Design of Interactive Systems (DIS)

Lecture 10: Techniques for designing interactive systems – Multimodal interface design

Dr. Kalpana Shankhwar, kalpana@iiitd.ac.in
Assistant Professor
Department of Human Centered Design,
IIIT Delhi

Part II Techniques for designing interactive systems

- **Chapter 7**: Understanding
- **Chapter 8**: Envisionment
- **Chapter 9**: Design
- **Chapter 10**: Evaluation
- **Chapter 11:** Task Analysis
- **Chapter 12**: Visual Interface Design
- **Chapter 13**: Multimodal Interface Design

Introduction

- In the design of interactive systems, designers will certainly be making use of technologies that go far beyond the **screen-based systems**.
- Designers will develop **multimedia experiences** using a variety of **modalities** (sound, vision, touch, etc.) combined in novel ways.
- Issues of designing for **multimodal** and **mixed reality systems**, at designing for sound, touch and at wearable computing are discussed in this chapter.

Aims

- The spectrum of media, modalities and realities
- The key design guidelines for designing for audition
- The role of touch, haptics and kinaesthetics
- Designing for tangible and wearable computing

Introduction

Sutcliffe (2012) distinguishes several key concepts of communication:

- *Message* is the content of a communication between a sender and a receiver.
- *Medium* is the means by which a message is delivered, and how the message is represented.
- *Modality* is the sense by which a message is sent or received by people or machines.

- The term ‘mixed reality’ was coined by Milgram *et al.* in 1994 to encompass a number of simulation technologies, including **augmented reality** (digital information added to the real world) and **augmented virtuality** (real information added to the digital world).

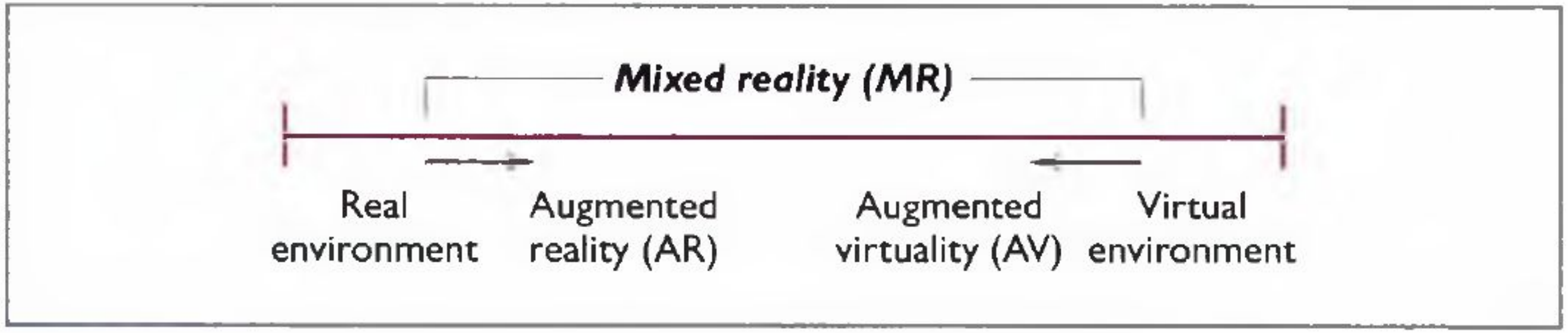
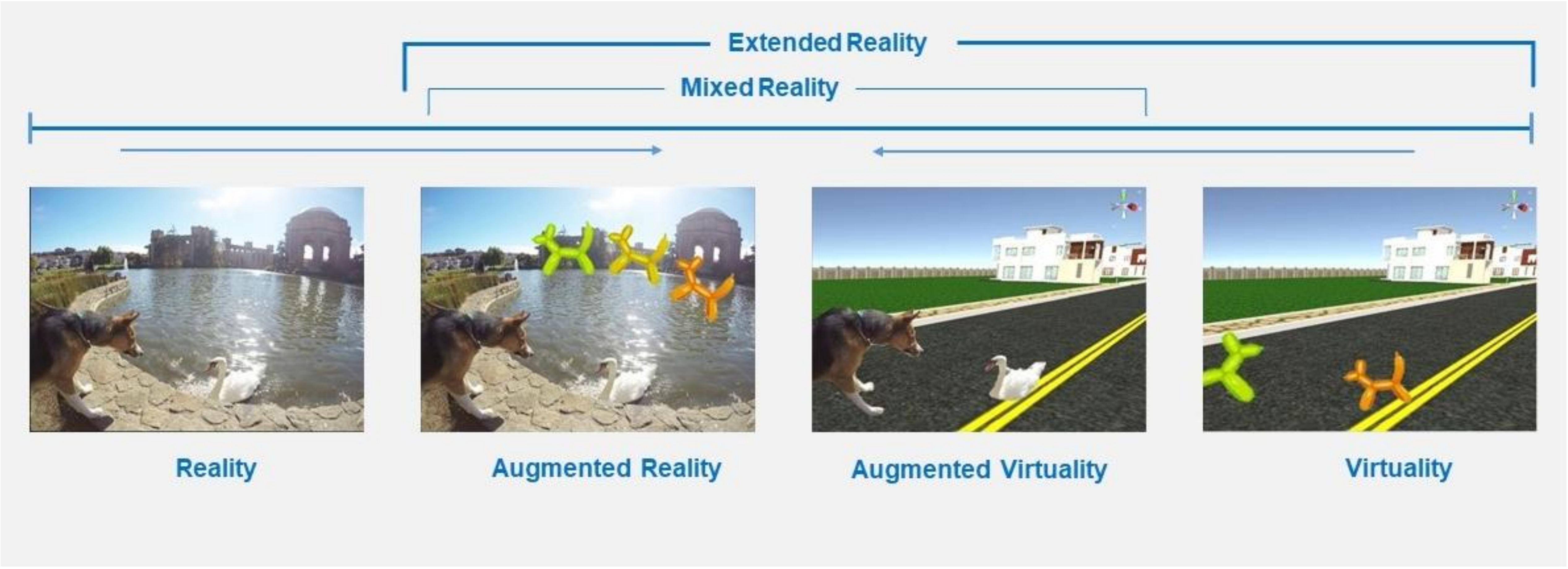


Figure 13.1 Reality–Virtuality (RV) Continuum



Mixed Reality

- In **augmented reality**, a live video stream can be enhanced with computer-generated objects (rendered so that they appear to be within the actual scene).
- Methods of presenting this visual information fall into the two main categories:

1. **Immersive** (where people see no view other than that of the mixed reality environment)
2. **Non-immersive** (where the mixed reality environment takes up only a portion of the field of view).

- The latter method can make use of a vast range of displays, including computer monitors, mobile devices and large screen displays.
- For immersive presentations people will wear a special helmet consisting a display, and which excludes any other view of the outside world.

Head mounted display

- These head-mounted displays (HMDs) are split into two categories:

1. ***video see-through*** (where the real world is recorded by a video camera and people are presented with a digital display)
2. ***optical see-through*** (where the display screens are semi-transparent, allowing a direct view of the real world and only adding computer graphics on top).

Head mounted display

Oculus Quest

HTC Vive

Valve Index

Head Mounted Display

HoloLens2

Epson

Google Glass

Modalities

- In this case computer-generated sounds can be supplied in such a way that they **appear to originate** from locations within the real environment.
- Common methods include the use of headphones or speaker arrangements, but there are more exotic technologies such as a ***hypersonic sound device***
- Of the remaining three senses, the sense of touch (or haptics) is the most developed field, with work ranging from the physical sensation of holding objects to simulating the sensation of touching different surfaces
- Smell has been simulated, but with limited success

Haptics

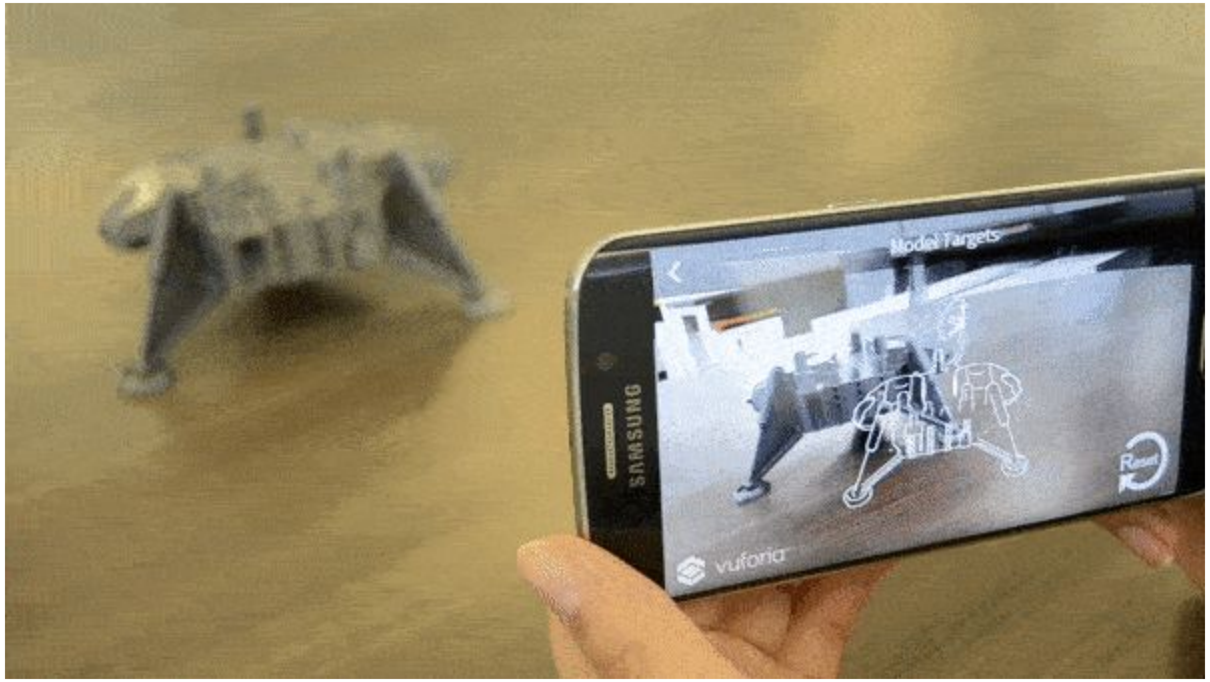
Interacting in mixed reality

- Interaction tools used in virtual reality include: '**spacemice**', which expand the **two degrees of freedom** in traditional mice (horizontal and vertical movement) to **six degrees of freedom** (horizontal, vertical, depth movements and yaw, pitch and roll rotations);
- 'Data gloves', where a glove is fitted with sensors to track **hand location** and **finger positions** and allows the grabbing and manipulation of virtual objects;
- Microsoft's Kinect allows for hand, arm and body gestures to interact with the content.

Interacting in mixed reality

- Mixed reality interaction demands the most from interaction designers
- One technical issue is that of **accurately aligning** the real and virtual environments: a process called ‘registration’.
- A notable example of this is the ARToolkit and Vuforia
- AR can not be successfully applied in remote surgery when the alignment of real and virtual worlds is a major technological

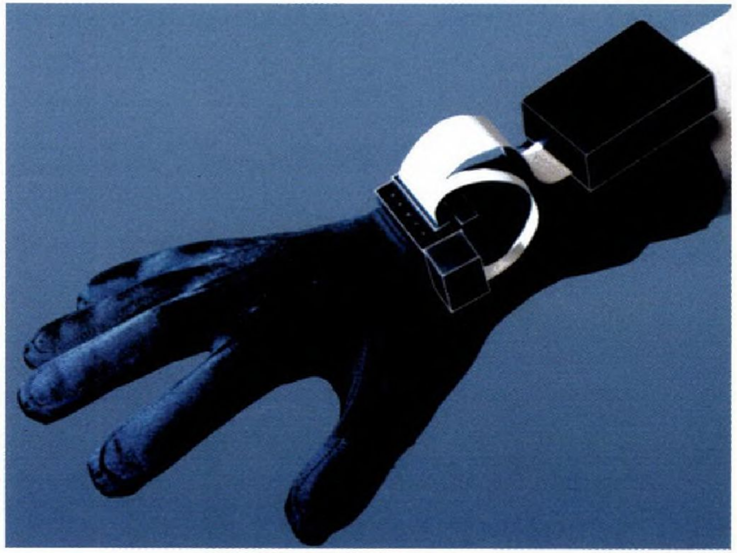
Interacting in mixed reality



Interacting in mixed reality

- Immersive virtual reality requires people to wear a **light-excluding helmet** (an HMD - head-mounted display) which houses the display, and a data glove which facilitates the manipulation of virtual objects within virtual reality.
- An HMD consists of two colour displays located in line with one’s eyes and a pair of stereo earphones.
- An HMD also has a **head tracker** which provides information about the wearer’s position and orientation in space.
- Gloves equipped with sensors (data gloves) are able to sense the movements of the hand, which are translated into corresponding movements in the virtual environment.


VIRTUALIT



Destructive
Disassembly
Training System

Interaction in mixed reality

- Non-immersive virtual reality (sometimes called desktop virtual reality) can be found in a wide range of desktop applications and games as it does not always require specialist input or output devices.
- Multimodal systems that do not mix realities, but combine gesture, speech, movement and sound, are increasingly common
- But these have their own issues to do with synchronizing the modalities
- One of the earliest systems was ‘Put That There’ (Bolt, 1980), which combined speech and gesture.

Using sound at the interface

- Sound is an increasingly important part of interface design in both mixed reality and multimodal systems.

Vision and hearing are interdependent

- Sight is a narrow, forward-facing, richly detailed picture of the world, while hearing provides information from all around us.
- An unexpected flash of light or a sudden movement orients our heads - and hence our hearing - to the source;
- The sound of a car approaching from behind makes us turn to look.
- Both sound and vision allow us to orient ourselves in the world.

Using sound at the interface

Reduce the load on the visual system

- It is now recognized that modern, large or even multiple-screen graphical interfaces use the human visual system very intensively
- To reduce this sensory overload, key information could be displayed using sound, again to redistribute the processing burden to other senses.

Using sound at the interface

Reduce the amount of information needed on screen

- One of the great design tensions in the creation of mobile and ubiquitous devices is to display a usable amount of information on a **small screen** - small as in palm-sized, or pocket-sized,
- When moving information from one place to another was expensive, the telegram ruled: ‘Send money. Urgent’
- Mobile and ubiquitous devices have very small screens which are unsuited to viewing large bodies of data.
- To minimize this problem, information could be presented in sound in order to free screen space.

Using sound at the interface

Reduce demands on visual attention

- Again in the context of mobile and ubiquitous devices, there is an uneasy need to switch from attending to the world - crossing the road, driving a car, following a stimulating presentation - to paying attention to the display of such devices
- The UK government made it an offense from December 2003 to drive a car while using a mobile phone

Using sound at the interface

The auditory sense is under-utilized

- We listen to highly complex musical structures such as symphonies and operas.
- These pieces of music comprise large complex structures and sub-structures.
- This suggests that there is, at least, the potential of using music to transmit complex information successfully.

Using sound at the interface

Sound is attention-grabbing

- While we can look away from an unpleasant sight, the same is not true of an unpleasant sound.
- The best we can do is to cover our ears.
- This makes sound very useful for attracting attention or communicating important information.

Using sound at the interface

To make computers more usable by visually disabled users

- While screen readers can be used to ‘read’ on-screen textual information, they cannot easily read graphical information.
- Providing some of this information in an auditory form can help alleviate this problem.

Earcons

- **Earcons** are a type of non-verbal audio (NVA) designed to convey a specific event, information or feedback to the user.
- Earcons are abstract, musical tones that can be used in structured combinations to create auditory messages.
- Numerous studies of the usefulness of earcons in providing cues in navigating menu structures have been conducted
- Participants in the study were asked to determine their location in the hierarchy by listening to an earcon.
- Results of this showed that participants could identify their location with greater than 80 percent accuracy.
- This suggests that earcons are a useful way of providing navigational information.

Auditory icons

- One of the most famous examples of auditory icons is the SonicFinder developed for Apple.
- Users were able to ‘tap’ objects in order to determine whether they are applications, disks or folders.
- It was possible to gauge their size depending upon how high-pitched they sounded (small objects sounded high- pitched while large objects sounded low-pitched)

Soundscapes

- The term ‘soundscape’ is derived from ‘landscape’ and can be defined as the **auditory environment** within which a listener is immersed.
- This differs from the more technical concept of ‘soundfield’, which can be defined as the auditory environment surrounding the sound source, which is normally considered in terms of sound pressure level, duration, location and frequency range.

Soundscapes

- An important issue in designing for sound is that of *discrimination.*
- While it is easy to talk about discriminating between low- and high-pitched tones
- It is quite another to discriminate between quite low and fairly low tones.
- There are a number of open questions about how well we can distinguish between different tones in context (in a busy office or a noisy reception area) and this is made worse by the obvious fact that **sounds are not persistent.**
- One of the strengths of the graphical user interface is the persistence of error messages
- Auditory user interfaces are, in contrast, transient

Speech-based interfaces

- Speech-based interfaces include speech output and speech input.
- Increasingly common in things like satellite navigation systems in cars and other areas such as announcements at railway stations, airports, etc
- Speech output uses a system that converts text to speech, TTS. Sounds are recorded from an individual and are then stitched together through the TTS to create whole messages.
- The best systems require people to train the automatic speech recognizer (ASR) to recognize their particular voice.
- This paves the way for natural language systems (NLS) where people can have conversations with their devices

Tangible interaction

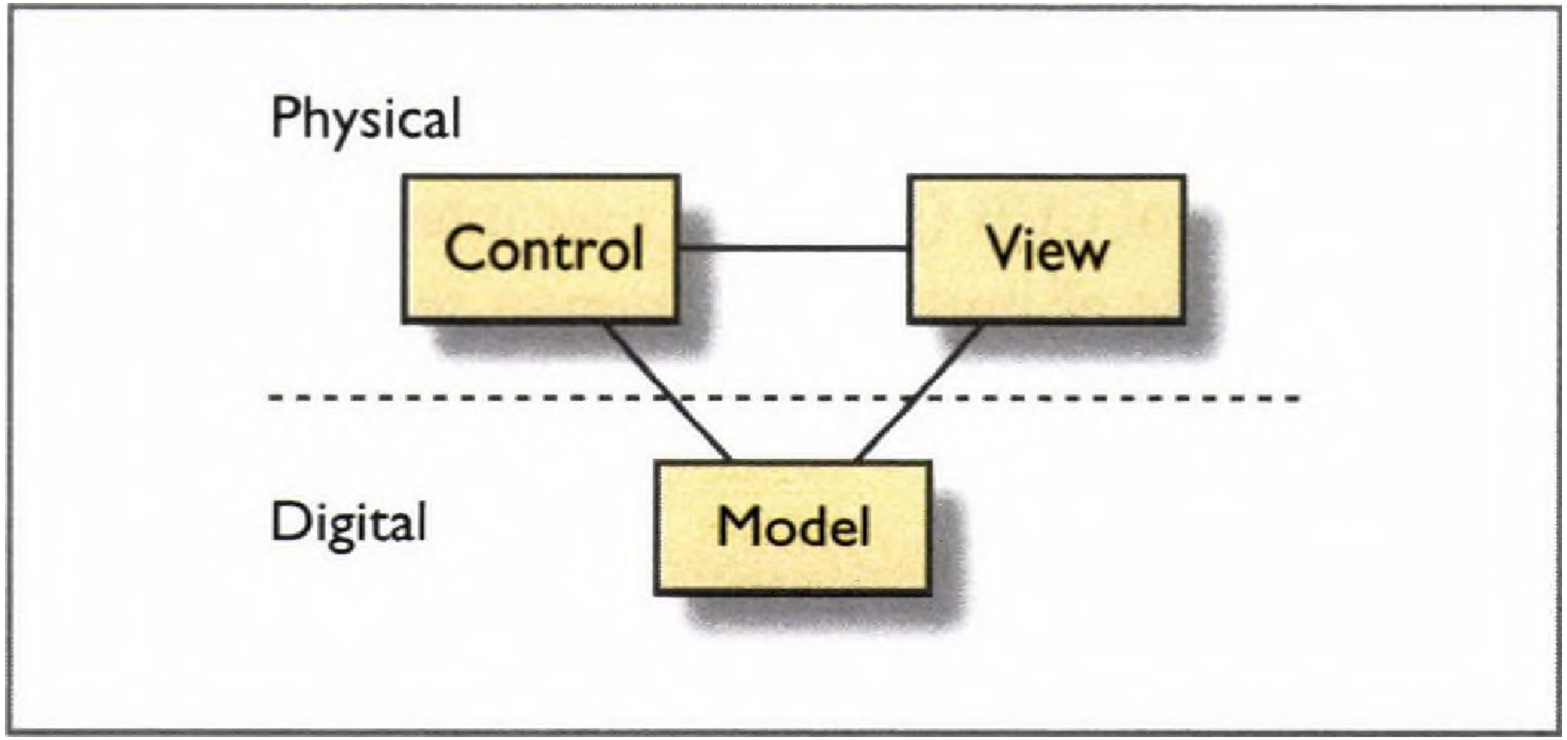
- Tangible means being able to be touched or grasped and being perceived through the sense of touch.
- With the introduction of multi-touch displays, TUIs promise to be increasingly important as they lead to interaction through physical objects and through gesture recognition.

Tangible interaction

- TUIs use physical representations - such as modelling clay, physical pens and so on
- physical drawing boards rather than pictures of them displayed on monitors
- instead of having to manipulate an image on a screen using a mouse and keyboard, people can draw directly onto surfaces using highlighter pens.
- As these tangible, graspable elements cannot, of course, perform computation on their own, they must be linked to a digital representation

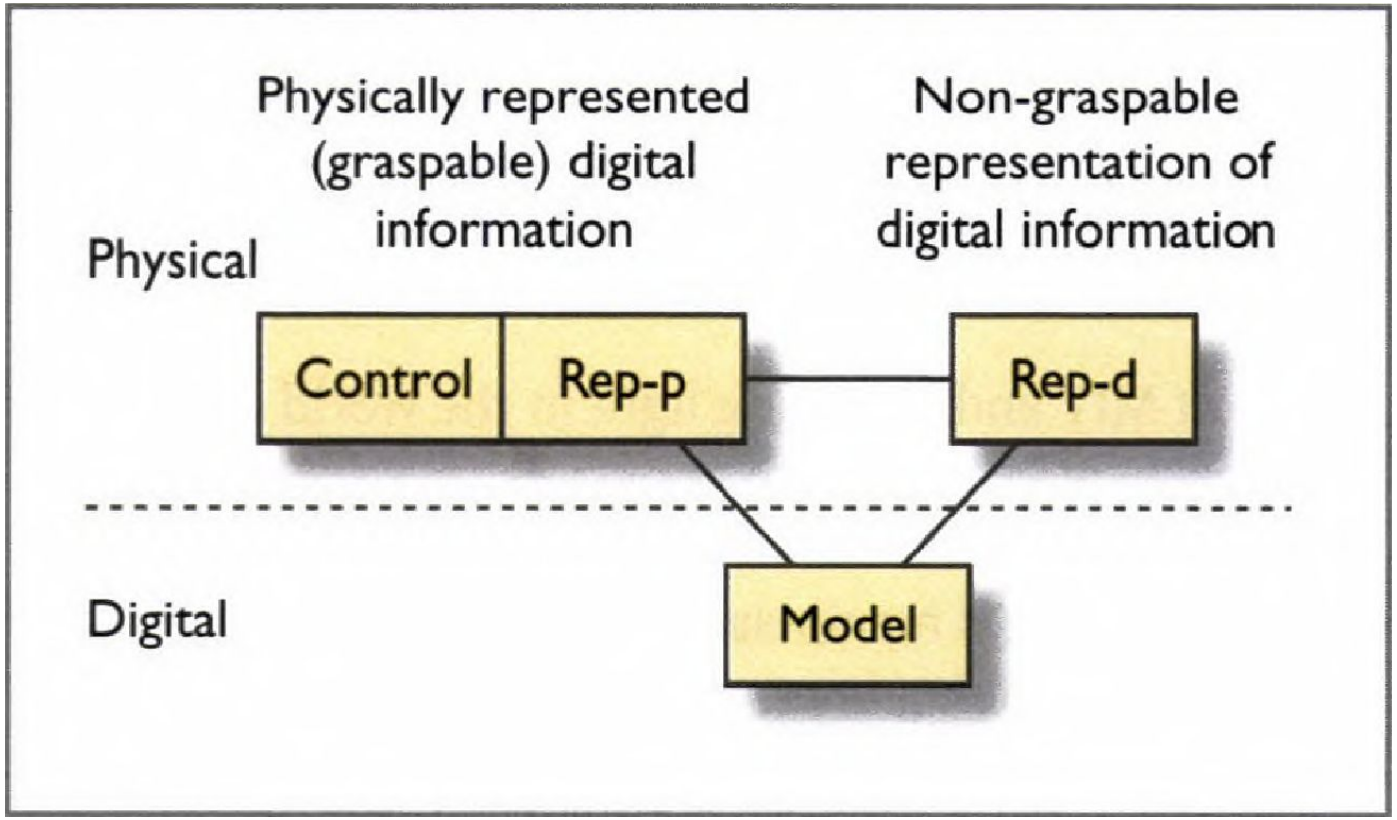
Tangible interaction

- In traditional GUIs we use peripheral devices such as a mouse or keyboard to control a digital representation of what we are working with (the model),
- the results of which are displayed on a screen or some other form of output (the view).



Tangible interaction

- TUIs in contrast have a more complex model
- view component is split between Rep-p (physical representation) and Rep-d (digital representation).



Tampere University

AR-based interaction for human-robot collaborative manufacturing

Antti Hietanen[1], Jyrki Latokartano[2], Roel Pieters[3], Minna Lanz[2], Joni-Kristian Kämäräinen[1]

Laboratories of Signal Processing[1] , Mechanical Engineering and Industrial Systems[2] and Automation and Hydraulic Engineering[3], Tampere University of Technology, Finland

First.Family@tut.fi

Gestural interaction and surface computing

- With the arrival of multi-touch surfaces - table tops, smartphones, tablets and interactive walls that recognize multiple touch points - a whole new era of interaction design is just beginning.
- The iPhone introduced gestures for ‘making things bigger’ (pinch with two fingers and draw them out) and ‘making things smaller’ (touch an object with two fingers and draw them in)
- Different applications demand different types of gesture according to the different activities that people are engaged in.

Table 13.1 iOS gestures

Gesture	Action
Tap	To press or select a control or item (analogous to a single mouse click).
Drag	To scroll or pan (that is, move side to side). To drag an element.
Flick	To scroll or pan quickly.
Swipe	With one finger, to reveal the Delete button in a table-view row, the hidden view in a split view (iPad only), or the Notification Center (from the top edge of the screen). With four fingers, to switch between apps on iPad.
Double tap	To zoom in and center a block of content or an image. To zoom out (if already zoomed in).
Pinch	Pinch open to zoom in. Pinch close to zoom out.
Touch and hold	In editable or selectable text, to display a magnified view for cursor positioning.
Shake	To initiate an undo or redo action.

Gestural interaction and surface computing

Figure 13.12 The N-wave gesture to open a browser; the red and black circles give feedback on the user's touches