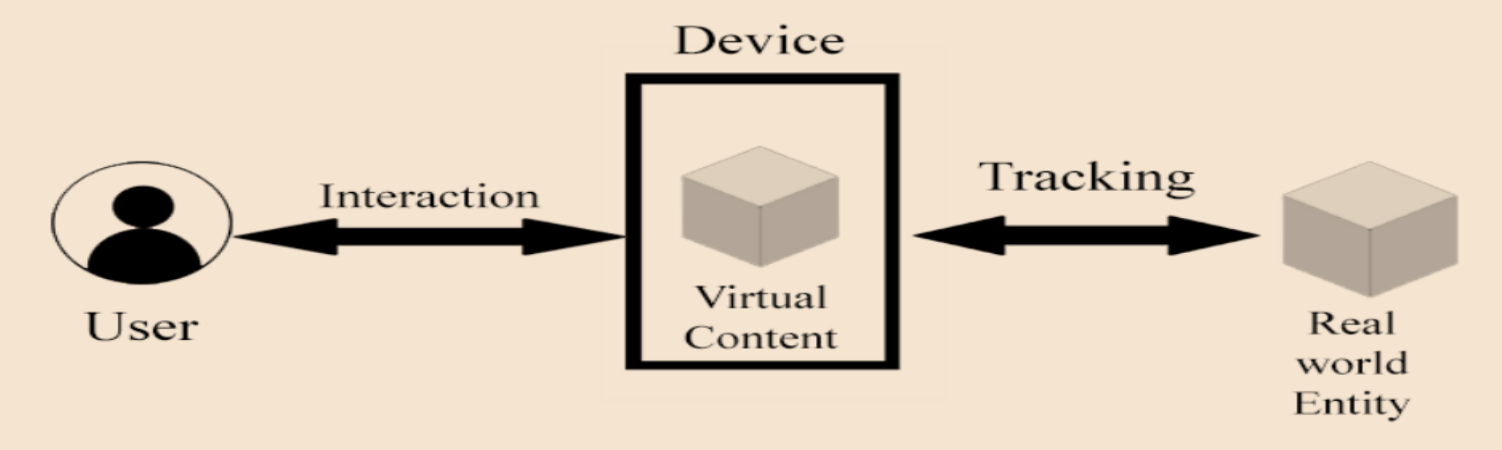
- **Input:** One or more sensors capture information from the real-world.
- **Processing:** The processing unit interprets the data acquired through different *hardware* and *software* resources.
- **Output:** It displays to the user about both information providing the feeling that virtual objects are part of the real-world.

Architecture of Augmented Reality

- Augmented Reality is technology in computer graphics which combines the real time environment with the digital one.
- In Virtual reality users totally experience new world while in augmented reality digital information display over the real environment.
- To experience Augmented Reality user need AR headset.

Architecture of Augmented Reality

- Augmented Reality is technology in computer graphics which combines the real time environment with the digital one.
- In Virtual reality users totally experience new world while in augmented reality digital information display over the real environment.
- To experience Augmented Reality user need AR headset.



Augmented Reality Architecture

Contd…

This Architecture comprised of all above components and interactive relationship between them helps to develop augmented reality working model.

1.User: The most essential part of augmented reality is its user. The user can be a student, doctor, employee. This user is responsible for creation of AR models.

2.Interaction: It is a process between device and user. The word itself consist of its meaning some action perform by one entity as result in creation or some action performed by other entity.

Contd…

3.Device: This component is responsible for creation, display and interaction of 3D models. The device can be portal or in static state. Example--- mobile, computer, AR headsets etc.

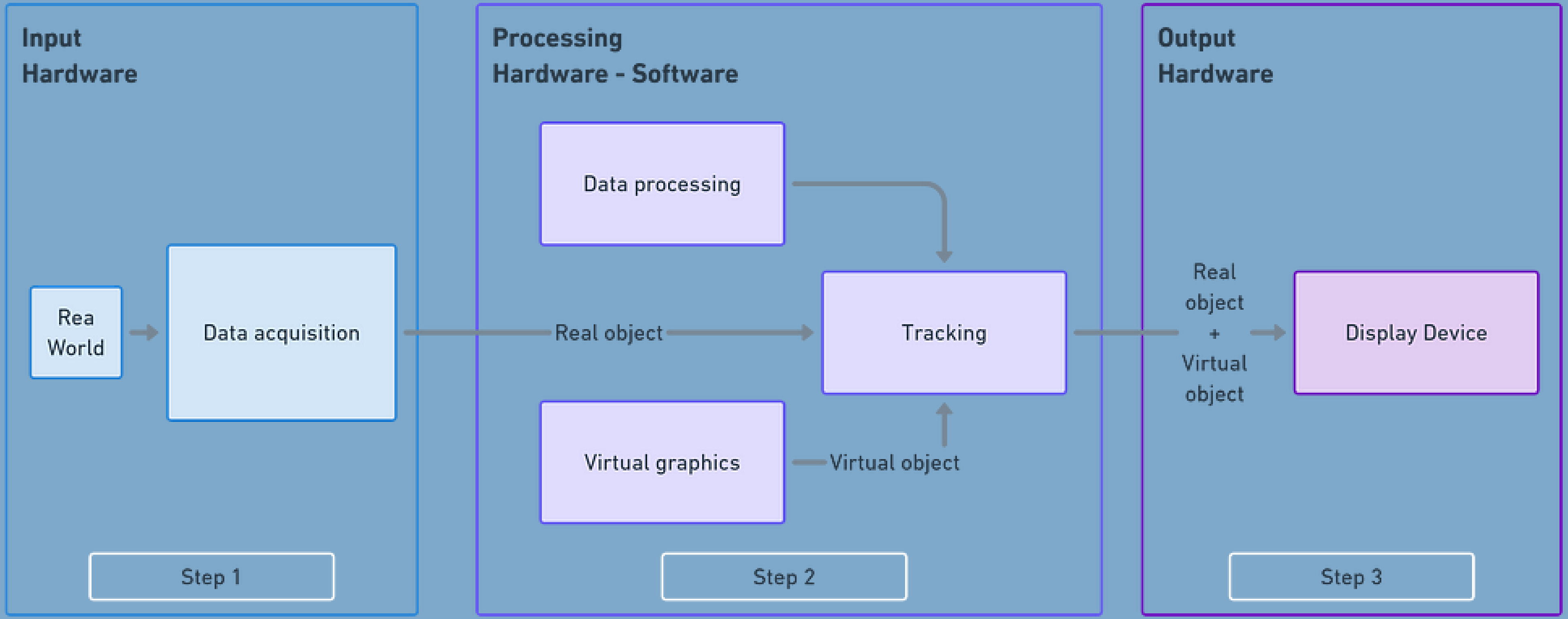
4.Virtual Content: The virtual content is nothing but the 3D model created or generated by the system or AR application. Virtual content is type of information that can be integrated in real world user's environment. This Virtual content can be 3D models, texture, text, images etc.

Contd...

5. Tracking: This component is basically process which makes possible creation of AR models. Tracking is sort of algorithm which help to determine the device where to place or integrate the 3D model in real world environment. There are many types of Tracking algorithm available which can be used in development of AR applications.

6. Real-life entity: The last component AR architecture is real world entities. This entities can be tree, book, fruits, computer or anything which is visible in screen. AR application does not change position of real life entity. It only integrate the digital information with this entities.

Components/Process of Augmented Reality



Contd..

❑ Hardware

The main characteristic of the hardware components is to **acquire** and **display** the data and information, and **process** it.

•**Input — Sensors:** Different types of sensors that respond to physical or chemical stimuli from the real environment and provide the necessary data for the development of the system.

Hardware	System	Type	Example
Input	Sensor	Optical	Cameras, Infra-red
		Magnetic	Compasses
		Inertial	Accelerometer, Gyroscope
		Others	GPS, Depth

Contd..

•**Output — Display:** The devices for displaying the information can be divided into **wearable** and **non-wearable**. But it can also be classified into **optical**, **video,** and projection devices.

Hardware	System	Type	Example
Output	Wearable	Optical	Helmets, Glasses
		Video	Head-Up Display
	Non-wearable		Smartphone, Tablets, PC
		Projection	Projectors

Contd..

Software

The main characteristic of software components is to interpret the acquired data to transform and augmented it.

•**High-level:** Currently, several tools help in the development of applications, known as SDK (Software Development Kit), such as **ARKit, ARCore, ARtoolKit, EasyAR, LayAR**.

•**Low-level** consists of different areas, which may vary according to the application or the developer’s needs, such as programming libraries, Computer vision, CG, Image processing, HCI.

Artificial Intelligence role in AR

- **Computer vision** algorithms, a subset of AI, play a crucial role in AR.
    - **Object Recognition and Tracking**: AI empowers AR devices to accurately recognize and track objects. Machine learning models swiftly identify and augment real-world objects, seamlessly blending the virtual and physical realms.
- **User tracking**
    - AI can track the user's movements and gestures.

Contd…

- **Personalization**
  - AI can analyze user data to provide tailored AR experiences.
  - ***For example***, Amazon uses AI to provide personalized recommendations to customers based on their purchase history and browsing behavior.

- **Text recognition and translation**
  - AI Optical Character Recognition (OCR) techniques combine with text-to-text translation engines such as DeepL. A visual tracker keeps track of the word and allows the translation to overlay the AR environment.

Types of Augmented Reality

- Marker-Based Augmented Reality
- Marker less Augmented Reality
- Projection-Based Augmented Reality
- Contour-Based Augmented Reality

Link:

https://www.youtube.com/watch?v=QMATJIlKnyE

Marker-Based Augmented Reality

- In this type, predefined visual markers (such as QR codes) trigger augmented experiences.
- Imagine scanning a QR code during a self-guided tour to learn more about a historical site.
- The app recognizes the marker and overlays relevant information based on its programming.
- These markers are also known as **fiducial markers**.

Example of Marker-Based AR

1) **Google Maps AR Navigation**: Imagine walking down the street while your phone overlays real-time directions onto your view. This system uses GPS and visual recognition to guide you.

2) **Microsoft HoloLens Spatial Mapping**: The HoloLens headset scans the environment, creating a 3D map. Users can interact with holograms that align precisely with real-world objects.

Marker-less Augmented Reality

- Markerless AR doesn't rely on specific markers but instead uses other cues for activation.
- Within markerless AR, we have several subtypes:

1) **Overlay AR**: Superimposes digital content onto the real world. For instance, you might see virtual furniture placed in your living room using an app.

NOTE: AR cues are designed to explicitly guide users to engage with the environment and actively learn their surroundings.

Contd…

2) Location-Based AR (AR with GPS): Integrates real-time location data to enhance experiences. It can provide directions, information about nearby landmarks, or geotagged content.

3) Surface-Based AR (World Tracking or SLAM): Maps and tracks surfaces in the environment. This enables 3D models to interact seamlessly with physical objects.

4) Spatial Tracking: Allows AR experiences to persist across different spaces.

Example of Marker-less AR

•**<u>Image Tracking</u>**: AR content is anchored to specific images or objects in the real world.

Example: **IKEA Place App**, which lets you visualize furniture in your home by placing 3D models on flat surfaces.

•**<u>Location-Based AR (AR with GPS)</u>**:Uses GPS data to provide context-aware experiences.

Example: Volkswagen's AR HUD Navigation System, which overlays driving directions onto the windshield.

Projection-Based Augmented Reality

- Projects immersive light onto flat surfaces to create 3D imagery.
- Combines projection with SLAM for accurate placement.

Example:

1) Notable applications include interactive art installations and architectural visualizations.
2) An advertisement projected onto a large building or another surface.

Contour-Based Augmented Reality

- Analyzes the contours of objects to align AR content.
- Useful for applications like virtual try-ons for clothing or accessories.

Example: a contour-based AR framework can be used in car navigation systems so that drivers can safely navigate low-visibility roads.

Algorithms in AR

- Simultaneous Localization and Mapping (SLAM)
- Depth Sensing
- Machine Learning in AR
- Reality-Virtuality Continuum
- Feature Detection

Simultaneous Localization and Mapping (SLAM)

•SLAM is a foundational algorithm in AR. It combines **computer vision techniques** to map and track the environment.

- **Mapping**: SLAM constructs a map of the surroundings by comparing visual features between camera frames.
- **Localization**: It simultaneously estimates the camera's position within this map.

•SLAM enables AR applications to understand the user's environment and overlay digital content seamlessly.

LINK: https://www.flyability.com/simultaneous-localization-and-mapping

Components of SLAM

1) Range Measurement

- All SLAM solutions include some kind of device or tool that allows a robot or other vehicle to observe and measure the environment around it.

2) Data Extraction

- After the range measurement, SLAM system must have some sort of software that helps to interpret that data.
- All of these “back-end” solutions essentially serve the same purpose though: they extract the sensory data collected by the range measurement device and use it to identify landmarks within an unknown environment.

Depth Sensing

- To make AR experiences more realistic, depth sensing plays a crucial role.
- Traditional AR algorithms use image markers and natural feature registration. However, they lack awareness of the environment's existing targets.
- Depth sensing enhances AR by providing **information about the distance to objects**, allowing more accurate placement of virtual elements.

Contd…

- Depth sensing algorithms in augmented reality (AR) use machine learning to compare multiple device images from different angles to estimate the distance to every pixel as a user moves their phone.
- This algorithm can increase depth processing even with minimal motion from a user.
- Depth sensing requires three closely-linked tasks: image acquisition, processing, and analysis.

Contd…

- AR uses depth sensors, accelerometers, cameras, gyroscopes, and light sensors to collect data on the user's surroundings.
- These sensors measure the distance to the objects, speed of the motion, direction and angle, and overall orientation in space.

Key points of Depth Sensing

•**Accuracy**: Depth sensing algorithms vary in accuracy, with stereo vision often providing high precision.
•**Hardware**: Depth sensors are integrated into devices for real-time depth estimation.
•**Integration with AR**: Depth maps help align virtual objects with the real world, improving realism.

Techniques of Depth Sensing

- **Structured Light**: Projects patterns onto the scene and analyzes their deformation to compute depth.
- **Time-of-Flight (ToF)**: Measures the time taken for light to travel to the object and back.
- **Stereo Vision**: Compares images from two cameras to calculate depth.
- **Depth from Monocular Images**: Infers depth from a single image using machine learning or geometric cues.

Use Cases of Depth Sensing

1) Object Occlusion:

- Occlusion refers to accurately rendering virtual objects behind real-world objects.
- For instance, consider placing a virtual object (let's call it "Andy") near a wooden trunk. Without occlusion, Andy might overlap unrealistically with the trunk.
- By leveraging depth information, we can render Andy with proper occlusion, making it seamlessly blend into its surroundings.

Contd…

2) Scene Transformation:

- Depth enables us to create immersive scenes where virtual elements interact with real-world objects.
- Imagine rendering virtual snowflakes settling on the arms and pillows of a user's couch or casting a living room in misty fog.

3) Distance and Depth of Field:

- The Depth API helps show depth cues like distance.
- By measuring distances, we can apply depth-of-field effects, such as blurring the background or foreground of a scene.

Machine Learning in AR

- AR often combines SLAM with sensor data from smartphone gyroscopes and accelerometers.
- Deep learning techniques can improve tracking reliability by fusing visual data with sensor information.
- By training models on large datasets, AR systems can better recognize and interpret real-world scenes.

Contd…

1) Machine Learning for Face Tracking

- Face tracking is a great example of how machine learning can enhance augmented reality.
- **For example**, Apple’s new iPhone X uses machine learning to enable Face ID, allowing users to unlock their phone by looking at it rather than entering a passcode.
- Likewise, Microsoft uses Face Tracking in its HoloLens, enabling users to use their faces as inputs for navigating and interacting with holograms.

Contd…

2) Natural Language Processing

- One of augmented reality's biggest hurdles is natural language processing (NLP).
- This allows AR apps to recognize and respond to text that appears in their environment.
- So if a user were wearing an AR headset and walking through a city, they could see text overlaid on signs, walls and billboards.
- The NLP algorithms would be able to read that text and, depending on what it says, could trigger a response or action.
- Contd….

Contd…

- **For example**, if someone walks by a restaurant with an available table then sees No tables available tonight printed across its storefront window, he or she might say Ok Google, show me restaurants with tables open nearby without having to do any additional typing or clicking.

Contd…

3) Facial Expression Recognition

- Facial recognition is a computer science term for technology that identifies and tracks faces in a digital image or video.
- It’s commonly used in security systems so you can unlock your front door using just your face, it’s also widely used by advertisers to make targeted offers to consumers.
- Facial recognition is an integral part of virtual reality (VR) because it allows you to identify friends and enemiews through cameras that are built into VR headsets.
- Facial recognition is also being used to detect and measure emotions, a feature that can be useful for some VR training applications.

•Contd….

Contd…

- **For example,** by using facial expression recognition technology and software, doctors can train for complex surgeries like heart transplants, which involve emotionally-charged interactions with patients.
- Facial recognition tech is so good now that it’s now being used in smart TVs to allow you to automatically log in simply by giving a quick glance at your screen.
- It’s also widely used in CCTV systems so law enforcement officials can identify suspects from large crowds and review video footage of crimes from different angles.

Reality-Virtuality Continuum

•Proposed by Paul Milgram, this continuum spans from completely virtual environments to pure reality.
•AR falls in the middle as "mixed reality," where digital elements coexist with the real world.

Examples include Google Glass, which displays digital content in front of the user's eye, and Microsoft HoloLens, providing real-world information through a headset.

Challenges of Augmented Reality

Implementation Cost: The perceived cost of implementing AR is a common concern. Customizing an AR app can range from a few thousand dollars to well into the six-figure range, depending on complexity.

For instance, IKEA's AR-powered app, which allows customers to virtually visualize furniture in their homes, likely cost between $30,000 and $60,000 to build and implement.

Challenges of Augmented Reality

- **Technology and Skills Gaps**: Executives express concerns about talent and technological gaps. Companies need guidance on training, use-cases, implementation, and maintenance. AR itself can be used to upskill a workforce, making companies more resilient to future disruptions.
- **Resource Shortage and Competing Priorities**: Devoting shortage of financial and human resources to AR implementation can be challenging. Balancing competing priorities while integrating AR into existing workflows requires strategic planning.

Challenges of Augmented Reality

•**Hardware Limitations**: The lack of consumer-ready AR headsets poses a significant challenge. Without suitable hardware, widespread adoption remains hindered.

•**Privacy and Security Concerns**: As AR interacts with the physical environment, privacy and security become critical. Protecting user data and ensuring secure interactions are ongoing challenges.

•**Content Creation**: Creating engaging and relevant AR content demands creativity and expertise. Designing interactive experiences that enhance user engagement is essential.

Object Grasp in VR

- In the context of **Virtual Reality (VR)**, **object grasp** refers to the interaction technique that allows users to manipulate and hold virtual objects within the VR environment.

1) Single Object Grasp
2) Multi Object Grasp

Single Object Grasp

- Most VR applications implement object grasping by attaching a single object to the controller.
- This approach suffices for many actions, such as picking up an apple or wielding a sword.
- However, it doesn’t allow users to grasp **multiple objects** simultaneously and release them in a controlled manner.

Multiple Object Grasp

- A **multi-object grasp technique** that enables users to hold and manipulate multiple objects in one hand within the virtual environment.
- This technique allows users to release the objects in a controlled way, mimicking real-world scenarios where tasks involve grasping and placing multiple items.
- Imagine laying out surgical instruments on a table or arranging items on shelves—these are the types of tasks that benefit from realistic multi-object grasping in VR.