[20240320] INFOMMMI - Multimodal interaction - 3-GS - USP

Course: BETA-INFOMMMI Multimodal interaction (INFOMMMI)

Note: This file only contains the questions covering lectures 5-7 (W. Hürst)

Duration:	2 hours
Number of questions:	14

7 Defining AR systems

[max. 7 points] In 1997, R. Azuma introduced a definition of Augmented Reality (AR) that defined it via three characteristics. What are these? *(It is sufficient to name them. No explanation of them is needed. The order in which you list them does not matter.)*

1 pt. **a.** [max. 1 point] First characteristic: Combines real and virtual

1 pt. **b.** [max. 1 point] Second characteristic: Is interactive in real time

1 pt. **c.** [max. 1 point] Third characteristic: Is registered in three dimensions

AR systems can be described with respect to these characteristics. Some do not fully fulfill all of them, others do fulfill them but to a rather limited degree. One AR system that we saw in the lecture is the IllumiRoom. Discuss this system with respect to the three criteria listed above. *(State if it does fulfill these criteria. If "no", shortly explain why not. If "yes" or "partial", shortly explain why or how, respectively. In case you need a refresh of your memory: The IllumiRoom was the spatial AR system from Microsoft that projected content that is associated with the content on your TV in the area surrounding it.)*

2 pt. **d.** [max. 2 points] First characteristic:

2 pt. **e.** [max. 2 points] Second characteristic:

2 pt. **f.** [max. 2 points] Third characteristic:

The answer to all is “yes, but …” or “partially” because all of them fulfill it but in each case, we can argue that there are limitations compared to a “perfect” AR system that does this in the best possible way. The explanation to what degree they fulfil the criteria varies (and is debatable, because it is not a linear characteristic that can be narrowed down to a single parameter). Therefore, all answers that reflected a good understanding of the definition as well as the system gave full credits.

Depending on the explanation, a “no” answer could also give full credits if it reflected a correct understanding of the criteria. For example, one could say that it is not interactive in real-time because the interaction is only indirectly via the TV and not directly with the augmented elements. I would still call this interactive (and thus answer “Yes, but …”), but the explanation would be the same as in the “no”-answer before.

There were also different examples in the video. Some were interactive, others not. If someone used a concrete example and the explanation reflected that they correctly understood the related criteria, full credits were given.

8
3 pt.

Different AR systems

[max. 3 points]

In the lecture, we mentioned a few AR systems that did not provide a first-person view (although the majority did). Give an example of such a system and shortly explain why a non-first-person view makes sense in this context. *(You can bring one of the examples from the lecture, but any other correct answer will give full credits as well.)*

Probably the most obvious example is the “magic” mirror that allows you to try on “virtual” clothes in a department store. A pure “first-person view AR system” would not make senses in that case. It would allow you to look down and create the illusion of you wearing, e.g., a shirt, but in a comparable real-world situation, we would also prefer looking at the mirror, that is, have a non-first-person view on it.

Some people mentioned projected maps in an AR navigation system (compared to a first-person navigation system that, e.g., puts arrows into the environment). If explained correctly, this is a nice example, too.

9 Tracking & 3D registration

In AR, ideally, we want to have "perfect" 3D registration. Two tracking-related aspects we need for this are calibration and tracking.

1 pt. **a.** [max. 1 point] Shortly explain what is meant by 3D registration.

1 pt. **b.** [max. 1 point] Shortly explain what is meant by calibration.

1 pt. **c.** [max. 1 point] Shortly explain what is meant by tracking.

See slide 39 in lecture 1 plus the explanation of 3D registration in relation to Azuma's AR definition.

10 Sensors for AR

Sensors have different characteristics. One of them is if they return "absolute" or "relative" data. *(In the following, a short answer, such as one sentence, can be sufficient to get full credits.)*

2 pt. **a.** [max. 2 points] Give an example for a sensor that returns absolute data. Name the sensor and explain why the data returned by it is characterized as "absolute."

2 pt. **b.** [max. 2 points] Give an example for a sensor that returns relative data. Name the sensor and explain why the data returned by it is characterized as "relative."

Absolute: reference coordinate system set in advance (e.g., room, world)
Examples: Basically, anything that delivers a fixed reference point in the real world. For example, GPS (gives absolute location in the world) or a marker tracked with the camera / computer vision (gives a fixed location in the real world, i.e., the marker's location and orientation)

Relative: reference coordinate system established dynamically (e.g., relative to a previous pose)
Examples: Basically, anything that gives you relative location or orientation information with respect to a reference point that cannot be associated with a fixed location in the real world. For example, accelerometer (gives relative orientation of a device with respect to gravity) or gyroscope (gives relative orientation with respect to an internal reference point)

Assume we have the following sensors: A magnetometer (aka digital compass), an accelerometer and gyroscope, and a GPS sensor. *(Note: This question is about sensors for tracking. The display technology does not matter. If cameras are needed for this display (e.g., for a video see-through system or handheld AR) they should not be used for the tracking.)*

3 pt. **c.** [max. 3 points] Give an example for an AR system where you need all of these sensors. For each sensor, shortly explain why we need it or what it is needed for. *(A short description that illustrates that you understood the problem is sufficient. A detailed discussion of the system is not needed.)*

3 pt. **d.** [max. 3 points] Give an example for an AR system that you can not build with these sensors. Shortly explain why. *(A short description that illustrates that you understood the problem is sufficient. A detailed discussion of the system is not needed. It can even be sufficient to just mention a characteristic of this system that cannot be built with it.)*

See lecture 2, slide 7, for an obvious example for c. Various other correct examples exist.
Examples for d. are given on the following slides and include basically everything that needs information about the environment (which you can't get with these sensors).

11 Optical see-through vs. video see-through displays

In AR with head-worn displays, latency is typically operationalized as the amount of time required for the digital display to update given a user's head motion.

2 pt. **a.** [max. 2 points] Describe a negative effect that can happen with optical see-through displays due to latency that does not happen for video see-through displays (aka passthrough video displays). Shortly explain your answer.

From the slides set about displays, lecture 6 (disadvantages of optical see-through displays):

- Lag & jitter between real & virtual images (”*swimming artifacts*”)

Explanation should focus on the fact that there is a difference between the real world and virtual elements (the latter appear “swimming” due to latency)

2 pt. **b.** [max. 2 points] Describe a negative effect that can happen with video see-through displays (aka passthrough video displays) due to latency that does not happen for optical see-through displays. Shortly explain your answer.

From the slides set about displays, lecture 6 (disadvantages of video see-through displays):

- Latency & jitter can lead to mismatch between eye & vestibular system (vestibular-visual conflict, like in VR)

Explanation should focus on the fact that everything that is visible might be displayed delayed due to latency, causing a mismatch what your body perceives from the real world and what your eyes see.

12 Display characteristics (video see-through displays)

In their paper "Seeing the World through Digital Prisms: Psychological Implications of Passthrough Video Usage in Mixed Reality" from 2024, Bailenson et al. take a critical look at passthrough video technology used for AR displays (aka video see-through displays).

1 pt. **a.** [max. 1 point] Displays have various characteristics. One of them is visual acuity, which, according to Bailenson et al. is commonly measured in Pixels-Per-Degree (PPD). Another is field of view (FOV). Shortly explain what FOV means in the context of an AR system.

See lecture 6, slide 35, for a formal definition in relation to AR. Other, informal descriptions exist and got full credits as well (if correct).

FOV = max. number of degrees of visual angle that can be seen instantaneously on a display

2 pt. **b.** [max. 2 points] Bailenson et al. mention that some AR systems, such as Apple's Vision Pro, sacrifice the FOV for a higher PPD. Give one possible reason or use case that justifies this decision. *(They mentioned one in the paper, but any convincing one will give full credits.)*

Every context that requires a high resolution. For example, reading text. This can only be done comfortably at a high resolution, but having a wide FOV is not necessary for it.

2 pt. **c.** [max. 2 points] Give one possible reason or use case why a manufacturer of a different headset might make the opposite decision, that is, build an AR headset with a wider FOV but a lower PPD.

Every context where people want or need to see what is going on in different directions. An example could be a navigation system that places direction information, such as arrows, directly on location, that is, integrate into the environment (e.g., placed on the floor). Some argued with safety (when navigating) or better immersion, which are a good points, too.

2 pt. **d.** [max. 2 points] Bailenson et al. state that there is a lack of research with respect to the longitudinal usage of headsets using passthrough video. Give one convincing (and concrete) reason why it is important that this gap in the common knowledge is filled. *("Concrete" just means that generic answers that apply to all research will give few or no credits.)*

See paper (Section “Looking Forward”)

2 pt. **e.** [max. 2 points] Name one problem of passthrough video for AR that they identified in their paper that does not exist with optical see-through display (or to a lesser degree). Name the problem and shortly explain it. *(A short explanation, e.g., one sentence, could be sufficient to get full credits.)*

Examples from the paper:

- Tunnel vision / small field of view
- Negative psychological outcome, e.g., impeding spatial understanding of a scene, decreasing feeling of presence
- Distortion: distortion of colors and lighting, hindering spatial information processing, problems with hand-eye coordination, producing nausea, oculomotor discomfort, disorientation
- Aftereffects that carry over into the real world
- Problems with distance estimation
- Simlator sickness
- Social issues (“glasshole”)

13 Interaction devices for AR

In the lecture, we categorized interaction devices into four possible categories. For each of these categories, give an example and shortly state an advantage of it compared to the others. *(A short explanation is sufficient. That is, only one advantage or benefit that is characteristic for this type of device is sufficient.)*

2 pt. **a.** [max. 2 points] Name an example for the category AR devices (handheld or worn) and a common advantage or benefit of this approach.

2 pt. **b.** [max. 2 points] Name an example for the category body parts and a common advantage or benefit of this approach.

2 pt. **c.** [max. 2 points] Name an example for the category special devices (handheld or worn) and a common advantage or benefit of this approach.

2 pt. **d.** [max. 2 points] Name an example for the category natural devices (part of the environment) and a common advantage or benefit of this approach.

See slide 9 and following ones in lecture 7.

14 Interaction techniques for manipulation

Many interaction approaches for the manipulation of objects in virtual reality can be categorized as either egocentric or exocentric. *(Note: You might want to read all sub-questions before answering them. The third one depends on the examples that you chose for the first two.)*

1 pt. **a.** [max. 1 point] Give an example for an egocentric interaction approach. *(Just name it. No explanation required.)*

1 pt. **b.** [max. 1 point] Give an example for an exocentric interaction approach. *(Just name it. No explanation required.)*

See slide 52 (and 56) in lecture 7.

2 pt. **c.** [max. 2 points] Give an example for a situation or use case where one would rather use the exocentric approach that you mentioned above instead of the egocentric one that you named. Shortly explain why.