INFOMMMI (Multimodal Interaction) 2017-2018

Exam questions, part 2
(lectures 5-7, W. Hürst)

(max. 40 points)

FIRST NAME:	LAST NAME:	STUDENT ID:

Please write your answers to these questions only on the pages for this part!

Don't forget to fill in your name and student ID in the dedicated boxes above on both parts!

Question 2-1: AR/VR comparison (max. 2 points)

Assume you are an AR/VR developer who is developing VR applications for the HTC Vive (which is a head-mounted display VR system) and AR applications for the Microsoft HoloLens (which is a head-mounted see-through AR system).

(Note: In the following, a short statement can be sufficient to get full credit. No detailed explanation is needed.)

Give one issue, characteristic, or aspect that is more difficult to deal with in VR (with the HTC Vive) than in AR (with the HoloLens) and shortly state why.

Give one issue, characteristic, or aspect that is more difficult to deal with in AR (with the HoloLens) than in VR (with the HTC Vive) and shortly state why.

Question 2-2: Displays / depth perception cues (max. 5 points)

In the lecture, we said that depth cues for 3D perception can be categorized in four groups: Oculomotor cues, binocular depth cues, motion related cues, and pictorial depth cues.

(Note: Below, an explanation is not needed. It is sufficient to state the display type or characteristic that the display(s) have to fulfill to be able to use these depth cues.)

What kind of display do you need if it is essential for your application to support oculomotor cues?

What kind of display do you need if it is essential for your application to support binocular depth cues?

What kind of display do you need if it is essential for your application to support pictorial depth cues?

(Note: Below, a short phrase or sentence illustrating this advantage can be sufficient to get full credit. A detailed explanation is not needed.)

Give one advantage that using an *OST (Optical-See-Through) display* might have compared to using a VST (Video-See-Through) display.

Give one advantage that using a *VST (Video-See-Through) display* might have compared to using an OST (Optical-See-Through) display.

Question 2-3: AR interaction (max. 3 points)

Interaction devices or concepts that we can use in AR include *special devices for 3D/6DOF interaction*, *6DOF hand tracking*, and *Tangible User Interfaces*.

(Note: Below, a short phrase or sentence illustrating this advantage can be sufficient to get full credit. A detailed explanation is not needed.)

Give one advantage that *special devices for 3D/6DOF interaction* have compared to the other two approaches.

Give one advantage that *6DOF hand tracking* has compared to the other two approaches.

Give one advantage that *Tangible User Interfaces* have compared to the other two approaches.

Question 2-4: True AR (max. 8 points)

In their paper “Breaking the Barriers to True Augmented Reality”, Sandor et al. describe their vision of “True AR” as an augmented reality that is indistinguishable from reality. They define it informally as “a modification of the user’s perception of their surroundings that cannot be detected by the user.”

(Note: A detailed elaboration is not needed for full credits in the following 2 questions as long as your example is convincing.)

Give an example or use case where such a True AR will likely be beneficial compared to the one we get with today’s systems and shortly state why.

Give an example or use case where such a True AR is not needed, not wanted, or could even be harmful and shortly state why.

The authors also discuss how an AR Turing Test could be designed (i.e., a test to verify if or to what degree an AR system could be classified as True AR). They propose three “dimensions along which the interactions in a useful test scenario can be restricted.” What are these three dimensions?

(Note: It is sufficient to write down the three phrases that are used in the paper. Yet, if you don’t remember the exact phrasing, you can also get full credit by giving a short informal description of the three aspects that one would need to test.)

The authors argue that Light Field Display technology is needed to achieve True AR for the visual channel. Give one reason why.

(Note: A short phrase can be sufficient to get full credit.)

Question 2-5: AR systems (max. 8 points)

AR systems can be realized in various ways by using different display technologies. One example is handheld AR (when using a mobile phone), another is projected or spatial AR (when using a projector).

(Note: For the 2 questions below, you don't have to describe a full game. Listing an aspect or characteristic that is likely easier to realize or that is a clear advantage compared to the other system can be sufficient to get full credits.)

Give one characteristic or advantage why a company developing AR applications might chose to develop an AR game *for a handheld device* and not for projected AR.

Give one characteristic or advantage why a company developing AR applications might chose to develop an AR game *for a projected AR system* and not for handheld AR.

Now we want to compare a handheld AR system that is created with a regular phone using the front facing camera, and an immersive AR system that is created with a head-mounted display such as the HoloLens using a See-Through-Display and various cameras for tracking. Assume that you want to implement an AR tower defense game. For this game, players have to place a marker on a table. The AR system shows a tower at the position of the marker. It also shows tanks approaching the tower from all sides. The goal of the game is to shoot and destroy all tanks before they reach the tower.

Give one problem that is hard or even impossible to solve when using *a regular phone* to implement this game, but that does not appear or is easier to resolve when using a HoloLens.

Give one problem that is hard or even impossible to solve when using *a HoloLens* to implement this game, but that does not appear or is easier to resolve when using a regular phone.

Question 2-6: Non-visual AR / Azuma's AR definition / sensors (max. 14 points)

In 1997, R. Azuma introduced a formal definition of AR by specifying three characteristics that each AR system should fulfill. Although his article only elaborates on visual displays, the definition was intended to cover other modalities as well. Recently, the speaker and headphone manufacturer Bose introduced AR glasses that do neither contain visual displays nor cameras, but are purely focused on audio. We shortly discussed it in the lecture. Here an excerpt from a related news article:

"Unlike other augmented reality products and platforms, Bose AR doesn't change what you see, but knows what you're looking at – without an integrated lens or phone camera," Bose said. "And rather than superimposing visual objects on the real world, Bose AR adds an audible layer of information and experiences, making every day better, easier, more meaningful, and more productive."

Discuss if Bose's AR system fulfills the characteristics of Azuma's definition by completing the sentences below (and crossing out the parts that are not correct).

(Note: The option "does only partly fulfill" has been added below because there might be cases that are debatable or when the above description might not provide all information to give a clear yes or no answer. It is more important that your answer reflects that you really understood the essence of this characteristic than just guessing the right answer.)

A first characteristic of Azuma's definition is:

The Bose AR system [**does fulfill** | **does only partly fulfill** | **does not fulfill**] this characteristic because:

A second characteristic of Azuma's definition is:

The Bose AR system [**does fulfill** | **does only partly fulfill** | **does not fulfill**] this characteristic because:

A third characteristic of Azuma's definition is:

The Bose AR system [**does fulfill** | **does only partly fulfill** | **does not fulfill**] this characteristic because:

In the news article cited above, they also list several potential use cases for such a system. For the two examples quoted below, list what kind of sensors have to be integrated into the device to realize it and shortly state why (i.e., what kind of data you get from the sensor(s) that is needed for this use case).

(Note: Only list sensors that are absolutely needed and don't just write down any sensor that comes to your mind. Adding ones that are obviously not needed will result in zero credit.)

Use case example 1: *"For travel, the Bose AR could simulate historic events at landmarks as you view them … You could hear a statue make a famous speech when you approach it."*

Sensor(s) needed and data they deliver:

Use case example 2: *"Bose AR would add useful information based on where you look. Like the forecast when you look up or information about restaurants on the street you look down."*

Sensor(s) needed and data they deliver:

Bose explicitly states that their device does not include any cameras. Give one use case example for audio AR that cannot be implemented for the Bose AR glasses because a camera would be needed as sensor.

(Note: You can either shortly describe an example case like above or just list a characteristic that is impossible to realize without a camera.)

Example / characteristic: