Lecture 10

CONTENTS

1. Augmented Reality Hardware – Displays – Audio Displays, Haptic Displays, Visual Displays, Other sensory displays, Visual Perception , Requirements and Characteristics, Spatial Display Model. Processors – Role of Processors, Processor System Architecture, Processor Specifications. Tracking & Sensors - Tracking, Calibration, and Registration, Characteristics of Tracking Technology, Stationary Tracking Systems, Mobile Sensors, Optical Tracking, Sensor Fusion.

02

Tracking :

- Every AR application has to detect and calculate the position and orientation of the user's head (or the camera) relative to the entities in the real world scene continuously. Additionally other parts of body may be tracked e.g., hands – to allow interaction, chest or legs. Three-dimensional objects have six degrees of freedom (DOF): position coordinates (x, y and z offsets) and orientation (yaw, pitch and roll angles for example). Tracking is significantly more difficult in AR than in VR because greater precision is required and the latency can not be tolerated. In AR, without accurate tracking the virtual objects will not be drawn in the correct location and the correct time, ruining the illusion that they coexist with the real objects.

- The most important properties of 6DOF trackers, to be considered for choosing the right device for the given application are :

- **Update rate** – defines how many measurements per second (measured in Hz) are made. Higher update rate values support smoother tracking of movements, but require more processing. The tracking rate must be not less than 10 frames per second to ensure accurate augmented scene.

- **Latency** – the amount of time (usually measured in ms) between the user‘s real (physical) action and the beginning of transmission of the report that represents this action. Lower values contribute to better performance.

- **Accuracy** – the measure of error in the reported position and orientation. Defined generally in absolute values (e.g., in mm for position, or in degrees for orientation). Smaller values mean better accuracy.

- **Range** – working volume, within which the tracker can measure position and orientation with its specified accuracy and resolution, and the angular coverage of the tracker

Sensor-based Tracking :

Sensor-based methods use not only the cameras, but also special sensors to determine the relationships between the viewer (or the camera), the display and the real world.

Magnetic tracker

A magnetic tracker is a position measurement device that uses a magnetic field produced by a stationary transmitter to determine the real-time position of a moving receiver element.

It consists of: a static part (emitter, sometimes called a source), a number of movable parts (receivers, sometimes called sensors), and a control station unit.

The assembly of emitter and receiver is very similar: they both consist of three mutually perpendicular antennae.

As the antennae of the emitter are provided with current, they generate magnetic fields that are picked up by the antennae of the receiver. The receiver sends its measurements to the control unit that calculates position and orientation of the given sensor.

There are two kinds of magnetic trackers that use either alternating current (AC) or direct current (DC) to generate magnetic fields as the communication medium

Acoustic (ultrasonic) tracker

An ultrasonic tracker is a position measurement device that uses an ultrasonic signal produced by a stationary transmitter to determine the real-time position of a moving receiver element.

Ultrasonic trackers have three components, a transmitter, a receiver, and an electronic unit, similar to their magnetic counterparts.

The difference is that the transmitter is a set of three ultrasonic speakers and the receiver is a set of three microphones.

Due to their simplicity, ultrasonic trackers represent a cheaper alternative to the magnetic ones.

Contd.

Mechanical tracker

A mechanical tracker consists of a serial or parallel kinematic structure composed of links interconnected using sensorized joints.

The dimensions of each link segment are known a priori and used by the direct kinematics computational models stored in the computer.

This model allows the determination of the position and orientation of one end of the mechanical tracker relative to the other, based on the real-time reading of the tracker joint sensors.

By attaching one end of the tracker to the desk or floor and the other to an object, the computer can track the object's 3D position relative to the fixed end of the arm.

Inertial tracker

Inertial tracker is a self-contained sensor that measures the rate of change in an object orientation. It may also measure the rate of change of an object translation velocity.

Inertial sensors have no range limitations, no line-of sight requirements, and no risk of interference from any magnetic or acoustic interference sources.

They can be sampled as fast as desired, and provide relatively highb and width motion measurement with negligible latency.

Even tiny low-cost MEMS inertial sensors measure motion with very low noise resulting in jitter-free tracking which looks very smooth to the eye.

Contd.

Optical tracker

An optical tracker is a non contact position measurement device that uses optical sensing to determine the real-time position/orientation of an object.

Similar to ultrasonic trackers, optical trackers work through triangulation, require direct line-of-sight, and are immune to metal interference. Optical trackers, however, offer significant advantages over their ultrasonic counterparts.

Their update rates are much higher and their latency is smaller than those of ultrasonic because light (whether visible or infrared) travels much faster than sound.

They are also capable of (much) larger work envelopers. If the tracker sensing component (charge-coupled device [CCD] camera, photodiode or other photo sensor) is fixed and some light beacons are placed on the user the tracker is said to be outside-looking-in.

By contrast, an inside-looking-out has the camera(s) attached to the tracked object or user.

Vision-based Tracking

Sensor-based methods do lack in either accuracy, robustness, range, drift or noise sensitivity.

Furthermore, there are practical limitations on the use of sensors, such as a user‘s limited area of movement and perturbation caused by the environment.

Computer vision has the potential to yield noninvasive, accurate, and low cost solutions without the need of any additional sensors. The relationship between the camera position, the real world and the virtual computer-generated objects is computed directly from the image sequences.

Different methods with variable degrees of accuracy and robustness can be applied.

Vision-based Tracking :

Sensor-based methods do lack in either accuracy, robustness, range, drift or noise sensitivity. Furthermore, there are practical limitations on the use of sensors, such as a user‘s limited area of movement and perturbation caused by the environment. Computer vision has the potential to yield noninvasive, accurate, and low cost solutions without the need of any additional sensors. The relationship between the camera position, the real world and the virtual computer-generated objects is computed directly from the image sequences. Different methods with variable degrees of accuracy and robustness can be applied.

Marker-based tracking

Camera tracking systems based on placing fiducial markers in the scene have been highly successful. Markers are constructed so that they are easily detected in each image frame and given some a priori information about the shapes or positions of the markers, the relative pose of the camera can be easily determined. This assumes that one or more fiducials are visible at all times.

Otherwise, the registration cannot be done. A drawback of this method is that it is not always possible to place fiducials. In fact, AR end users do not like them because it is not always possible to modify the environment before running the application.

The addition in the scene of fiducials, also called landmarks or markers, greatly helps accomplish two main tasks: extracting information from the image and estimating the pose of the tracked objects.

Markers constitute image features that are easy to extract. They also provide reliable, easy to exploit measurements for the pose estimation.

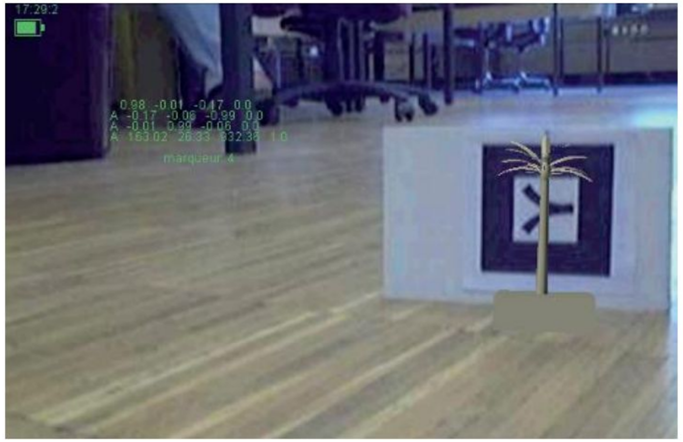
There are two types of fiducials: —

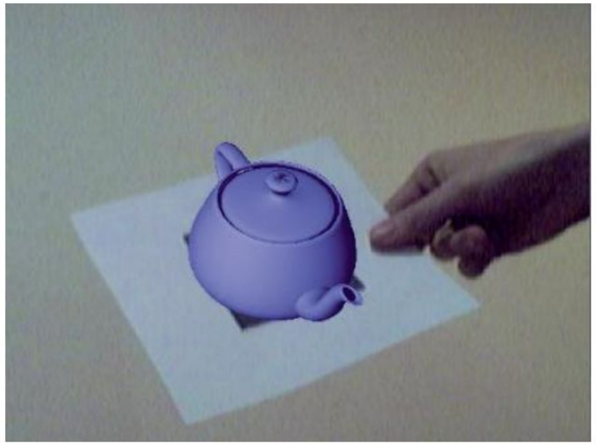
- point fiducials‖ and
- planar fiducials.

Point fiducials are commonly circular markers because the appearance of circular patterns is relatively invariant under perspective distortion, facilitating their detection. Point fiducials give one-point correspondence between the scene and the image To facilitate their identification, the markers can be arranged in a distinctive geometric pattern.

Planar rectangular fiducials have gained popularity because they yield a robust, low-cost solution for real-time 3D tracking.

Contd.





Marker-based tracking using ARToolkit

Marker-less (natural feature) tracking

In some applications, it is either undesirable or very difficult to add markers to the scene, for example, outdoor environments or unknown environments. In such cases it is necessary to rely on naturally present object features such as points, corners, edges, or texture. Systems based on natural features extend the tracking range and are typically more stable as there are more features available to track the camera pose from. However, this increases the complexity of the tracking algorithm. Finding and following feature points or edges can be difficult because they are hard to detect, and in many cases there are not enough of them on typical objects. Features may come from the background or from unknown objects.

Total or even partial occlusion of the tracked objects typically results in tracking failure. The camera can easily move too fast so that the images are motion-blurred; the lighting during a shot can change significantly; reflections and specularities may confuse the tracker. It should also be considered that an object may drastically change its aspect due to displacement, for example, changes in perspective or appearance of hidden parts when changing the point of view. In such cases, the features to be followed always change, and the tracker must deal with features coming in and out of the picture.

Contd.

Marker-less (natural feature) tracking (Contd.)

We can decide which tracker we will use depending on not only the nature and requirements of the application, but also the environment in which it will be performed. For indoor applications, since the dimensions of the environment are predefined (e.g. a laboratory room), physical restrictions on user‘s movements and location can be conveniently determined by the AR system. In addition, an indoor environment can be deemed “prepared“ by placing sensors, optical markers, or location tracking cameras at important locations. Unlike indoor environments, a user operating outdoors has an unlimited number of possible locations and orientations. Tracking in unprepared environments is difficult for three reasons.

First, if the user operates outdoors and traverses long distances, the resources available to the system may be limited due to mobility constraints (size, weight, and power).

Second, lighting conditions, weather, and temperature are all factors to consider in unprepared environments. Third and most importantly, the system designer cannot control the environment. It may not be possible to modify the environment, such as placing special fiducial markers at known locations in the environment.

Registration :

Registration is the process of rendering virtual objects onto the user‘s view based on the values calculated from the tracking process. Visually the real-time constraint is manifested in the user viewing an augmented image in which the virtual parts are rendered without any visible jumps. To appear without any jumps, a standard rule of thumb is that the graphics system must be able to render the virtual scene at least 10 times per second. Although tracking technology is a primary concern to the registration issue in AR, data latency can also bring about unexpected registration problems.

Registration approaches are named depending on the name of the tracking approaches used.

So we can say that registration methods are:

- sensor-based registration and
- vision-based registration.

Interaction :

While great advances have been made in AR display technologies and tracking techniques, few systems provide tools that let the user interact with and modify AR content. AR technology typically includes input devices that enable users to control and interact with the synthetic world. Input devices can be as simple as keyboards or as complex as tracking systems that are used to determine the location and orientation of real objects in the environment.

In fact, writing this part was a difficult mission because there was no clear and specific categorization of the recent input devices. There were three different classifications have been presented. Hiroshi et. al. classified input interfaces for wearable computers into contact and non-contact devices. Their classification was limited only to wearable computers. Sriram and Wijnand presented an excellent classification of spatial object manipulation techniques.

This classification considered 3D spatial contact devices only and didn't consider 2D input devices or non-contact spatial techniques. Also, Tomasz and Michael, who classified input devices into 3D and desktop input devices didn't mention non-contact spatial devices. Furthermore, the three papers are very old. So the recent technologies were not considered in them. As a result, we will present a new classification which integrates between the previous classifications and consider the most recent interaction technologies. We have classified interaction interfaces mainly into 2D (or non-spatial) and 3D (or spatial) interfaces.

Calibration :

To represent virtual objects properly in the real scene, AR systems should calibrate the camera with intrinsic and extrinsic parameters.

User would certainly get very disoriented if virtual objects would be floating in front of real objects but yet they would feel to be further away. Faulty camera parameters also undermine chances for successful occlusion detection. Camera calibration involves the numerical calculation of both types of parameters for each camera.

1. Intrinsic parameters :

   The intrinsic parameters are about how the camera will convert objects within the camera's field of view into an image, and they are independent on position and orientation. The intrinsic parameters are related to the internal geometry of the camera:

   for example, the focal length, the center position in the pixel image, and the pixel size of the resolution.

1. Extrinsic parameters :

   The extrinsic parameters are related to the external properties of the camera, such as camera position and camera orientation in space. These parameters uniquely identify the transformation between the unknown camera coordinate system and the known world coordinate system.

   Using these extrinsic parameters, we can define a transformation matrix, which consists of a 3x3 rotation matrix and a 3D translation vector

Thank You