COMPUTER VISION FOR AUGMENTED REALITY-

What is computer vision?

Computer vision AR focuses on enabling machines to interpret and understand the visual world, much like humans do. It involves developing algorithms and techniques that allow users to process and visualize images to process the virtual environment, extract meaningful insights, and make informed decisions based on visual data.

For instance, consider an example of a self-driving car equipped with computer vision AR technology. As the car navigates through the busy roads of an urban city, its cameras capture a continuous stream of images.

The computer vision system processes these images in real time to identify various objects and make sense of the road environment. It can detect pedestrians crossing the street, recognize traffic signs and signals, differentiate between cars and motorcycles, and assess the distance and relative speed of nearby vehicles. In this scenario, computer vision is the technology that allows self-driving cars to ‘’see’’ and ‘’understand’’ the road scene.

The role of computer vision in AR

Computer vision plays a transformative role in AR, as it forms the foundation of how digital content is seamlessly integrated into the real-world environment. Below we have listed a few ways in which computer vision is contributing to the field of AR.

1)Object recognition: Computer vision in augmented reality serves a pivotal role by identifying and detecting real-world objects, seamlessly integrating virtual elements into them. This process, known as object detection, is a fundamental aspect of crafting immersive and lifelike AR experiences. The procedure unfolds as cameras capture images or videos of the user’s surroundings, which are then swiftly scrutinized by real-time computer vision algorithms.

These sophisticated algorithms harness techniques such as edge detection, pattern recognition, and machine learning to discern objects based on their visual attributes — comprising size, form, and hue. Once an object is successfully identified, AR systems can overlay virtual content onto it. This content might encompass diverse elements, from dynamic 3D models and animations to informative textual displays, enriching the user's perception of reality with digital enhancements.

2) Localization and mapping (SLAM): Augmented reality computer vision plays a pivotal role in simultaneous localization and mapping (SLAM), a technique that enables AR devices to understand their own position and orientation within the environment in real-time. This is essential for ensuring that virtual content aligns correctly with the user's perspective as they move around. SLAM helps AR experiences remain stable and accurate, even in dynamic environments.

For instance, you are wearing a headset in a museum. As you move, the headset uses computer vision, the technology behind augmented reality, to understand where you are and what you are looking at. This process, called SLAM, helps the headset create a map of the museum and know its own position. When virtual content, like a guide, appears next to an exhibit, SLAM ensures that the virtual content stays in the right place even as you move around. This makes the AR experience stable and accurate as if the virtual and real worlds are perfectly connected.

3) Gesture and motion recognition: Computer vision allows AR systems to track user gestures, facial expressions, and body movements. This tracking enables users to interact naturally with virtual content. For instance, users can manipulate virtual objects using hand gestures, enhancing the interactive aspect of augmented reality technology.

4) Depth estimation: Computer vision assists in estimating the depth of objects in the environment, enabling AR devices to understand the three-dimensional layout of the scene. This depth information is crucial for placing virtual objects at the correct distances, ensuring they appear realistically integrated into the physical space.

5) Environmental adaptation: In the world of Augmented Reality technology, making virtual objects feel like they belong in the real world is all about adapting to the environment. This is where computer vision comes into play. Imagine seeing virtual objects that react to changes in lighting, behave like real ones in reflections, and cast lifelike shadows.

<u>MARKER TRACKING -</u>

What is marker-based AR?

Marker-based Augmented Reality is an approach where images, referred to as markers or triggers, serve as reference points for superimposing digital content onto the physical environment. The camera recognizes these markers with the help of computer vision and tracking features. It then uses them as anchors to overlay virtual elements, such as 3D models, animations, or information, onto the corresponding marker locations.

How does it work?

It operates through the synergistic efforts of cameras, computer vision algorithms, and the Augmented Reality marker image itself, following a sequence of fundamental steps:

Marker Detection

The camera continuously scans the physical environment, actively searching for predefined markers. These markers are typically high-contrast patterns or designs that are easily recognizable by the system.

Pose Estimation

Once a marker is successfully detected, the computer vision techniques calculate the precise position and orientation of the marker relative to the camera's viewpoint. This spatial relationship is crucial for accurate content alignment.

Content Overlay

With the marker's pose accurately determined, it then seamlessly overlays digital content onto the physical marker in real-time. This augmented content can take various forms, such as 3D models, videos, animations, or textual information, effectively blending the virtual and real worlds.

What are its widespread applications?

Major applications of marker-based AR are in the fields like:

Education

Marker-based AR can transform traditional educational materials into immersive and interactive learning experiences. For instance, biology textbooks can come alive when viewed through a smartphone or tablet, with 3D models of organisms superimposed onto designated markers. Students can explore these virtual representations from multiple angles, gaining a deeper comprehension of complex structures and processes, ultimately enhancing their understanding of the subject matter.

Marketing and Advertising

Brands leverage it as an innovative way to engage consumers and promote their products. By incorporating markers on product packaging, companies can unlock additional digital content, such as promotional videos, interactive experiences, or even virtual try-on experiences, when scanned with a dedicated mobile application. This approach captivates the audience and provides a unique and memorable brand experience.

Gaming and Entertainment

It can be integrated into games and entertainment experiences, creating interactive and immersive environments. For example, board games or card games can come alive with virtual characters, animations, or special effects when viewed through a compatible AR application, adding an extra layer of excitement and engagement for players.

Retail and E-commerce

Markers can be used in retail settings or on product catalogs to provide customers with additional information, virtual try-on experiences, or even augmented reality product demonstrations. This approach can enhance the shopping experience, aid in purchase decisions, and potentially increase sales.

Architecture and Construction

Architects and construction professionals can leverage it to visualize proposed designs or modifications on-site. By placing markers at specific locations, they can overlay virtual representations of buildings, structures, or landscaping elements, allowing for better communication and collaboration with clients or stakeholders.

Benefits of Marker based AR

Some well-used benefits include:

Amplifies branding, marketing, and advertising content

You could be adding oomph to your marketing collateral and branding content beyond just plain images or videos on the internet with **marker-based web AR** that users can access on the browser without installation.

Increases customer interaction

Interaction between customers and the brand is extremely important, and this is not possible with plain images or videos. With marker-based AR, customer interaction with the products and brands leaves a lasting impression.

Embed interactive AR content

With AR markers, interactive content can be easily embedded in materials related to advertisement and customer engagement. It can vary from an immersive experience to a product demonstration.

Types of AR marker-based content

Some of the most used examples of how market-based AR is used in the market are:

QR codes

QR codes have emerged as a highly effective and widely preferred tool for triggering AR. It is an emerging tactic for marketing campaigns. With their ease of use and accessibility, these codes have become immensely popular. Brands of all sizes and across diverse industries use them in many ways. Also, they offer a seamless and convenient way to create anchors for AR content, helping brands engage with their customers in a more interactive and immersive manner.

Logos and brand visuals as Markers

Logos are a great example of how to create an anchor. They represent the brand and can become an asset in creating AR experiences. By using materials related to the brand, such as logos and other brand elements, companies can create an immersive experience. It helps keep users engaged and invested in the brand. This approach helps to build brand recognition and loyalty. Users are more likely to remember a brand that offers engaging and unique experiences.

Smart packaging

Smart packaging has become increasingly popular in the market due to its versatility. You can embed AR experiences on product packaging. Users can access it simply through their smartphone by scanning the package. Smart packaging is also easy to manufacture and distribute and is less expensive than most traditional immersive experience triggers.

2.4.5. Multiple Camera Infrared Tracking

In general, the known points in the world will not be constrained to a plane, as assumed in the previous section on tracking of flat markers. For tracking arbitrary objects, we require general pose estimation, which addresses the problem of determining the camera pose from 2D-3D correspondences between known points q, in world coordinates and their projections p, in image coordinates.

In this section, we describe a simple infrared tracking system designed to track rigid body markers composed of four or more retro-reflective spheres (an approach introduced in Chapter 3). It uses an outside-in setup with multiple infrared cameras [Dorfmüller 1999]. A minimum of two cameras in a known configuration a calibrated stereo camera rig is required. With this strategy, the additional input and wider coverage of the scene from multiple viewing angles will improve the tracking quality and the working volume. In practice, four cameras set up in the corners of a laboratory space are a popular configuration. Use of more than two cameras will improve the performance of the system, but is not fundamentally different from the stereo case.

The stereo camera tracking pipeline consists of the following steps:

1. Blob detection in all images to locate the spheres of the rigid body markers

2. Establishment of point correspondences between blobs using epipolar geometry between the cameras

3. Triangulation to obtain 3D candidate points from the multiple 2D points

4. Matching of 3D candidate points to 3D target points

5. Determination of the target's pose using absolute orientation (as described, for example, by Horn [1987] and Umeyama [1991])

2.4.6. Natural Feature Tracking by Detection

- Unlike AR solutions that use markers as their basis for recognition, natural-feature tracking solutions can be applied to almost any image as long as the image is complex enough.

- An example of a natural-feature tracking application is a mobile application that can recognize a movie poster.

- With natural-feature tracking, the application can analyze the poster and identify it by comparing the poster image to similar images. In contrast, a marker-based solution requires a special identifier to be included on the poster; it would be the marker that provides the identification rather than the poster image.

2.4.8. Simultaneous Localization and Tracking

The simplest form of model-free tracking, which can be seen as a precursor to simultaneous localization and mapping (SLAM), is sometimes called visual odometry. In a nutshell, visual odometry means continuous 6DOF tracking of a camera pose relative to an arbitrary starting point. This approach originally comes from the field of mobile robotics. Visual odometry computes a 3D reconstruction of the environment, but uses it just to support the incremental tracking. A basic visual odometry pipeline encompasses the following steps:

1. Detect interest points in the first frame.

2. Track the interest point in 2D from the previous frame

3. Determine the essential matrix between the current and previous frames from the feature correspondences with a five-point algorithm

4. Recover the incremental camera pose from the essential matrix.

5. Since the essential matrix determines the translation part of the pose only up to scale, this scale must be estimated separately, so that it is consistent throughout the tracked image sequence. To achieve this aim, 3D point locations are triangulated from multiple 3D observations of the same image feature over time (see "Triangulation from More Than Two Cameras"). This approach is called structure from motion (SFM).

6. Proceed to the next frame.

2.4.9. Outdoor Tracking

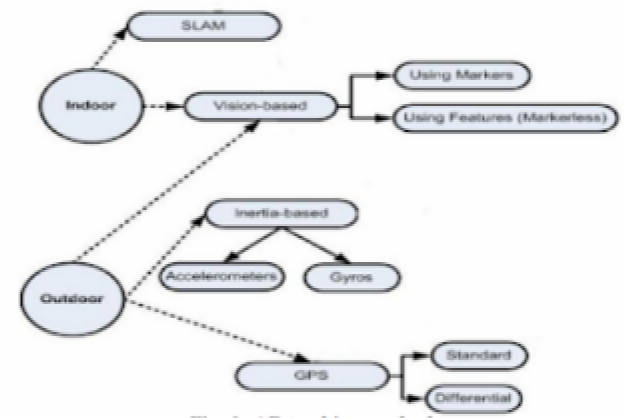


Fig. 1. AR tracking methods.

- Indoor environments are usually more predictable whereas the outdoor environments are limitless in terms of location and orientation.
- As stated earlier, GPS is a good tracking option when working outdoors.
- A differential GPS and a compass was used for position and orientation judgement. Latitudes and longitudes of several viewpoints were collected in a database along with the set of images captured at different times of the year with varying light conditions.
- Reference images were utilized for video tracking and matching was performed to discover these reference images for the outdoor AR system.
- A video image was examined with the reference images and a matching score was achieved. For the finest matching score, the 2D transformation was measured and the current camera position and orientation were deducted.
- This transformation was utilized to register the model on the video frame. The matching technique was based on Fourier Transformation to be robust against variation in lighting conditions hence it was limited to only 2D transformations like rotation and translation.
- This technique had a fixed number of computations therefore it was appropriate for real-time operation without using markers, it yet worked on 10Hz which is a low rate for real-time display.

SOFTWARE USED TO CREATE CONTENT OF AR APPLIACTION

There are several software options for creating augmented reality (AR) applications, including:

ARCore

A platform for building AR apps on mobile devices, with a Cloud Anchors API that allows multiple users to place virtual content in the same real-world location

Vuforia

An AR software development kit (SDK) that uses computer vision to detect and track images and 3D objects

ARKit

A tool with image detection and tracking functionality that allows apps to anchor virtual content on real-world surfaces

Wikitude

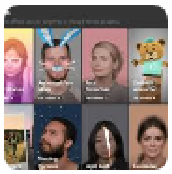
A tool for developing location-centric AR experiences, with an SDK that allows developers to implement geolocation features, image tracking, and object recognition

Unity

A platform for creating AR and virtual reality (VR) applications, with high-fidelity visuals, realistic audio, and intuitive controls

Kudan

An SDK that uses SLAM to recognize 3D objects and pictures, and is easy to integrate



EasyAR

An affordable platform for developing AR apps, with a free SDK that includes multi-target simultaneous detection and tracking

Unreal Engine

A game engine that offers advanced graphics capabilities for creating content for AR, VR, and MR

AR Foundation

A set of tools from Unity that allows you to build AR apps in a multi-platform way, with features such as object tracking, plane detection, and image tracking