Tracking

AUGMENTED REALITY

Group No. 2

Team members

2021UCS1511 - Siddhant Basu
2021UCS1512 - Vievaan Khattar
2021UCS1513 - Farhan Khan
2021UCS1514 - Zaid Khan
2021UCS1515 - Vikrant Dabas
2021UCS1516 - Megha Gaur
2021UCS1517 - Komal
2021UCS1518 - Alex Guria
2021UCS1519 - Aditya Giri
2021UCS1520 - Sharil Malik

Index

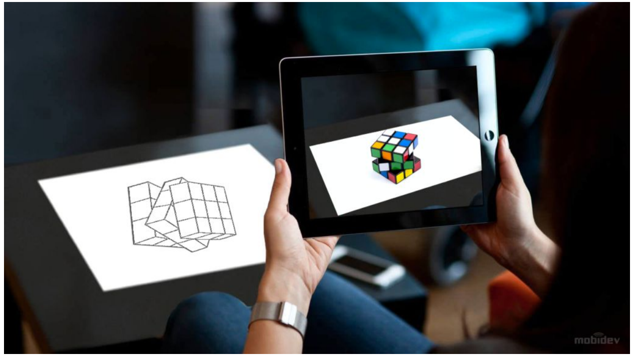
Introduction to Tracking in Augmented Reality

Siddhant Basu - 2021UCS1511

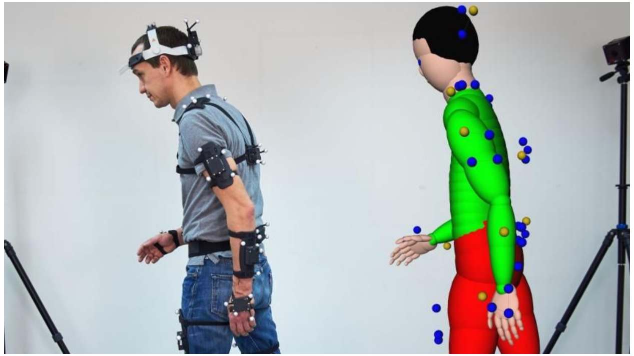
What is AR Tracking?

Definition : AR tracking refers to determining the position and orientation of a device relative to the real world.

Importance : Critical for placing virtual objects accurately in the physical environment.

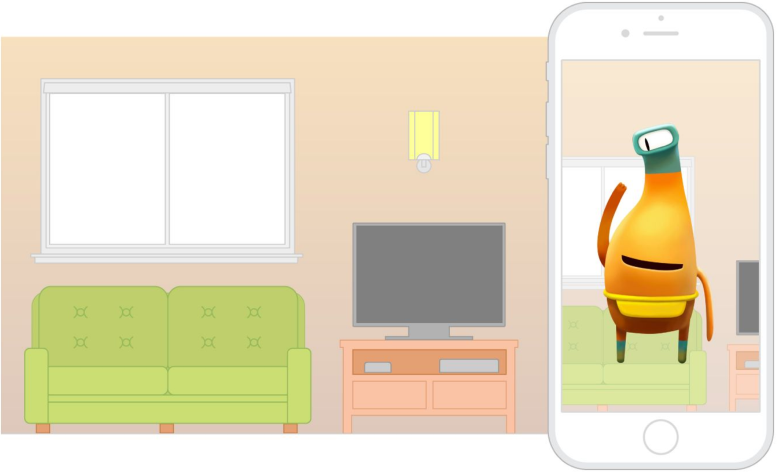


Positional Tracking



Realtime Tracking

Importance of Accurate Tracking



Ensures virtual objects stay anchored in the real world. Enables immersive and realistic AR experiences.

Impacts use cases like navigation, gaming, industrial applications.

Types of Tracking Methods

Sensor Based Tracking

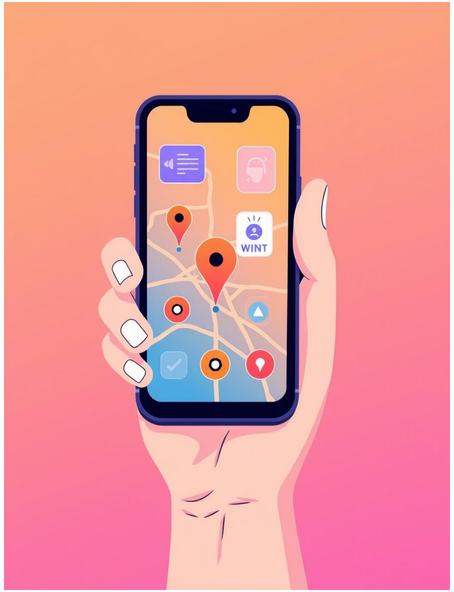
Visual Based Tracking

Hybrid Tracking

Sensor-Based Tracking: Overview

Sensor-based tracking involves using sensors to determine the position, orientation, or movement of an object or person. This technology has become increasingly prevalent in various applications, including robotics, healthcare, and sports.

Sensor-based tracking systems are crucial in various AR applications for ensuring accurate movement and orientation data. Each sensor type has its strengths and weaknesses, often being combined to improve performance.



Types of sensor based tracking

1.Inertial Tracker: Utilizes accelerometers and gyroscopes

2.Acoustic Tracker: Uses ultrasonic sound waves

3.Mechanical Tracker: Operates via physical mechanisms

Magnetic Tracker: Leveraging Magnetic Fields

Magnetic Tracking is a method of sensor-based tracking that uses magnetic fields to determine the position and orientation of objects.

How It Works:

- A **magnetic field** is generated by a source (transmitter).
- Sensors (receivers) measure the strength and direction of the magnetic field to calculate the position and orientation of the object being tracked.
- Since magnetic fields can penetrate through most materials, **line-of-sight** is not always necessary for tracking.

Pros

No Line-of-Sight Required: Magnetic tracking works even if there are obstacles between the source and the sensor.
High Precision: Can provide highly accurate position data, especially in small spaces.
Low Latency: Quick response times, making it suitable for real-time applications like AR and VR

Cons

Interference: Magnetic fields can be disrupted by nearby metals or electrical equipment, reducing accuracy.
Limited Range: The accuracy of magnetic tracking decreases over longer distances.
Calibration Issues: Requires careful calibration to minimize errors due to environmental interference.

Acoustic (Ultrasonic) Tracker: Tracking with Sound Waves

Acoustic (Ultrasonic) Tracking is a method that uses high-frequency sound waves (ultrasound) to track the position and movement of objects. It is commonly used in applications where short-range tracking is required.

How It Works:

- **Ultrasonic transmitters** emit high-frequency sound waves, which are inaudible to humans.
- These sound waves travel through the air and are detected by **ultrasonic receivers**.
- By measuring the **time it takes** for the sound waves to travel from the transmitter to the receiver (Time of Flight), the system can calculate the distance between them.
- Using multiple transmitters and receivers, the system can determine the **precise position** and movement of the tracked object.

Pros

- **Short-range accuracy**: Very effective for precise tracking over short distances.
- **Low cost**: Relatively inexpensive compared to other tracking technologies.
- **No interference from light**: Works well in environments where optical tracking may fail, such as in low-light or high-reflection conditions.

Cons

Limited Range: Effective only over short distances, typically a few meters.
Susceptible to Noise: Accuracy can be affected by background noise or other sounds that interfere with the ultrasonic waves.
Obstacle Interference: Sound waves can be blocked or reflected by physical objects, reducing the accuracy of tracking.

Optical Tracking

Optical tracking uses cameras to capture and analyze the position and movement of objects by detecting physical markers or natural features.

How It Works:

- Cameras track **visual markers** (e.g., QR codes, LEDs) or **natural features** on objects in the environment.
- The system uses **computer vision algorithms** to interpret the data and determine the object's position and orientation.
- It provides highly detailed and accurate tracking, especially when combined with depth sensors.

Pros:

- **High Accuracy:** Excellent for tracking fine movements and positions.
- **Real-time Feedback:** Offers fast updates, making it ideal for augmented reality (AR) and virtual reality (VR) applications.
- **Scalable Range:** Can cover both small and large areas depending on camera setup.

Cons:

- **Line-of-Sight Required:** Tracking is interrupted if the object is not visible to the camera.
- **Sensitive to Lighting Conditions:** Performance can degrade in poor lighting or when there are reflections.
- **Complex Setup:** Requires careful camera placement and calibration for reliable tracking.

Inertial Tracking

Inertial tracking relies on sensors like **accelerometers** and **gyroscopes** to detect motion and orientation changes.

How It Works:

- **Accelerometers** measure changes in velocity and direction (linear acceleration).
- **Gyroscopes** track rotational movements, detecting changes in orientation.
- Data from these sensors are processed to estimate the object's position and orientation in real-time.

Pros:

- **No External References:** Inertial tracking works independently without needing cameras or external markers.
- **Fast Response Time:** Provides immediate feedback, ideal for high-speed movement.
- **Portable and Lightweight:** Inertial sensors are small and can be embedded in devices like smartphones, VR headsets, or wearables.

Cons:

- **Drift Over Time:** Without recalibration, the tracking can become less accurate due to cumulative sensor errors (drift).
- **Limited Positional Accuracy:** While orientation tracking is accurate, position tracking is less precise compared to optical systems.

Optical and Inertial Tracking Combined: In many systems, optical and inertial tracking are **combined** to leverage the strengths of both:

- Optical tracking provides **positional accuracy**, while inertial sensors ensure **smooth orientation tracking**.
- This hybrid approach reduces drift and allows for robust tracking even when one method experiences interference.

Marker Based Tracking

Marker-based tracking is a technique used in augmented reality (AR) where the system relies on predefined markers, usually images or patterns, to recognize and track objects or locations in the physical world.

These markers are typically 2D patterns like QR codes or fiducial markers, which are easy for the system to detect due to their high contrast and distinct shapes. When a camera captures the marker, the AR software identifies it by matching it against a known database of markers.

Once recognized, the system can determine the marker's position and orientation in real time. This information allows digital content, such as 3D models or animations, to be precisely placed and aligned with the physical marker, creating the illusion that the virtual content exists in the real world.

Marker-based tracking is widely used in AR applications for its reliability and ease of implementation, especially in controlled environments where the markers can be placed in advance.

Zaid Khan 2021UCS1514

Point Fiducials

Point fiducials are specific, easily identifiable points in an image or physical environment or set by the user in the scene that serves as reference markers for tracking or measurement in various computer vision and augmented reality (AR) systems.

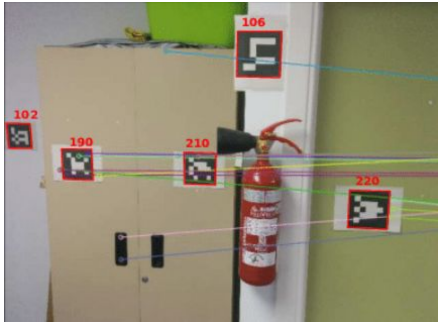
System can reliably detect these points and their orientations and use them to understand the spatial information. Fiducials help establish a coordinate system in which the position and orientation of objects can be accurately tracked relative to the points.

Planar Fiducials

Planar fiducials are flat, 2D markers used in augmented reality (AR) and computer vision systems for object tracking, localization, and alignment.

These fiducials are often designed with distinct geometric patterns, such as grids, checkerboards, or QR codes, that can be easily recognized and processed by a camera or sensor.

Unlike point fiducials, which are individual points, planar fiducials provide a surface with multiple reference points distributed over a plane. This allows the system to determine both the position and orientation (pose) of the marker in space with greater accuracy.



Vikrant Dabas - 2021UCS1515

Working of Marker-Based AR

The working of Marker-Based tracking in Augmented Reality can be broken down into the following steps:

Marker Detection	Pose Estimation	Content Overlay
Marker Detection involves high-contrast patterns using image processing techniques to identify predefined markers	Pose Estimation calculates the marker's position and orientation relative to the camera	In Content Overlay, digital elements are rendered in real-time on top of the physical markers

Marker Detection

Camera Scanning : The camera continuously scans and captures frames from the environment to search for predefined markers. For example, we use high-contrast images with black-white patterns for easy detection.

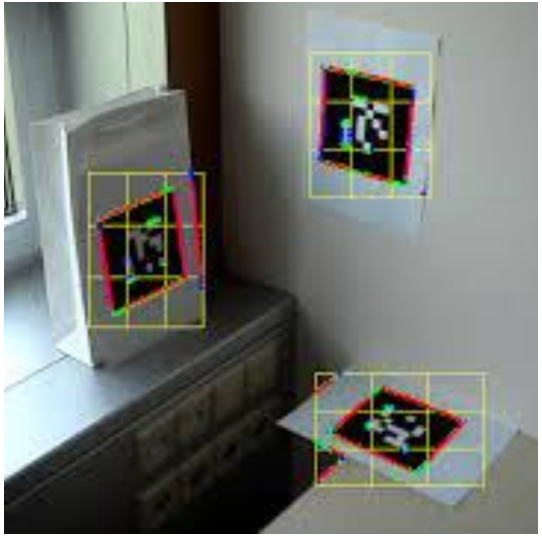
Image Processing : The frames captured by the camera are then processed to detect markers. This is done by converting the image to binary format to highlight the edges and contours. These shapes in the binary image are then detected by various algorithms. Each detected contour is analyzed and compared to known markers to confirm a match

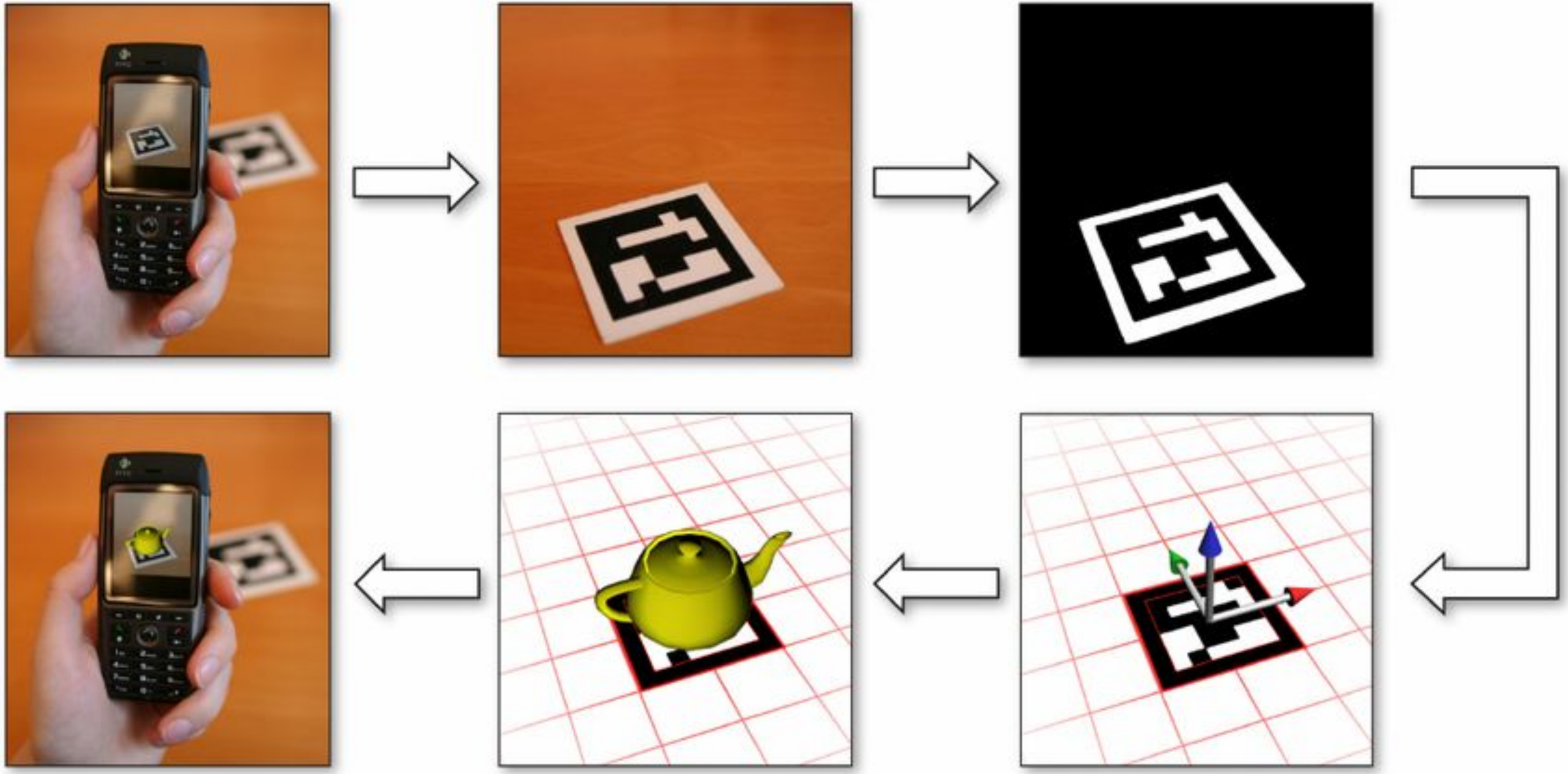
Pose Estimation

Pose : Pose refers to the marker's position and orientation relative to the camera, defined by six degrees of freedom (3D position + 3D orientation).

Algorithms : Pose estimation is done using various algorithms, these include Homography Transformation which calculates the transformation matrix that relates the marker's coordinates in the image to its real-world coordinates and Matrix Multiplication used to transform points in 3D space.

Content Overlay : Pose Estimation gives a precise spatial relationship for accurate rendering of digital content.





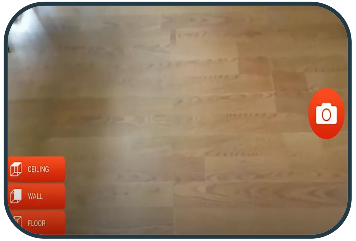
Marker-less Tracking

Marker–less tracking detects and tracks objects or environments without physical markers, using natural features like corners, edges, or textures.

Advanced algorithms interpret these features to determine the position and orientation of virtual objects, making it adaptable to different environments.

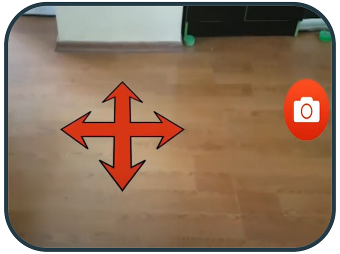
Fundamental processes

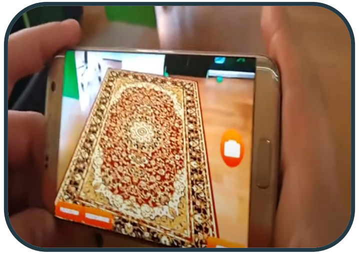
- **SLAM (Simultaneous Localization and Mapping)**: Tracks the device's position while mapping the environment in real time, anchoring virtual content to natural features.
- **Feature Detection**: Recognizes patterns like edges or textures to serve as reference points.
- **Environment Mapping**: Updates the map as the device moves for accurate virtual object placement.
- **Continuous Tracking**: Monitors movement to keep virtual content aligned with the real world.



No object

Plane detection and Indicator





Carpet is placed in the right location.

SLAM - Simultaneous Localization and mapping

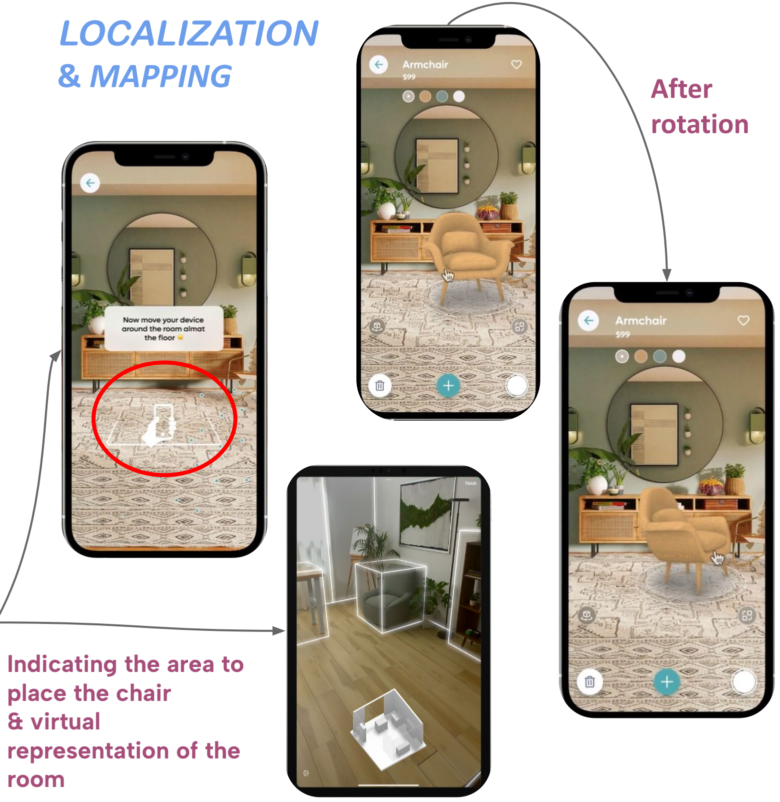
Simultaneous Localization and Mapping (SLAM) is a technique in AR that allows a device to track its position in real-time while simultaneously mapping its environment.

Localization:

It is the **process of determining** the device's **position** in a given space. In markerless AR, the system **continuously tracks the position and orientation** of the device (e.g., a smartphone or AR headset) relative to its surroundings. This enables the AR **experience to remain stable as users** move through the environment.

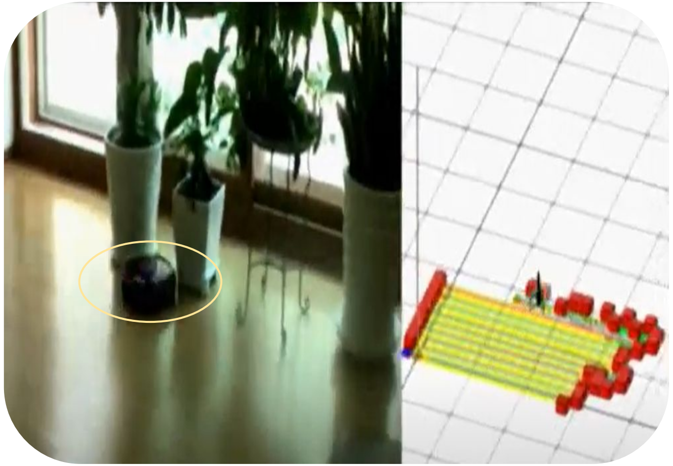
Mapping:

Mapping refers to **building a virtual representation** or map **of the environment** in real time. The device **scans and analyzes features in the physical space**, such as walls, furniture, and floors, to create this map. This map is crucial for **accurately placing** and anchoring virtual objects in the real world.



Visual odometry

- Determines **camera pose and movement** by analyzing **image sequences**.
- It's a model-free method that uses 3D reconstruction for **continuous 6DOF tracking**.
- Basic steps include detecting interest points, tracking them, calculating the essential matrix, recovering camera pose, triangulating 3D points, and repeating for subsequent frames.



Vacuum robot mapping the room

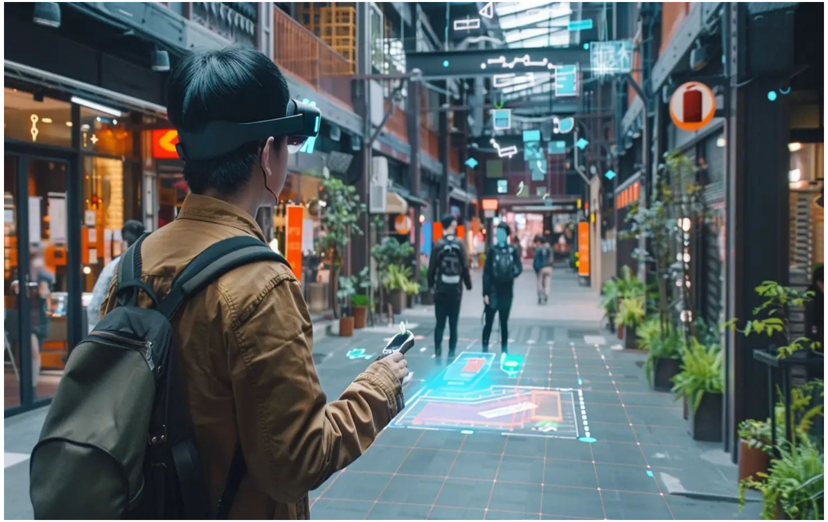
Advantages and limitations of markerless tracking

Advantages	Limitations
• No need for physical markers. • Works in dynamic environments. • Provides a natural user experience. • Low setup costs.	• Struggles in featureless spaces. • Affected by poor lighting. • High computational demands. • May lose accuracy over time.

Outdoor Tracking in AR

Outdoor tracking in augmented reality (AR) involves determining a user's location and orientation in the real world and overlaying virtual content on top of it.

Outdoor tracking in mobile AR applications primarily relies on a combination of **GPS**, **inertial tracking**, and **vision-based techniques** to create an accurate and immersive augmented reality experience.



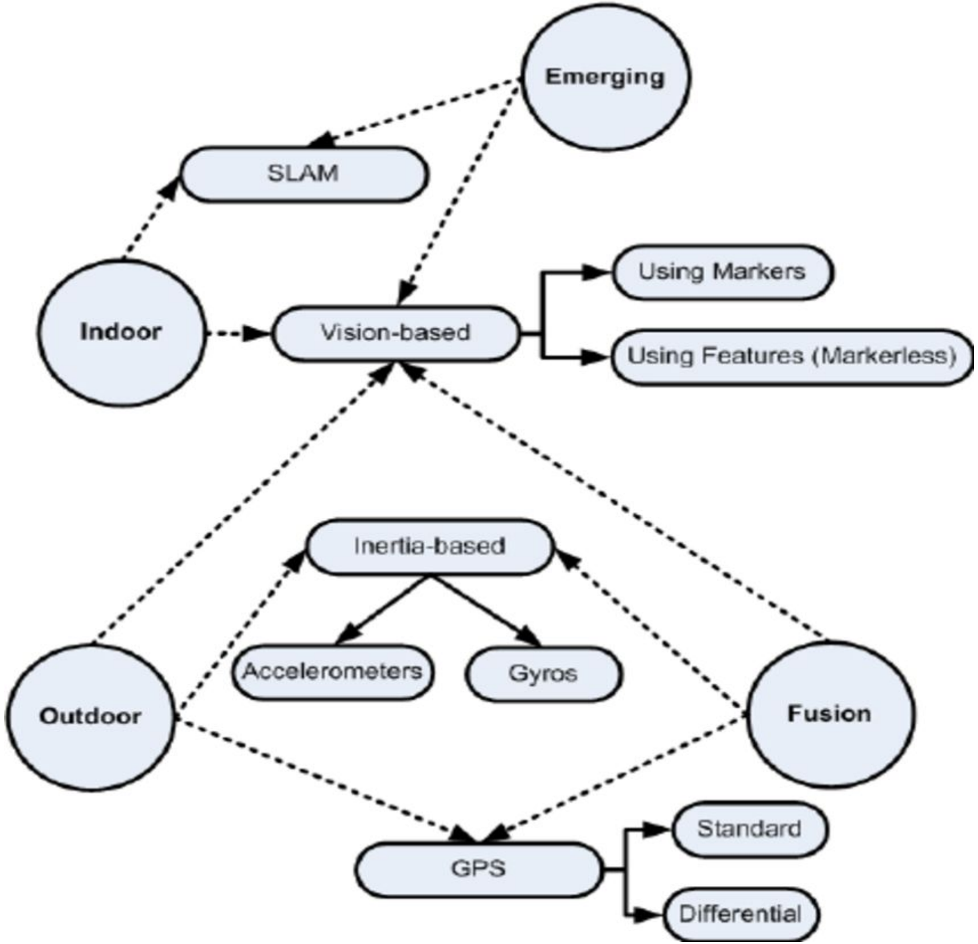
1. Indoor environments are more predictable, while outdoor environments are limitless in terms of location and orientation.

2. GPS, including differential GPS and a compass, was used for outdoor position and orientation tracking.

3. A database of latitudes, longitudes, and images captured in varying light conditions was created for outdoor AR systems.

4. Video tracking used reference images, with a matching score calculated through 2D transformations like rotation and translation.

5. The matching technique, based on Fourier Transformation, allowed for real-time tracking but was limited to a 10Hz refresh rate.


Emerging
SLAM
Indoor
Vision-based
Using Markers
Using Features (Markerless)
Inertia-based
Accelerometers
Gyros
Outdoor
Fusion
GPS
Standard
Differential

Combination of Techniques

GPS

Provides coarse location information, but is unreliable in urban areas or indoors.

Ex: Pokemon GO

Inertial Sensors

Track device movement and orientation using accelerometers and gyroscopes, offering accurate short-term positioning.

Ex: Google Map Live View

Vision-Based Techniques

Use cameras to detect and track features in the environment, improving accuracy and providing detailed positioning data.

Ex: Nantics AR platform

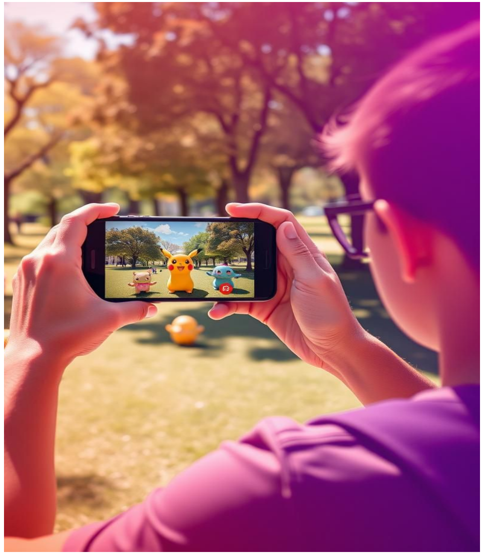
Use in Mobile Apps

Mobile AR applications often combine all three tracking methods (GPS, inertial tracking, and vision-based techniques) to deliver an experience where virtual elements appear as if they're seamlessly integrated into the real world. The combination compensates for the limitations of each method, improving accuracy and user immersion.

Use in Games AR Apps Like Pokémon GO

A popular AR game that relies on accurate outdoor tracking for placing virtual creatures in the real world.

Players explore their surroundings, encounter virtual creatures, and collect items, all within the real world.Players explore their surroundings to find and capture Pokémon at specific locations, which are mapped out using the game's GPS tracking system.



Challenges

- **Accuracy**

  Maintaining precise location and orientation is critical for a seamless AR experience.

- **Scalability**

  Outdoor tracking solutions should work reliably across a wide range of devices and environments.

- **GPS Limitations**

  GPS signals can be weak or unavailable in urban canyons, indoor spaces.

Visual Tracking

Visual tracking in Augmented Reality (AR) systems refers to the technology and methods used to determine the position and orientation of objects using **predefined visual markers** or specific visual elements that have been intentionally placed in the environment. These can be QR codes, fiducial markers, or any other identifiable graphics.

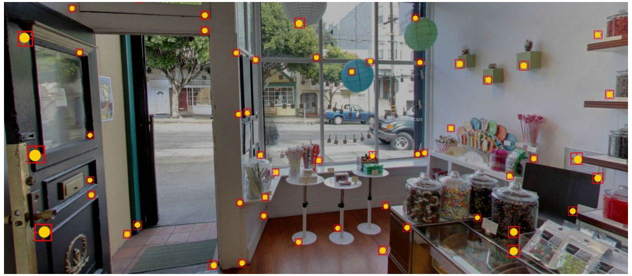
The AR system detects and recognizes these markers, using them as reference points to accurately determine the device's position and orientation. Because the markers are known and fixed, the system can achieve high precision and stability, but it requires prior setup, hence this method of tracking is more **static in nature**.

Feature-Based Tracking

Use of feature-Based Tracking

Feature-based tracking in AR systems is a method that relies on **detecting and tracking natural features** in the environment, such as edges, corners, textures, and other distinct visual points. Unlike marker-based approaches like visual tracking, it doesn't require predefined markers; instead, it uses these natural features to understand the device's position and orientation in real time. This flexibility allows for more **dynamic AR experiences.**

It is generally used in **markerless AR** systems, which makes them more flexible.A good example would be the SLAM (Simultaneous Localization and Mapping) system and various object recognition techniques.



Hybrid Tracking

Hybrid tracking in augmented reality (AR) refers to the use of multiple tracking methods or technologies in combination to enhance the overall performance of AR systems.

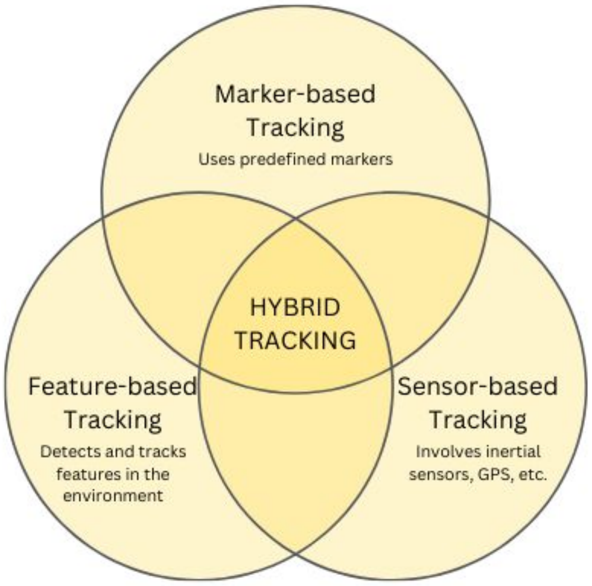
Enhanced Robustness: Combines methods to improve reliability in diverse conditions.

Seamless Transitions: Enables smooth shifts between tracking methods, minimizing disruptions.

Improved Accuracy: Merges tracking sources for precise, stable experiences.

Wider Applicability: Suitable for both indoor and outdoor AR applications.

Reduced Tracking Loss: Minimizes tracking loss, keeping AR content anchored to the real world.



Initialization and Recovery in Hybrid Tracking

Initialization Strategies:

Process of starting the tracking of a real-world object or environment.

- **Feature Initialization**: Detects and tracks distinctive environmental features to estimate camera pose.
- **GPS Initialization**: Uses GPS data for initial location and orientation, especially outdoors.
- **SLAM Initialization**: Builds a map and estimates camera pose simultaneously in unknown environments.
- **Inertial Initialization**: Utilizes accelerometers and gyroscopes for initial orientation, refined by visual tracking.

Recovery Strategies:

Recovery strategies are essential for regaining tracking when it is lost or disrupted, due to occlusion, abrupt camera movements, or changes in lighting conditions.

- **Re-Initialization**: Restarts tracking using initialization strategies when tracking is lost.
- **Sensor Fusion**: Combines data from multiple sensors to maintain accurate tracking.
- **Predictive Tracking**: Uses past movements to estimate future positions for smoother recovery.
- **Visual SLAM Loop Closure**: Recognizes previously mapped areas to recover tracking after a loss.

Applications of AR Tracking

Real-time tracking in surgeries and medical procedures.

Example: **AccuVein** – an AR device that helps doctors locate veins using a combination of optical and inertial tracking.

Virtual try-ons for clothing and makeup.

Example: **IKEA Place** – Uses SLAM to allow users to visualize furniture in their homes, adjusting to real-world space and dimensions.

Applications of AR Tracking

Augmented gaming experiences that blend virtual elements with real-world environments.

Example: **Pokémon GO** – GPS and camera-based tracking to locate and interact with virtual Pokémon in physical spaces.
Eg: a combination of optical and inertial tracking.

Enhanced learning experiences with interactive content overlaid in real environments.

Example: **zSpace** – Combines visual tracking and stereoscopic displays to create an immersive educational environment.