(INFOMMMI) Multimodal Interaction - 8 april 2019

Course: Multimodal interaction (INFOMMMI)

Questions 1-9 cover lectures 1-4 by Peter Werkhoven (max. 60 points).
Questions 10-13 cover lectures 5-8 by Wolfgang Hürst (max. 40 points).
Note that questions 10-13 contain multiple subquestions. Thus plan your time accordingly.

Good luck!

Number of questions: 13

You can score a total of 100 points for this exam, you need 50 points to pass the exam.

**Note that this file only contains the exam questions
covering lectures 5-8 by W. Huerst.**

10   Comparing AR with VR

Assume you are a developer of AR and VR systems. Name one aspect that is generally more difficult to achieve with AR than with VR. Give a short explanation why. *One sentence can be sufficient to get full credits. Note that this question is about system development, not usage. Thus, make sure your answer addresses an aspect that one comes across when developing it (i.e., not a feature or usage-related aspect of the system).*

2 pt.   **a.**   [max. 2 pts] Aspect that is generally considered to be **harder** to achieve **in AR** than in VR:

Now name another aspect that is generally more difficult to achieve with VR than with AR. Give a short explanation why. *Again, a short answer is sufficient, and make sure to focus on a development-related aspect, not a usage-related one.*

2 pt.   **b.**   [max. 2 pts] Aspect that is generally considered to be **harder** to achieve **in VR** than in AR:

In the lecture, we compared the typical "interaction loop" of a VR system ("tracking – simulation – feedback – control") with the one for an AR system. In the following, assume we are comparing two head-mounted displays; for example, the HTC Vive for the VR system, and a Microsoft HoloLens for the AR system.

3 pt.   **c.**   [max. 3 pts] For **tracking**, give one example where **there is *no* real difference**; for example, an aspect of the implementation where we can use the very same algorithms in both cases (VR and AR).

3 pt.   **d.**   [max. 3 pts] For **simulation**, give one example where **there is a difference** between VR and AR; for example, a situation where we would need to use a different algorithm to realize it in AR, or a situation that is different in VR and AR, thus requiring us to use a totally different approach.

11   Comparing different AR systems

When using **head mounted displays** for augmented reality, we can distinguish between optical see-through displays (OSTs) and video see-through displays (VSTs).

2 pt. **a.**   [max. 2 pts] Give one occlusion problem that appears with OSTs but **not** with VSTs. Shortly explain your answer (*2 short sentences could be enough to get full credits*).

2 pt. **b.**   [max. 2 pts] Give one occlusion problem that does appear with **both** OSTs and VSTs. Shortly explain how one usually deals with it (*2 short sentences could be enough to get full credits*).

Instead of AR with head mounted displays, we can also create so-called **handheld AR** by using our smartphone. Here, the live video stream of the phone's camera is enriched (i.e., augmented) with graphics integrated into the live video (i.e., reality).

Assume the university asks you to develop an AR information browser that gives people on campus live information about the buildings. That is, people walking around on campus can start the app on their phone, point it at a building, and then get graphical or textual information associated to all buildings that are currently shown in the live video stream.

3 pt. **c.**   [max. 3 pts] What kind of sensors from the phone would you use to implement this? Name each sensor and specify the data that you would use from it (*short phrases for each sensor could be sufficient to get full credits; no lengthy explanation is needed*).

Note that the AR app only needs to work outside. It is not expected that it works indoors. Further note that there are obviously different ways to achieve this. Make sure that your solution is sufficiently reliable and efficient (i.e., uses the minimum amount of resources of the device).

In the lecture, we saw an approach where your smartphone is used to create an OST head mounted display by putting it into a cheap "frame", similarly to how you can use a Google Cardboard to create a cheap virtual reality headset (to refresh your memory: the device we looked at was called *Aryzon* and used mirrors, reflective glasses, and stereoscopic lenses in a cheap frame to achieve this).

Assume a museum wants to develop an AR app that allows people using such a device to get additional information for some pieces of their exhibition. That is, people looking at, for example, an art piece via these goggles get additional graphical or textual information associated with it.

3 pt. **d.**   [max. 3 pts] Shortly state why this situation is different from the one above (i.e., the university's AR information browser). How would you realize this app, i.e., what information would you need to display the augmented information in the app and how would you get it? *(A short explanation is sufficient. No need for lengthy elaborations.)*

Now assume we want to use a smartphone with a Google Cardboard-like frame to create a VST head mounted display.

2 pt. **e.**   [max. 2 pts] Shortly explain why it might be a good idea to use an additional camera mounted on the cardboard frame for this instead of using the one integrated in the smartphone. *(1-2 sentences could be sufficient to get full credits.)*

12 AR interaction

Interaction in AR (and VR) is usually done in 3D. (*Note: also read the second sub-part of the question before answering the first one.*)

1 pt. **a.** [max. 1 pt] Give one reason why this is generally more difficult than standard human-computer interaction (e.g., with keyboard and mouse).

Two techniques to interact with objects that are out of reach in VR and AR are *ray casting* and *cone casting (*aka *flashlight)*.

2 pt. **b.** [max. 2 pts] What problem does cone casting solve that commonly exists with ray casting? Shortly explain (*2 sentences should be enough to get full credits*).

2 pt. **c.** [max. 2 pts] What is a potential problem or disadvantage of this approach?

4 pt. **d.** [max. 4 pts] In their paper "Augmented Reality vs Virtual Reality for 3D Object Manipulation" from 2015, Krichenbauer et al. compared task completion time for selection and transformation tasks in VR and AR. Their results show that when using a 3D input device, completion time in AR was generally lower than in VR. Shortly state what might have been the reason for this. (*Note: you do not have to write down the details of the results here. General statements, similarly to the one used by the authors in their hypothesis, can be sufficient to get full credits.*)

13 Multimodal AR (smell/olfactory AR)

In the last lecture, we talked about a device produced by the company VASQO that attaches below a VR headset and produces the illusion of smell by spraying small amounts of fragrance. Now assume the company would also make such a device for an AR headset, such as MicroSoft's HoloLens. For the olfactory part of such a multimodal AR system, list each of the three criteria from Azuma's AR definition and shortly state if they would be fulfilled or not.

(*No lengthy explanation is needed, as long as it becomes clear that you understood the definition and how to apply it to this concrete context. Note that without having tested the device yourself, you might not be able to judge certain things just based on the description from the lecture and the video that we saw. If that is the case, it is perfectly okay to start your answer with something like: "Assuming this is implemented in a way that ... ".*)

1 pt. **a.** [max. 1 pt] First characteristic according to Azuma:

2 pt. **b.** [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.

1 pt. **c.** [max. 1 pt] Second characteristic according to Azuma:

2 pt. **d.** [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.

1 pt. **e.** [max. 1 pt] Third characteristic according to Azuma:

2 pt. **f.** [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.

13   Multimodal AR (smell/olfactory AR)

In the last lecture, we talked about a device produced by the company VASQO that attaches below a VR headset and produces the illusion of smell by spraying small amounts of fragrance. Now assume the company would also make such a device for an AR headset, such as MicroSoft's HoloLens. For the olfactory part of such a multimodal AR system, list each of the three criteria from Azuma's AR definition and shortly state if they would be fulfilled or not.

(*No lengthy explanation is needed, as long as it becomes clear that you understood the definition and how to apply it to this concrete context. Note that without having tested the device yourself, you might not be able to judge certain things just based on the description from the lecture and the video that we saw. If that is the case, it is perfectly okay to start your answer with something like: "Assuming this is implemented in a way that ... ".*)

1 pt.   **a.**   [max. 1 pt] First characteristic according to Azuma:

2 pt.   **b.**   [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.

1 pt.   **c.**   [max. 1 pt] Second characteristic according to Azuma:

2 pt.   **d.**   [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.

1 pt.   **e.**   [max. 1 pt] Third characteristic according to Azuma:

2 pt.   **f.**   [max. 2 pts] Is this fulfilled by the olfactory part of the system? Shortly explain your answer.