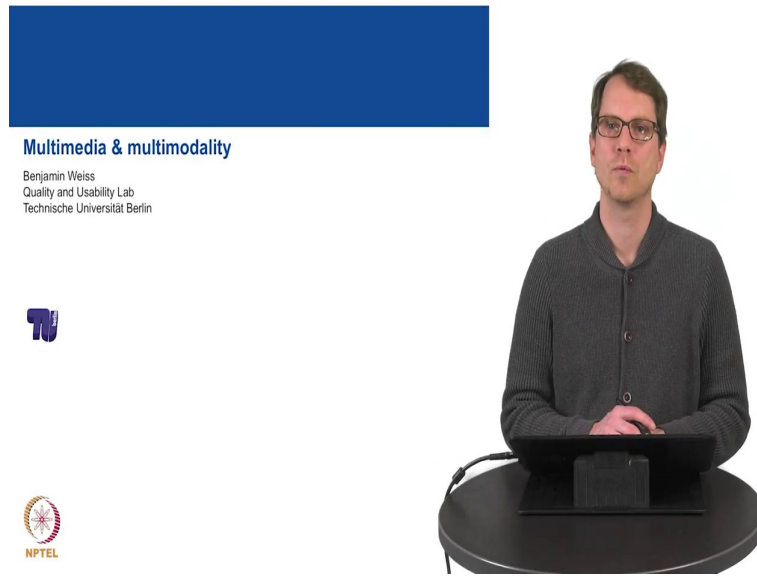


Multimodal Interaction
Multimedia and Multimodality
Professor Benjamin Weiss
Quality and Usability Lab
Technische Universität Berlin
Media versus Modality

(Refer Slide Time: 00:19)



Multimedia & multimodality

Benjamin Weiss
Quality and Usability Lab
Technische Universität Berlin

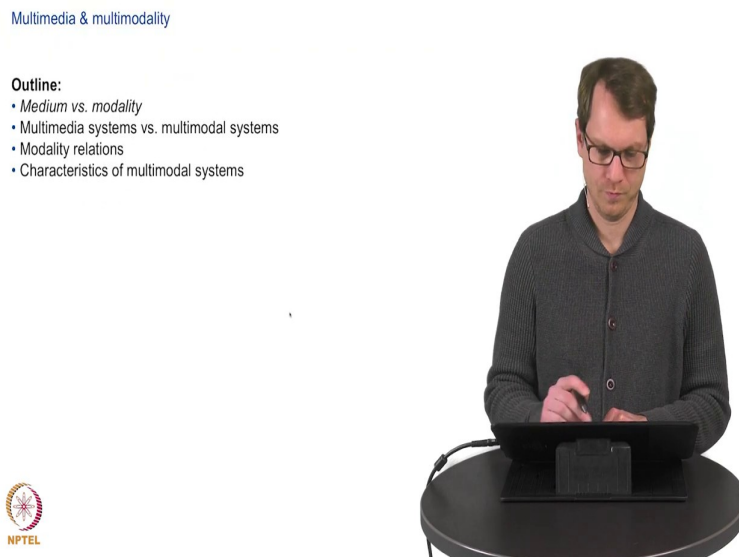
TU

NPTEL

The slide features a blue header, the title 'Multimedia & multimodality', the presenter's name and affiliation, the logos for TU and NPTEL, and a video frame of Professor Benjamin Weiss sitting at a desk with a laptop.

Welcome to the first week of this course. It is about multimedia and multimodality and the difference between the two.

(Refer Slide Time: 00:25)



Multimedia & multimodality

Outline:

- *Medium vs. modality*
- Multimedia systems vs. multimodal systems
- Modality relations
- Characteristics of multimodal systems

NPTEL

The slide features the title 'Multimedia & multimodality', an 'Outline' section with four bullet points, the NPTEL logo, and a video frame of Professor Benjamin Weiss sitting at a desk with a laptop.

This is the outline for this week. I will start with the definition of medium and modality and then continue with characteristics of multimedia systems and multimodal systems.

I will then present two approaches to relate different modalities to each other and I will complete with characteristics of multimodal systems, in particular the different benefits we have with the multimodal systems compared to ordinary ones.

(Refer Slide Time: 00:52)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information



Let's start with the Medium. A medium is a technical means to store, transfer or convey information. This might be for example, the physical channel.

(Refer Slide Time: 01:04)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information
- the physical channel



So, for instance, the air for acoustics information or the light for visual information, but also

(Refer Slide Time: 01:12)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information
- the physical channel
- representation of a message to the user



the medium is a representation of a message to the user.

So, consider for example, the same information which can be presented as text, as a graph or even as video clip. This will be, this would be all different moda, this would be all different media.

(Refer Slide Time: 01:30)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information
- the physical channel
- representation of a message to the user

Modality:

- a modality is a human sensory channel



In contrast to this, the modality is a human sensory channel

(Refer Slide Time: 01:35)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information
- the physical channel
- representation of a message to the user

Modality:

- a modality is a human sensory channel
- visual, auditory, olfactory (smell), haptic (e.g. touch), gustatory (taste), vestibular (balance)



and we have several of these human sensory channels, for example visual, auditory, olfactorial channel, for example the smell, the haptic channel which is for example, the touch but also temperature sensation, and even pain sensation. There is the gustatory channel which is taste, and the vestibular channel which is related to the balance perception.

As you can see, the same example does not apply here for the modality. So if we have the same kind of information presented either as a video clip or as a text or a graphic, this would be all the same modality, namely the visual channel while they are different media used.

Still,

(Refer Slide Time: 02:22)

Medium vs. modality

Medium:

- a medium is a technical means to store, transport and convey information
- the physical channel
- representation of a message to the user

Modality:

- a modality is a human sensory channel
- visual, auditory, olfactory (smell), haptic (e.g. touch), gustatory (taste), vestibular (balance)

Modality (alternative definition):

- a modality is a user input mode to interact with a system



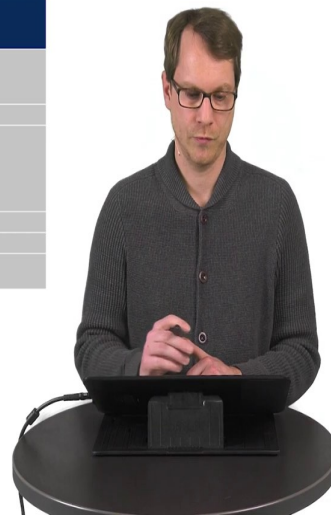
in human-computer interaction, we have an alternative definition of modality. I will explain a little bit later why this is so, namely modality as a user input mode to interact with the system.

So let us have a look at this table.

(Refer Slide Time: 02:38)

Medium vs. modality

Sense	Organ	Sensory channel (Modality)	Sensor	Example Modalities
Vision	Eyes	Visual	Camera Keyboard/pen	3D gesture recognition Face recogn./lipreading Text/handwriting recogn.
Hearing	Ears	Auditory	Microphone	Speech recognition
Touch	Skin	Haptic - Tactile - Deep sensitivity - Temp. sensitivity - Pain sensitivity	Touchscreen Accel./glove Temp. sensor	2D gesture recognition Position/motion recogn.
Olfaction	Nose	Olfactory		
Taste	Tongue	Gustatory		
Balance	Equilibrium org.	Vestibular		



On the left column you can see the different human senses like vision, hearing or touch and the appropriate assigned organ that is used to capture this information, like the eyes, ears and skin.

If we now go to the alternative definition of modality, then it is all about the capabilities of the system to mimic such human sensory channels. For example, we need a microphone to replicate hearing, or, we need a camera to replicate vision.

(Refer Slide Time: 03:13)

Medium vs. modality

Sense	Organ	Sensory channel (Modality)	Sensor	Example Modalities
Vision	Eyes	Visual	Camera Keyboard/pen	3D gesture recognition Face recogn./lipreading Text/handwriting recogn.
Hearing	Ears	Auditory	Microphone	Speech recognition
Touch	Skin	Haptic - Tactile - Deep sensitivity - Temp. sensitivity - Pain sensitivity	Touchscreen Accel./glove Temp. sensor	2D gesture recognition Position/motion recogn.
Olfaction	Nose	Olfactory		
Taste	Tongue	Gustatory		
Balance	Equilibrium org.	Vestibular		

So, on the utmost right column you can see the different modalities that are following this alternative definition of modality. This would be for example 3D gesture recognition or face recognition which could be captured by a camera, or even speech recognition with a microphone.

For all these we need of course some classifiers. There is also position or motion tracking of 3D gesture recognition like on the touch screen.