

Multi Modal Interaction...

Issues and Themes...

So what do we mean by Multimodal Interaction?

- The user is able to interact with the 'Interactive Computer System' using a range of modalities, e.g. Speech, gesture.
- Combination/fusion also important, e.g. Point at light while saying "down"...
- Wikipedia Def:
 - **Multimodal interaction** provides the user with multiple modes of interfacing with a system beyond the traditional keyboard and mouse input/output. The most common such interface combines a visual modality (e.g. a display, keyboard, and mouse) with a voice modality (speech recognition for input, speech synthesis and recorded audio for output). However other modalities, such as pen-based input or haptic input/output, may be used. Multimodal user interfaces are a research area in human-computer interaction.

An example...

- Control Using Hand Gestures...



Multimodal Interaction and AI...

- Multimodal Interaction is often associated with AI
- Consider the Gesture example
 - Complex 'pattern detection' algorithms (often based on fuzzy logic) are frequently required in order to determine whether or not a given action has been carried out by the user (or that user has a given goal)
 - False positives – system detects that user has performed action X when he/she hasn't
 - False negatives – system detects that user has not performed action X when he/she has
 - I was pointing at the light when I said "down" but 'it' turned the volume of the TV down instead...

Why MMI?

- A more natural means for interaction...
 - Person at home: "Switch off (while pointing at the TV)"
 - Visitor to Lancaster: "What's this building (while pointing at Lancaster Castle)"
 - Anthropomorphic
 - System takes on human qualities, e.g. Senses, intelligence etc.
 - Can result in certain emotional responses from the user...
 - Remember the Microsoft Paperclip!
 - The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places -Byron Reeves and Clifford Nass



Why MMI?

- Can be more efficient...
 - "Switch off (while pointing at the TV)"
 - User doesn't have to walk across to the TV or even pick-up the remote control
 - In the future it may even be a case of the user simply looking at the TV (and issuing the command) and the system detecting the focus
 - Fingerprint recognition on a PDA
 - Saves user having to enter a password via the keyboard

Why MMI?

- Can be accessibility issues...
 - “Switch off (while pointing at the TV)”
 - User doesn't have to walk across to the TV or even pick-up the remote control
 - Reducing demand on certain skills
 - Handwriting or Speech recognition...
 - User doesn't have to be a good typist
- Can be Safety Elements
 - Current studies into MMI at the dentists...
 - Infection can result from touching a computer keyboard...

Enabling Technologies...

- Whole range...
 - Video Cameras (often multiple)
 - Microphones (often multiple)
 - Digitiser Tablets
 - Now starting to handle multiple/simultaneous 'touches'
 - DiamondTouch display
 - Data gloves
 - Accelerometers.
 - Now appearing in Nokia phones...
 - Wii
 - Etc.
 - Etc.

Minority Report...

- What were examples of multi-modal interaction in Minority Report?
- Did we see examples of the issues we've talked about arising in the film?
 - False Positive for Gesture Recognition
 - Accessibility Issues?