Eigene Ausarbeitung der Prüfungsfragen. Großteils in Englisch geschrieben und aus den Vorlesungsfolien, früheren Ausarbeitungen und eigenen Recherchen zusammengestellt. Übernehme keine Verantwortung für Richtigkeit oder Vollständigkeit, habe bei der Prüfung gut abgeschnitten.

Unterschied zwischen optischen und magnetischen Trackern, bzw. beide Techniken erklären

• Electric current in a coil creates a magnetic field.
• Magnetic Flux (magnetischer fluss) is created in receiver
• Magnetic flux: Function of distance and orientation relative to coil
• To measure position and orientation of reciever in space, emitters consist of 3 orthogonal coils and recievers of 3 sensors -> combination of three elementary orthogonal directions

Advantages:
• No line of sight restrictions – continuous data
• small, leightweight sensors
• wireless
• robust, off the shelf availability
• High update rate (200 Hz or more)
• relatively inexpensive

Disadvantages:
• External noise
• Limited field strength
• limited working volume and mediocre accuracy

Cybersickness
• Cybersickness is a form of motion sickness present when people interact with virtual environments.
• Bad correspondence of virtual and actual positions in AR (inaccurate tracking, bad registration) can lead to cybersickness.
• Extended use of HMDs (max. 30-60min) leads to cybersickness
• Up to 95% of regular VR users are affected by it

Was sind die Gründe für ihr Auftreten und welche Symptome hat Cyber Sickness?

Technische Fehler wie Tracker Errors (Unterschiede zwischen Benutzerbewegungen und Avatarbewegungen), großer oder geringe HMDAuflosung und kleiner FOV (wobei auch sehr große FOVs problematisch sein konnen) konnen das „Erkranken“ an der Cybersickness verursachen.
Auch das Alter, der Gesundheitszustand und die generelle Anfalligkeit kann ausschlaggebend sein.

Unterschiede zwischen realen positionen und bewegungen und der virtuellen darstellung. akkomodation des auges auf falsche distanzen (conflicting stereo cues) etc.

Cybersickness symptoms include: nausea, vomiting(severe cases), disorientation, vertigo, eye strain and dizziness.

Was ist Presence? Warum ist sie wichtig?

The feeling of really being at the location that the displays specify for the user, a “place illusion”
Let’s the participant experience objects and processes of the virtual world in a way that they seem to present reality more than the real world in which the VE is embedded.
-> Participant will behave in the virtual environment similarly to how he/she would behave in reality
VE should be a visiting place, not just about seeing images

Measurement:
- Questionnaires
- Behavioural measures
- Physiological measures: social phobia

- Biofeedback measures: heart rhythm, blood pressure, skin conductivity,…
- Conflicting signals (wrong shadows) -> measuring BIPs (Breaks in presence)

CAVE: excellent results (few BIPs, no BIPs when engaging)

Crossfire und SLI erklären und die Modi
Multiple boards with multi GPUs connected with a bridge

Performance increase by maximum of 1.8x

Two modes:
- Split Frame Rendering (SFR) - Scissors: Splits each frame and sends half the load to each of the graphics cards
- Alternate Frame Rendering (AFR): Frame 1 – Card 1, Frame 2 – Card 2, alternating

Marker Tracking:

Rectangular markers provide 4 corner points –> enough for pose estimation
Lots of open source solutions available, fast, and has been used for over 10 years
Steps:
- Get camera frame
- Binarisation of image using threshold
- Mark connected regions (check min/max area size)
- Undistort
- Create contour
- Fiducial detection
- Rectangle fitting
- Template matching using the downscaled undistorted greyscale pattern within the rectangle borders
- Render the virtual object matching the pattern, based on the pose (orientation, position)

Problems: lighting, optical noise, ambiguity, low resolution
Solution: Infrared light camera (works also in dark environments)
Active (emits light) and passive (reflect light) markers

Wie funktioniert Wii Controller?

- 3-axis accelerometer
- Optical infrared sensor
  - Uses IR-LEDs on sensorbar for recalibration
  - Distance to sensorbar is calculated
- Rumble functionality

Erkläre passives und aktives Stereo. Was sind die Vor- und Nachteile? Welche Auto-stereoskopische Display Technologien gibt es? Beschreibe diese kurz.

Active Stereo:
- Active switching between left and right eye, e.g. using shutter glasses

Passive Stereo:
- Anaglyph (red/green) stereo: colour filters on left and right eye let only certain colours get through – left and right images are colour encoded.
- Polarised light (from projectors with filters) gets through only on the specific side of the glasses (linear +-45° - cant tilt head, or circular or elliptical polarisation) works because of polarisation filters on each side

• Infinitec – interference filter technology using specific wavelengths of r,g and b for each eye/glass

Aktives Stereo: Shutter-glasses (CAVE, Projection Walls, 3D TV, 3D Monitor, cinema)
Passives Stereo: Anaglyph(red/green), polarised glasses (linear (+-45°), circular or elliptical polarisation), Infitec – interference filter technology using specific wavelengths of r,g and b for each eye/glass

LCD Shutterglasses: Inexpensive, eye sees only its "own" image (no ghosting),
Polarisation glasses: Cheap glasses but needs two projectors (more expensive than stereo-ready monitor/TV), if linear: can't tilt head but little ghosting, if circular polarisation: can tilt head but more ghosting
Infitec: Filter for projectors + glasses. Glasses quite expensive. Apparently works well though

Welche arten von Stereo gibt es? - Volumetric und Lentrikular

Passive-Active

Cave, Projection walls, monitor, display

Auto-stereoscopic displays:
• Lenticular: Requires no glasses, some offer up to 48 views others have head-tracking uses a grid of lenses -> directing pixel light at viewers. Good 3D only at sweet spot. Normally single-user, if multi-user: lower 3D quality
• Volumetric displays:
- Swept Volume: 3D scene is decomposed into slices, slices are projected under fast motion. The eye fuses the slices into a 3D volume
- Stack of planes: slices are behind each other and a parts of the images are projected onto each slice
- Volumetric 3D LED displays

Interaction: Volumetric displays are good for multi-user purposes and are invariant regarding view angles. The rear-side of objects can be viewed therefore as well. Lenticular lenses are better for single user interaction.

Wie schaut die Einkaufsliste aus um 15 Kindergartenkindern den neuesten 3d disney film zu zeigen

Nvidia 3D vision glasses, 3D Vision-ready monitor, Compatibly Nvidia card, PC with Windows Vista or newer Windows

Or 3D glasses + 3D TV

Wie kann die Microsoft Kinect ein Tiefenbild erzeugen?

Uses 3D Sensor by Primesense. Projects a pattern of infrared dots onto the scene. Uses randomised sub-patterns. The IR-camera detects the shift of the pattern in the image and calculates the depth from it. Heavy achine-learning technology required for "guessing" the pose.

Traversieren von Scenengraphen und wofür gann das Traversieren noch verwendet werden außer fürs rendern?

Scenegraph = data structure
-> All nodes gets traversed hierarchically = processed
-> Each nodes has methods implemented (rendering, calculations, file writing,event-handling, bounding box calculations, etc)

Ordered vs unordered: top-down & left-right traversal vs only top-down

Useful for transformation hierarchies, instancing, specific attributes (materials, texture, etc)

Is also used in Studierstube: During traversal each node executes the AR Framework code. The usage of the AR code is encapsulated in the nodes.

Erkläre die Manipulation

Universal Interaction Tasks:

- Selection: Picking objects from a set of objects
  Goals: - Indicate action on an object – Make object active – Travel to object location – Set up manipulation
- Manipulation: Manipulating object properties (such as position, orientation, shape, colour etc)
- Navigation: Travel (motor component) , Wayfinding (cognitive component, decision-making)
- System control: changing state or mode
- Symbolic input (keyboards, pens, gestures, sign language, speech)
- Modelling and other tasks

Goals:

- Object placement
  - Design
  - Layout
  - Grouping
- Tool usage
- Travel

User performance affected by:

- Required translation distance
- Required amount of rotation
- Required precision of placement

Metaphors:

- Virtual hand (natural, easy placement, limited reach, fatiguing, 1:1 mapping of position)
- Ray casting (little effort, exact position and orientation difficult)
- Indirect depth control (e.g.: mouse wheel, infinite reach and no tiring, not natural and separate DOFs)

CAVE, Aufbau etc.

Computer Assisted Virtual Environment.
3-6 screens are arranged around the user (e.g.: front wall and side walls) creating a virtual room. This creates immersion. Screens are connected to a single or multiple PCs (clusters)
Users can be tracked

Advantages

- high resolution, big fov (field of view)

- peripherical seeing is used
- can use shutter glasses for stereo
- user can move in the CAVE freely
- space for requisites
- virtual and real objects can be used
- multiple people can be inside the room at once

Disadvantage

- expensive, depending on the hardware used it costs at least some thousands of euros
- requires a lot of space
- projectors need to be calibrated regularly (lenses and bulbs can change position over time, so that there can be visible seems on the edges between the walls after a while)
- only 1-2 people can be tracked
- stereoscopic viewing can be difficult (requires omnistereoscopic rendering either via vertex-offset + tessellation or via a per-pixel solution or using rendering of vertical slices)
- No direct interaction such as with an HMD possible (e.g.: moving around an object)
- Real objects can occlude virtual ones

Erkläre den Begriff DOF. Was kann mit einem System, das 6 DOF hat berechnet werden?

Degrees of Freedom
Number of independent measurements
e.g.: 3-DOF can consist of x-,y-,and z-axis measurement = position
Mit einem 6-DOF system können 6 unabhängige werte gemessen warden zb. Rotation (roll, pitch, yaw) und Position (x,y,z)

Was sind die Charakteristika von HMDs? Was muss man beim Kauf eines HMDs beachten?

- HMD consist of either one or two displays plus special optics in front of the user's eyes
- Users do not see the real world
- Provides a stereoscopic view that moves with the users head.

Stereo transmission: Line-interleaved, side-by-side, over-under, two oculars, field-sequential

There are also see-through versions of HMDs:
- Optical see-through (Reflection on curved mirrors
- Video see-through

…which can be used for AR

Important characteristics:
- resolution (for each eye)
- FoV (field of view)
- head-tracker (none vs. 3-DOF vs. 6-DOF)

Prices vary a lot. Actual value of product might not be equal

Best product atm.: Oculus Rift (300 USD,
FOV: 110° diagonal, 90° vertical
Full HD - vertically split for left and right eye so 960x1080 per eye
220 grams
3DOF orientation + position tracking for consumer version confirmed)

Was sind die einzelnen Schritte in der Software, die ein optisches Tracking System mit retroreflektiven Markern durchläuft? (Optical Tracking System Design) Was sind hybride Tracker? Erkläre es anhand des Sony Move Game Controllers.

Example for a stunt-car: Use rigid constellations of retro-reflecting markers which are attached to the car
Cameras from different views record the car

Optical Tracking System Design

- Camera calibration (extrinsic and intrinsic parameters)
- Segmentation and feature identification: detecting blobs
  Detection using binary thresholding. Problem: overlapping blobs
- Feature correlation: finding correspondences between blobs in multiple views (using epipolar geometry)
- Projective reconstruction: acquire 3 DOF marker positions
  Rays do not intersect in reality, Least-squares approach → 3D point cloud
- Model-fitting: find pre-calibrated rigid constellations in the 3D marker cloud
- Pose estimation: obtain 6-DOF pose for each rigid constellation

Alternative to such a system:
Radio Frequency trackers: Triangulation similar to acoustic (ultrasound) tracking
Phase difference time of flight
transmitting signals over multiple bands of frequencies simultaneously
Active tags used : unique identifiers
LPM – Local position measurement by Albatec:
500x500m area, 1000Hz, +-5cm accuracy

Hybrid Trackers:
Combining multiple forms of tracking because no tracker is good at everything

Sony Move:

- Inertial sensor (gyro, accelerometer, magnetometer)
  - Accelerometers : Mass on spring in tube, measures relative navigation from a starting point, based on sensing the acceleration
  - Gyroscope: Angular velocity
  - Magnetometer: compass, earth's magnetic field -> 2DOF
  - → Altogether measuring the orientation
- 60Hz camera does optical tracking of the coloured sphere
  - High accuracy (cm/mm)
  - Different colours

Was ist statische und dynamische Genauigkeit? Beschreibe diese kurz.

Static accuracy:
Acccuracy without movement. Maximum deviation of fixed tracker position (sender) from fixed reference value
Dynamic accuracy:
Accuracy as sensor is being moved. Depends on static accuracy.

Welche 3 Möglichkeiten gibt es mit einer einzelnen Kamera Tiefenbilder zu erzeugen? Erkläre kurz die Funktionsweise.

- Time of Flight (TOF)
  - Compute TOF for each pixel (IR Pulse) , computes depth based on time the ray pulse needs to return to the sending device
  - approximately 7m range, e.g.: 0.5m – 7m Range
  - 320x240 resolution and 256 depth values
- Structured light
  - e.g: Microsoft Kinect
  - Projects a pattern into the scene (IR pattern) with randomised sub-patterns
  - Camera records infrared images. The shift in the pattern of an image is detected
  - Heavy machine learning is used to "guess" poses
- Light field camera
  - Microlens array is used to capture 4D light field information of the scene
  - Allows reconstruction of the whole light field situation

Was sind die Schritte der modernen Graphik-Pipeline? Erkläre kurz die einzelnen Schritte.

OpenGL 3.2+ Core Profile pipeline:
1. Application stage
2. Command buffering and interpretation
3. Vertex Shader (the input vertex can be transformed, most commonly from object space into projection space. Vertices can also be displaced and per-vertex calculations can be done)
4. Tessellation Control Shader (the tessellation control shader defines how much tessellation of the incoming primitive (usually a triangle or a quad) is necessary, or if any is necessary at all, by defining the outer and inner tessellation factors. It can also set these values to 0 to fully discard the primitive as a form of culling)
5. Tessellator (tessellates the primitive)
6. Tessellation Evaluation Shader (The tessellated or non-tessellated primitive attributes are sent as inputs but usually require a form of barycentric interpolation to get the interpolated position, normals and texture coordinates, which has to be done in this shader. Following this step the shader can do everything the vertex shader does)
7. Geometry Shader (Works similarly to the tessellation shader (but all in one shader) but is good only when only small amounts of additional geometry is to be added based on the incoming primitive, and also only usable when no transform feedbacks are necessary)
8.Rasterizer (The vertices that should now be in projection space undergo culling and perspective division. It is determined which pixels the primitive covers and the values are interpolated perspectively for these
9. The fragment shader determines the colour of the final pixel on the screen. It has the interpolated inputs and uses them usually to do things such as fetching texture values for texturing and to perform per-pixel lighting calculations for shading.

Welches Tracking Verfahren würdest du verwenden, wenn Personen mit einer Genauigkeit von 50cm in einem mehrstöckigen Gebäude getrackt werden sollen? Beschreibe die Funtionsweise dieses Systems. Wie sieht die Einkaufsliste aus, um dieses System einzusetzen?

Either:
Albatec: Local Position Measurement (LPM)
- Uses Ultra-Wideband transmission
- Active tags for each person in the house
- Unique identifier per tag
- Phase difference TOF
- Using multiple sensors around the house

Intersense IS-1200:

hybrid of: Inertial tracking + Marker tracking (optical)
very accurate 2-5mm
Inside-out tracking
Unlimited Fiducials can be placed
6-DOF
Requires Intersense optical camera, windows server as host, the Intersense software and printed out fiducials
Disadvantages: lots of markers needed

Erkläre die Funktionalität von GPS. Was ist der Unterschied zwischen GPS und DGPS?

• Large scale tracking based on radio signals, based on distance measurement to satellites
• 24 satellites in orbit + Master Control Station on ground, 6 monitoring stations and 4 antennas
• 4-5 satellites are used to determine the user's position (theoretically at least 3 required)
• Satellite has atomic clock. GPS receivers have less accurate clock. Therefore more satellites needed
• Accuracy of SPS: 100m, PPS has at least 10m accurate tracking

+ Cheap and easy, uniform accuracy of 10-100m
- Direct line of sight needed (doesn't work indoors), doesn't work in narrow streets and forests, bad resolution and precision

DGPS = Differential GPS
Uses additional fixed ground stations to refine the resolution (in Austria : EPOSA)
Theoretically 0.1m, in reality 3-5m accuracy

Augmented Reality - Azuma Definition

• Augmented reality combines the real and virtual world
• Interactive in real-time
• Registered in 3D (real and virtual objects are in a 3D relationship to each other)
Requires precise, fast, robust and accurate tracking and registration (combining real and virtual world)

Collaborative AR/VR
The virtual space is shared among multiple people : a social interaction between them is possible
Assists social interaction and cooperation (natural communication + supports teamwork)

Distributed AR/VR
Allows big groups of people to interact with each other across large distances. Flexible hardware setups and flexible distribution of computing power and resources possible.

Mixed Reality
Often used as synonym for AR: Any combination of VR and real world or coupling of real world with the virtual world

What can be measured using inertial systems:

Accelerometer: measures acceleration which can be integrated twice to measure rotations relative to a starting orientation (3DOF, roll pitch yaw)
Gyroscope: Measures angular velocity based on momentum. Vibrating gyroscope uses carioles effect to determine orientation
Electronic compass: earth's magnetic field (2DOF)

Projektionstisch, mehrere Nutzer:

Workbench:
• One or two screens, can be desk or a large display
• high resolution,
• intuitive display for certain applications
• allows table-top placement of props
• Can be shared by multiple users
• Pen-based and touch input possible
• Large FOV

Two-User Workbench
Difficult to do stereo:
Doubled interleaving
• Refresh rates
• Cross-talk
• Reduced brightness

Multi-user approach:
Private screens and framebuffers per user
or time-sequential display of images (special glasses required to separate images) or screen partitioning with additional optics

Was ist WIM, Vorteile u. Nachteile

World in miniature:
• All manipulation can be done in reach
• However it doesn't scale well for large environments.
• Indirect manipulation

Voodoo dolls

• Two-handed pinch gloves
• Interaction of the gloves on image plane selects object and creates a voodoo object
• Voodoo object can be manipulated and moved between hands
Original object is transformed towards the new manipulated stands when the manipulation task is finished

Simple Virtual Hand

• Natural and easy placement with 1:1 position mapping
• Limited reach, fatiguing overshoot

Go-Go Technik erklären

Implements a metaphor for interactively letting the arm grow so that it can reach farther in the virtual world. When the hand goes beyond a certain distance from the body it starts growing (polynominal function)
Variations: Fast Go-Go, Stretch Go-Go, Indirect Go-Go
Natural and efficient manipulation but getting hand in the correct position to grab objects is difficult
Ray-casting

• Requires little effort
• Exact positioning and orientation is difficult

Arm retains at a constant length
Originating from the hand a ray is cast into the virtual scene which can be used to select/pick up objects and manipulate them. With added reeling (indirect depth control), the distance can be changed at which objects are selected
! No tiring and infinite reach possible, but it separates the DOFs and it isn't natural
! Easy to grab objects but manipulation is difficult !

Homer technique:
"Hand-Centered object manipulation extending ray-casting"
Ray-casting + arm extension
• Select: via ray-casting
• Virtual hand moves to object
• Natural hand motions manipulate the object
• Easy selection and manipulation but hard to move objects away from you

Moving objects: Indirect HOMER (joystick buttons) or arm motion linear mapping (Direct HOMER)

Scaled-world grab:
Selection by image plane: world scaled down around virtual hand (or user avatar scaled up)
easy and natural manipulation
hard to move objects away

Floating menus:
2D menu floating in 3D:
3D selection for a 1D task
Can be difficult to find
Occluding environment
User knowledge needed

Pop-up menus:
pie-menu with 3D selector
Shadow stick is rotated to point in desired pie section (selection is shown)

Pen & Tablet
Menu hidden in physical object (tablet)
Menu can be hidden by actual physical interaction
Interface easy to understand because known from desktops
Any 2D interface possible
Handwriting recognition possible
clear interface for input

Nenne 3 Anforderung eines AR/VR Frameworks in Zusammenhang mit Studierstube

Has to use the scene graph library
For DIV it requires reliable network communication (replicate scenegraph over multiple machines=

Multi-user support
Support for distribution
3D event system
3D widgets