NEW DIRECTIONS IN USER INTERFACE DESIGN AND TECHNOLOGY

-

Tapio Takala, HUT/CS/TML
http://www.cs.hut.fi/~tta/

1. Paradigm changes in user interfaces
2. Extensions of the WIMP metaphor
3. Interfaces in 3D
4. Multimodality: audio and haptics
5. Mobile computers
6. Smart environments
7. Agents and Artificial Intelligence
8. Narrative interfaces

1. Paradigm changes in UI

- Current dominant paradigm = WIMP
- Metaphors, user's conceptual models
  - spoken / written dialogue
  - spatial organization, e.g. desktop
  - coaching, process control
- Expected general trends
  - multimodality, more senses
  - new input devices
  - embedded/invisible interfaces
  - artificial intelligence
  - paradigm changes in UI software

The current dominant paradigm

WIMP = Windows - Icons - Menus - Pointers

- **Basic features**
  - 2D visual display + point & click device (mouse)
  - direct manipulation
  - hierarchical window structure
  - virtual devices (widgets)
    - menu
    - button
    - text box
    - scroll bar
  - event-based control loop
    - window manager
    - device handlers
  - gestural interaction techniques
    - drag & drop
    - selection by ”sweeping”
    - opening by double click

About UI metaphors

- **purpose**:
  - user's **mental model** of how an interactive system behaves

- common metaphors in current UIs
  - spoken / written dialogue (as between two humans)
  - spatial organization (e.g. desktop)
  - physical manipulation (WYSIWYG) on 2D screen
  - coaching, process control (e.g. games)

- what real-life features most current UI implementations do **NOT** cover
  - fragmentary conversations, with interrupts and changes of discourse
  - fuzzy concepts and indirect references to things
  - concept of time, real-time behavior (except in action games)
  - computer as a peer (personality with its own agenda and intelligence)
  - 3D space, bodily motion

Use of senses in current interfaces

	input (to computer)	*output (display)*
sight	webcam **offline images**	**UI (widgets) and content (text, images, video, etc.)**
hearing	speech recognition **content recording**	alert signals (UI) **soundtracks (content)**
smell	**?**	**?**
taste	**?**	**?**
touch	**keyboard** **mouse**	force feedback (in some game consoles)

- very asymmetric: tactile input + visual output; sound underutilised
- exception: multimedia content (sound and images in and out)
- NOTE: the fifth ”feel” sense actually covers several different senses – touch (skin), kinesthetics (limb position/force), balance (inner ear vestibular organ), and inner body state (pain, breath, hunger, temperature, etc.)

About multimodality

- **multisensory** = use of several different senses: vision, hearing, taste, smell, and different bodily senses (touch, kinesthetic, vestibular, etc.)
- **multimedia** = simultaneous use of different media types (text, graphics, audio, video, speech, etc.) – term mostly used for output (input neglected)
- **multimodality** = simultaneous use of different types of human sensorymotoric affordances (not only one for input and one for output)
  - not necessarily multisensory, nor multimedia, - mainly dealing with input
    - example: keyboard and mouse are both tactile, but used differently ➔ different modes
  - includes efferent abilities: tactile (fingers), body/limb motion, gaze, etc.
- current situation
  - dominant configuration: visual display (+marginally auditory), and tactile input in two modalities (keyboard+mouse)
- trends
  - senses: **visual** dominates, **auditory** and **haptic** (skin+kinesthetic force) developing, others (taste, smell, vestibular etc.) remain marginal for a long time
  - effectors: **tactile** dominates, **voice** and (hand)**gestural** developing, whole body motion has potential, others (gaze, EEG, etc.) still only experimental

New input devices

- current standard devices: keyboard + mouse
- available but less used:
  - pen/stylus, touch pad, joystick, track ball (replacing mouse)
  - 3D trackers (magnetic/optical/ultrasound/mechanical)
  - force/torque handles (e.g. *Trackpoint, Spaceball*)
  - speech recognition
  - obsolete: light pen, separate function buttons, thumb wheels, etc.
- new potential
  - image capture
  - non-speech sound
  - motion sensors (accelerometers, and others)
  - positioning (GPS, and others)
  - biosensors
  - chemical sensors, smell

Paradigm changes in UI software

Current

- single-threaded I/O
- discrete tokens
- precise tokens
- sequence, not time
- explicit user commands

Future

- parallel, asynchronous dialogues; may be interrelated
- continuous inputs and responses (plus discrete)
- probabilistic input, not easily tokenized
- real-time requirements, deadline-based
- passive monitoring of the user

source: Butler, Jacob & John CHI'98

2. Extensions of the WIMP metaphor

W
I
M
P

- multifocal / hyperbolic displays
- semitransparent windows
- magic lenses & toolglasses
- infinite zoom
- intelligent lay-outs, constraints
- visual programming
- two-handed input
- gesture recognition

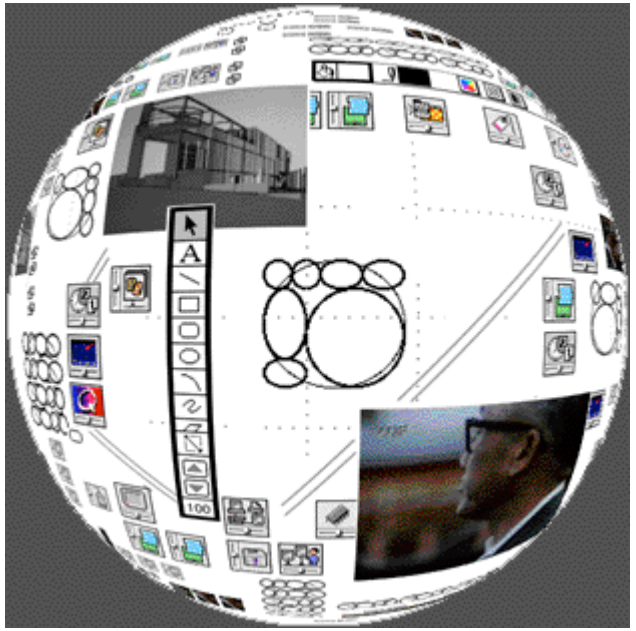
Within the same basic framework, make extensions by adding/changing functionality of the WIMP elements

Multifocal / hyperbolic displays

W
I
M
P

- *purpose*: see more data within limited display area
- **features**
  - shows important information larger on display, while keeping the surrounding space still visible
  - acts like multifocal goggles, or magnifying glass moving on the display
  - smooth animated transitions are essential to avoid distraction
- *pioneering work:*
  - *Office of the professional (SIGCHI'83)*
- *related work: perspective wall*
  - *http://www2.iicm.edu/ivis/ivis/node17.htm*
- demo application: http://ucjeps.berkeley.edu/map2.html
  - similar usage also in mindmapping tools

shortcut bar from Max OS X



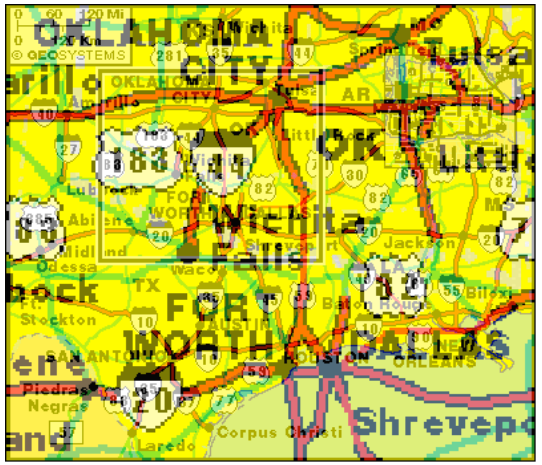
http://www.acm.org/sigchi/chi95/Electronic/documnts/shortppr/vho_fg1.gif

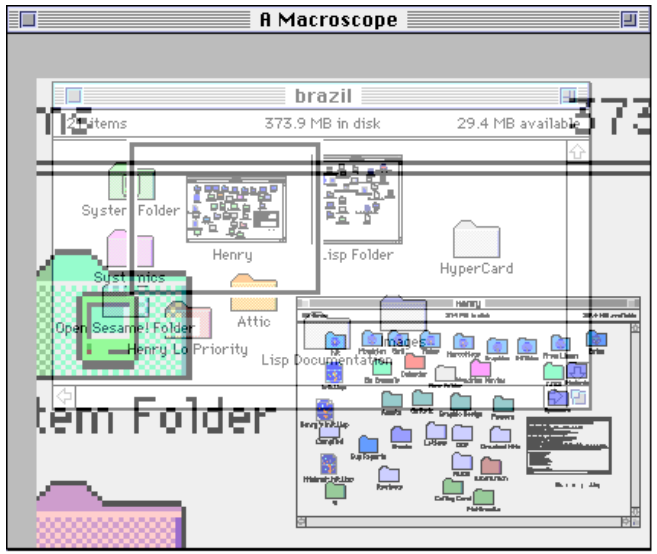
Semitransparent windows

W
I
M
P

- *purpose*:
  let the user see more than fits the screen
- **features**
  - multiple objects in same space without occlusion
  - see same object in different scales
- slow animation helps to visually separate the overlaid images
- applications
  - map reading

    http://agents.www.media.mit.edu/people/lieber/Lieberary/Macroscope/Powers-UIST.rtf
  - desktop crowd
    - Windows 2000 & XP
    - Mac OS X Aqua http://www.apple.com/macosx/whatyoucando/finder.html

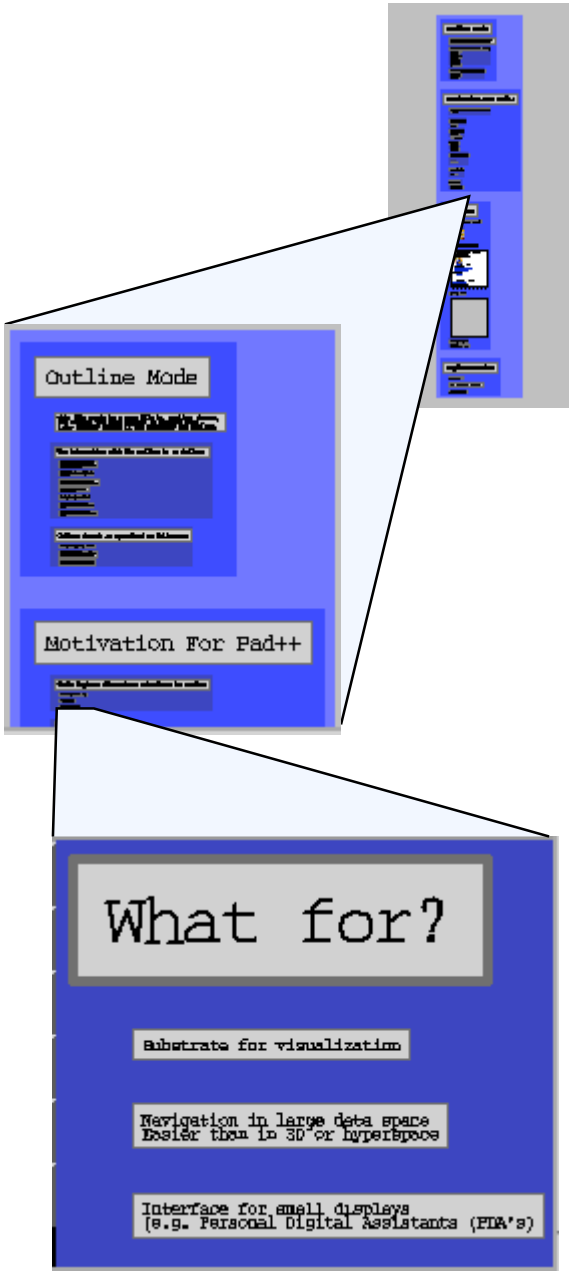




Magic lenses and toolglasses

W
I
M
P

- *purpose*: show interactively selected hidden information
- **features**
  - **magic lens** = movable area on the screen acting as filter, through which additional things or a modified view can be seen
  - **toolglass** modifies the effect of a manipulation tool
  - also studied in 3D http://www-2.cs.cmu.edu/~stage3/publications/96/conferences/uist/lenses/
- applications
  - multi-purpose maps
  - scientific visualization of dense data
  - annotations in technical documents
  - maintenance information in drawings
- *reference, including demos:*
  - http://www.parc.xerox.com/istl/projects/MagicLenses/default.html

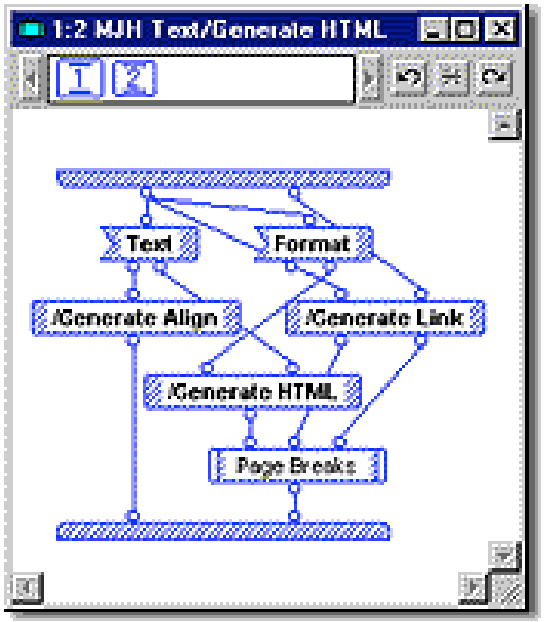


Infinite zoom

W I M P

- *purpose*:
  manage unlimited data within limited display area
- **features**
  - shows important information larger by zooming in
  - avoids the non-linear distortion of hyperbolic display
    ➔ cheaper to implement
  - smooth transitions are essential
  - no inherent limitation of the resolution of displayed data
    - works best with dynamically regenerated or multiresolution images
- main application: outline view of hierarchical structured documents
- *Pad++*
  - http://mrl.nyu.edu/projects/zui/
  - http://www.cs.umd.edu/hcil/pad++/
- Jazz
  - http://www.cs.umd.edu/hcil/jazz/play/swingJazz/swingjazz.html

Visual programming

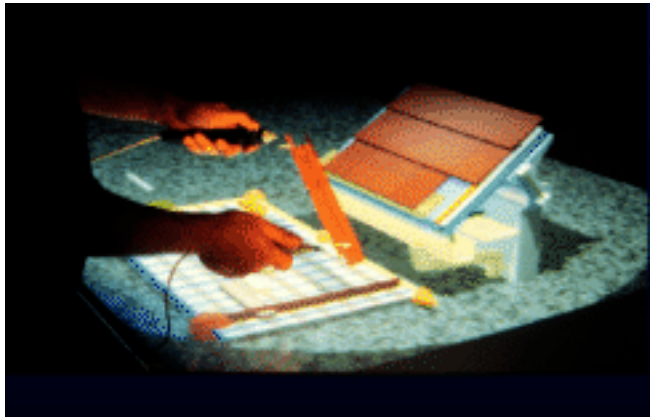


- **why programming** as a user interface?
  - avoid tedious repetitive interaction in repetitive tasks
- drawbacks of textual programming
  - not intuitive for novice users
  - typing is prone to errors
  - from the code it is not easy to understand relations of objects
  - requires naming of everything – sometimes artificially
    - e.g. loop index, used for nothing but counting the loop
- ***visual programming*** =
  describe programming concepts (data and control structures) as graphical elements (boxes, arrows, icons, spatial overlays, etc.)
- main uses
  - **dataflow** programming (image/audio processing, data visualization)
  - **spreadsheets**: pointing used to define cell references (a visual interface to textual representation)
- general references
  - http://cui.unige.ch/Visual/
  - http://www.cs.berkeley.edu/~maratb/cs263/paper/paper.html

Two-handed input

W
I
M
P

- more degrees of control by two simultaneous pointing devices
  - similar aim: stylus with tilt and pressure sensors, for more natural drawing
- applications
  - simultaneous positioning and rotation (or sizing) of an object
  - aligning objects to each other
  - specify action and selection simultaneously
    - e.g. Toolglass
  - specify action and attributes/modifyers simultaneously
    - e.g. color or brush width on one hand, drawing path on the other
- usually not supported by OS and HW drivers
  - *how about having two mice?*
- references
  - http://www.billbuxton.com/ToCHI2H.html
  - http://graphics.stanford.edu/projects/RWB/talks/iptw-97/



Gesture recognition

W
I
M
P

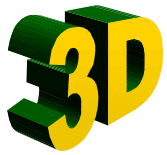
- **gesture** = movement with mouse, stylus or other device
- **features**
  - the shape (and speed) of a gesture determines its meaning
  - more than drag-and-drop (which only uses end points of gesture)
- most common uses
  - hand-drawn character recognition
  - short-hands for drawing diagram objects, or expressing commands
    - *Cocktail napkin (UIST'96)* http://depts.washington.edu/dmachine/presentations/napkin-s.mov
- other applications (non-WIMP)
  - dynamic manipulation (e.g. throwing) of physically-modeled objects
    - *Dynawall* http://www.darmstadt.gmd.de/ambiente/activities/dynawall.html
  - sign language
  - continuous control of processes (e.g. conducting music)
    - *DIVA (Siggraph'97)* http://www.tml.hut.fi/Research/DIVA/past/vsop.html
  - context recognition from sensory data
  - design of soft 3D objects: *Teddy* http://www.mtl.t.u-tokyo.ac.jp/~takeo/papers/siggraph99.pdf

3. Interfaces in 3D

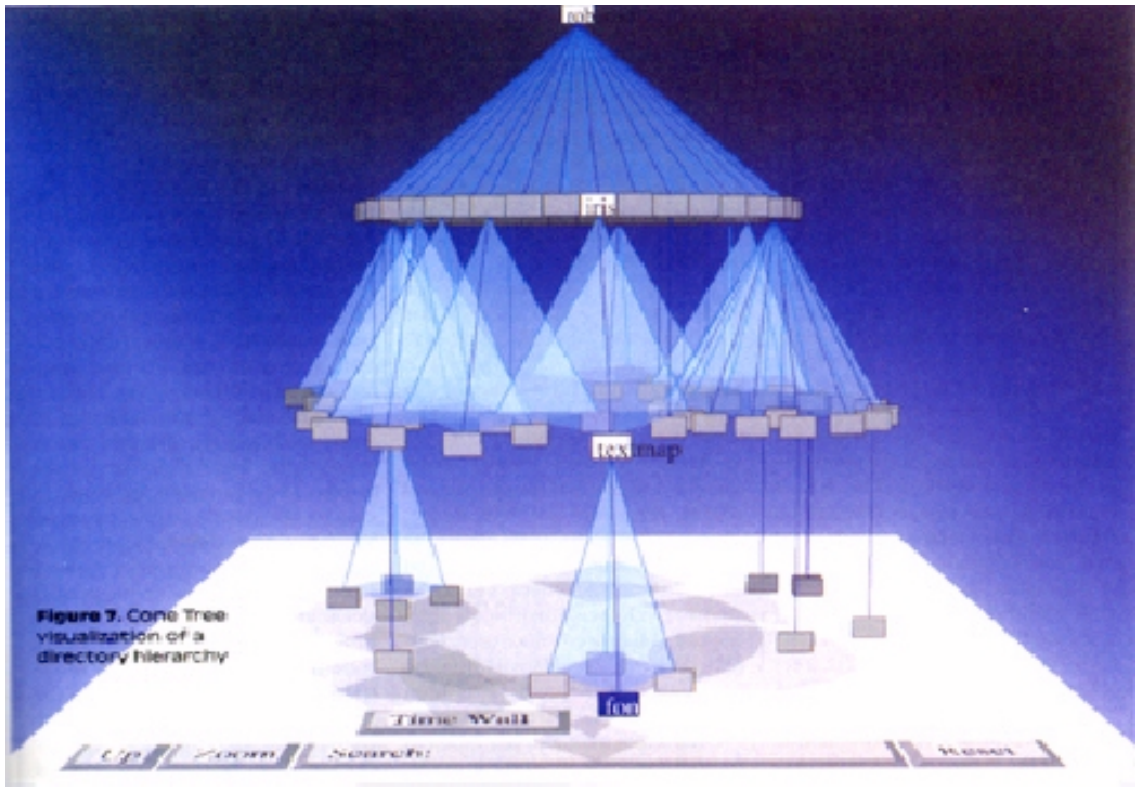
- spatial organization of information
- 3D widgets
- virtual reality = virtual world simulation + immersive UI
- augmented reality
- physical props as input devices

- **general benefits**
  - more information fits in 3D than 2D
  - often more natural, realistic mental framework
- **potential problems**
  - requires special hardware – usability problems if done ”through” 2D
  - too many degrees of freedom – rotation may cause disorientation

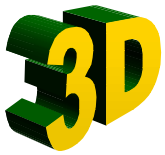
Spatial organization of information



- cone trees
  - http://www2.iicm.edu/ivis/ivis/node30.htm
  - *Fix and Float (UIST'97)* *http://www1.acm.org/pubs/articles/proceedings/uist/263407/p149-robertson/p149-robertson.pdf*
- office room metaphor http://www.cpsc.ucalgary.ca/grouplab/papers/1998/98-RoomMetaphor/report_98_611_02/roommetaphor.pdf
  - helps groupwork
- virtual spaces
  - MUDs
  - portals
  - possible platform: VRML
- **features**
  - utilize spatial memorizing
  - save screen space
  - note: hardly works without a visual map

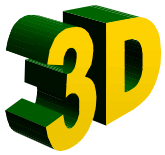


3D widgets



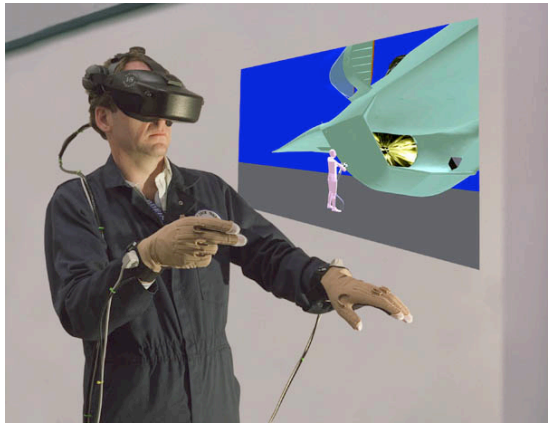
- virtual I/O devices, controlled indirectly by real devices (cf. WIMP)
- promising area, but relatively little research done
  - http://www.cs.brown.edu/research/graphics/research/widgetlib/
  - http://www.cevis.uni-bremen.de/~bernhard/papers/maerz97.html
- problems
  - transforming 2D widgets to 3D usually doesn't pay off
  - how to get appropriate feedback (esp. haptic)
  - no standard toolkits yet
- examples
  - course material http://w5.cs.uni-sb.de/~butz/teaching/sg-SS00/sg-ss00-05/sld025.htm
  - *Aperture selection (UIST'96)*

Virtual Reality I

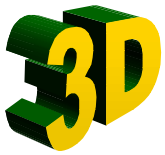


virtual reality = virtual world simulation + immersive UI

- visual display techniques
  - HMD (Head Mounted Display)
  - VMD (Virtual Model Display)
  - SID (Spatially Immersive Display)
    - *CAVE* http://evlweb.eecs.uic.edu/pape/CAVE/
- other output devices
  - audio
  - haptics
- input techniques
  - position/motion tracking
  - data glove
  - wand
  - locomotion

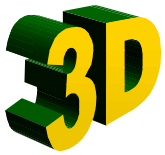


Virtual Reality II



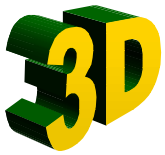
- applications
  - training simulators
  - entertainment, games
  - CAD, virtual prototyping
  - medical
  - visualization, presentation
- problems
  - great hype in early 90's – very slow success
  - expensive and cumbersome hardware
  - not so widely applicable as expected
  - interaction methodology not stabilized (see 3D widgets)
- links
  - a gallery of 3D interaction
    http://www.immersion.com/products/3d/interaction/gallery.shtml

Augmented reality
(mixed reality, enhanced reality)



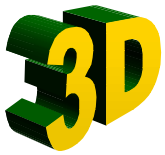
- **features**: adding virtual components on top of a real view
  - usually requires accurate spatial/directional registration
- recent survey: http://www.cs.unc.edu/~azuma/cga2001.pdf
- typical uses
  - annotations to real world (e.g. virtual graffiti, maintenance manual)
  - magic lenses (e.g. see inside patient during a surgery operation)
    - http://www.cs.unc.edu/~us/
  - navigation aids
  - real objects as markers to virtual information
    - Magic Books http://www.hitl.washington.edu/magicbook/
  - real desktop with virtual elements: *DiditalDesk*
- references:
  - link collection http://www.csl.sony.co.jp/project/ar/ref.html
  - Office of the Future http://www.cs.unc.edu/Research/stc/office/

Physical props as input devices



- techniques for virtual and augmented realities
- **features**: provide fake tactile feedback from virtual objects
  - avoid the ”ghostly” feeling of typical virtual objects
  - manipulation of virtual objects without visual attention
- references:
  - Tangible Media Group http://tangible.media.mit.edu/projects.htm
- examples
  - architecture: ”emancipated pixels” http://delivery.acm.org/10.1145/320000/311593/p385-underkoffler.pdf?key1=311593&key2=7530335101&coll=portal&dl=ACM&CFID=1553397&CFTOKEN=47504943
  - LEGO interface toolkit http://www.cs.unc.edu/~welch/class/comp239/media/pdf/interaction/lego_paper.pdf
  - real objects as markers to virtual information
  - toys as interface
- integration of physical and digital experience
  - http://www.equator.ac.uk/index.htm (research)
  - http://www.derivativeinc.com (entertainment)

3D gesture recognition



- quite little studied area
  - a sample study: http://www.crg.cs.nott.ac.uk/people/Mike.Craven/jamcon2000.pdf
  - potential research areas
    - gesture recognition in general
    - contextual inference (non-explicit gestural communication)
- application examples
  - sign language with cyberglove http://www-2.cs.cmu.edu/afs/cs/project/space/www/hmm/hmm.html
    - a resource page http://www.cse.unsw.edu.au/~waleed/gsl-rec/
  - conducting virtual orchestra http://www.tml.hut.fi/Research/DIVA/past/

4. Multimodal display

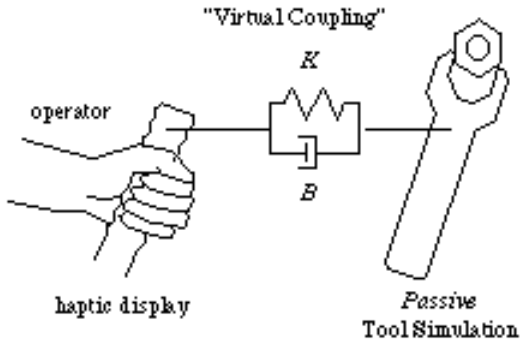
(input not considered here)

audio

- auditory display: use audio to convey information to a user
- important for visually impaired, and for situations where visual attention cannot be paid
- topics
  - speech synthesis
  - auditory icons, Earcons
  - sonification
  - spatial auralization (3D sound)

haptic

- surface touch (tactile) or force feedback to user's actions
- commonly used in power steering of vehicles



- also important for tactile devices (e.g. feeling when a button is pushed)

Speech synthesis

- **features**: convert written text to audible speech
- the process:

  - must be tuned for each language
  - short messages handled easier with recorded sound
    - example: phone answering machine
- applications
  - text reader for sight impaired
    - may be specified with document structure: http://www.w3.org/TR/REC-CSS2/aural.html
  - automatic guidance & instructions
    - vehicles, elevators, ATM
- current status: quite reliable, but dull to listen
  - reference http://www.utoronto.ca/atrc/reference/tech/textspeech.html

Auditory icons

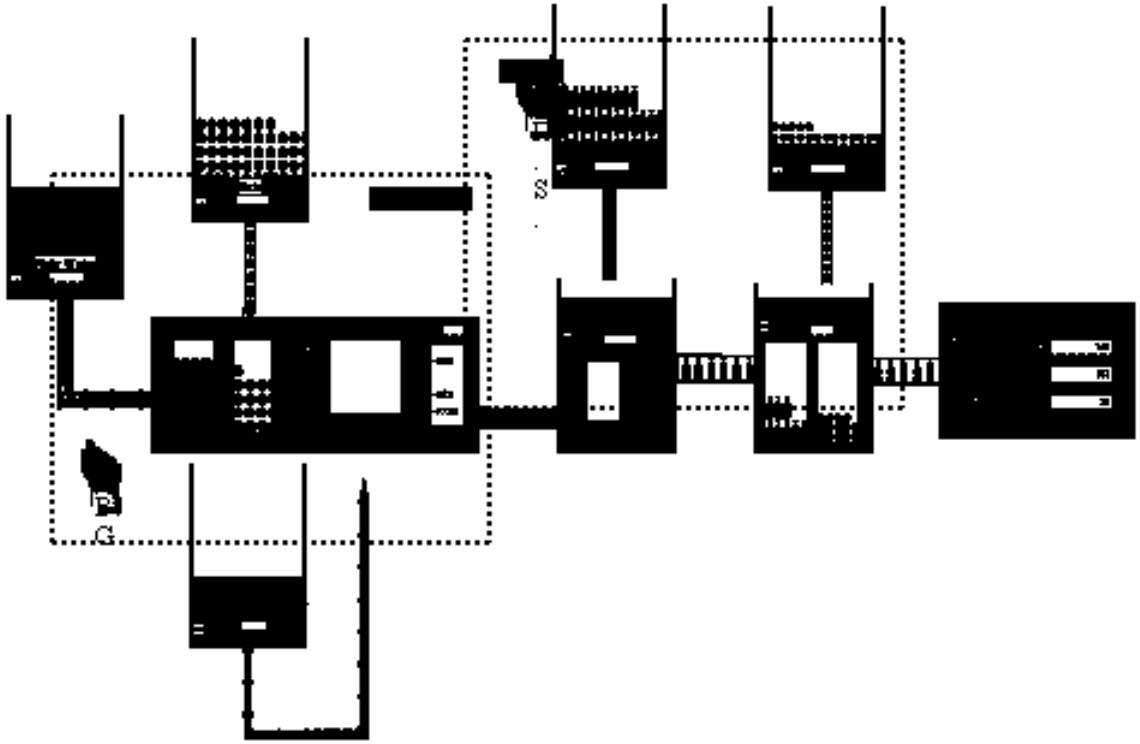
- **auditory icons**, Earcons
  = recognizable sounds used to mark an item or event
    - overview: http://www.acm.org/sigchi/chi95/Electronic/documnts/shortppr/mca_bdy.htm
  - can also carry information
    - for example, the sound marking a folder may depend on its contents
  - Earcon = auditory icon based on melodic structure
    - examples: http://www.dcs.gla.ac.uk/~stephen/earconexperiment1/earcon_expts_1.shtml
- **WARNING**: should be used with care – can be very annoying!
- applications
  - phone's ringing tone, alert sound for incoming email
  - counterpart feature of visual icons: *SonicFinder*
  - give information about out-of-sight events
  - auditory navigation through menus or information spaces
  - also used for continuous processes (cf. animated icons) ➔ see sonification
    - stationary ambient sounds are non-disturbing, abrupt changes are easily recognized

Sonification

- *sonification* = presentation of data on an auditory display (cf. visualization)
  - NSF/ICAD report 1997 http://www.icad.org/websiteV2.0/References/nsf.html
- features
  - additional display dimensions for multidimensional data
  - the ear is very sensitive to timing, detects
    - synchronization of events
    - repeating patterns (rhythm)
  - also detects spectral changes of noisy data
- sample applications
  - seismographic data (1000 x speeded playback)
  - monitoring of continuous processes
    - LA atmospheric demo
    - *ARKola*
    - *Cohen: Out to Lunch*

ARKola

CHI'91



The figure is a screen snap of the simulated bottling plant used in the ARKola experiment. Each of the shaded rectangular objects represents on of the machines in the factory. For example, the left-most one is the source of "fizzy water," and the right-most one is the "shipping and finances department." The two larger dotted rectangles illustrate what part of the factory each of the two "operators" is currently viewing. Operators can move their view independently, but neither can see the whole factory at one time. At the moment, there is little overlap between the two views. The two hands seen in the figure (lower left and upper center) are controlled by the operators' mice. Which is controlled by which operator is indicated by the initials under the hands ("B.G." and "R.S." respectively). Given that the factory was larger than the operators could satisfactorily monitor visually, the sounds made by the various machines played an important role in helping the operators detect, diagnose and repair problems. (Figure compliments of Bill Gaver, Rank Xerox EuroPARC.)

excerpt from http://www.billbuxton.com/2cultures.html

5. Mobile computers

wearable = attached to body/clothe, otherwise not in active use

- wearable devices
  - clothes
  - wristwatch
  - VR devices:
    - HMD, trackers, dataglove
- input techniques
  - touching/pulling
  - bodily motion
  - eye/gaze tracking
  - biosensors / stimulators

handheld = carry-along, maybe temporarily wearable (e.g. in pocket)

- handhelds: PDA, mobile phone
- input devices
  - tilt and g-sensors for input
  - breath controller (in music)
- potential gestural techniques
  - touch & click
  - multiuser pick and drop
  - device as a pointer in AR
  - throwing a message
  - squeezing out information
  - etc.

6. Smart environments

- Non-intrusive interaction: the environment is watching/listening you
- open (public place, traffic network) or closed (room, vehicle) environment
- Sensing technologies
  - image processing & understanding
  - voice / speech recognition
  - light/sound beams, sensing floor, etc.
- Examples
  - *Videoplace* (Krueger 1983)
  - *Put that there (SIGCHI'83)*
  - *DiditalDesk (UIST'91)*
  - *Mandala (Siggraph'94)* http://www.vividgroup.com/
  - *Holowall (UIST'97)* *http://www.csl.sony.co.jp/person/rekimoto/holowall/*
  - *Kukakumma Muumaassa* http://www.kukakumma.net/

007 Agents and Intelligence

- **agents** are daemons (= hidden background processes attending to various tasks)
  - http://www.freebsd.org/copyright/daemon.html

- level of **Artificial Intelligence** and task complexity may vary
- the more powerful agent, the more catastrophic is a malfunction!
- cooperation with human client requires discoursive knowledge about the user (wishes, intent, current activity, mood, social situation etc.)
- techniques for understanding the user
  - contextual inference
  - intention recognition
  - affective computing
- applications
  - servers
  - entertainment
  - cooperative tasks

Agent applications

- servers, intelligent web sites
  - search robots / engines
  - semantic net technology
- entertainment
  - virtual pets http://virtualpet.com/vp/
    - *Tamagochi*
    - *Fin Fin by Fujitsu:* http://www.gamezilla.com/reviews/f/finfin.asp
    - *MS parrot:* http://www.research.microsoft.com/research/ui/persona/home.htm
  - games
    - artificial enemies / competitors / supporters
  - virtual services
    - chat peer, psychiatric simulator
- cooperative tasks
  - virtual butler: automatic home control, news filtering, etc.
  - virtual me: agent as presence server, avatar

8. Narrative interfaces

- **Basic idea:**
  interaction follows a narrative story,
  which may be affected by the user
- The UI may have its own agenda, role and personality
- Little research results available, mostly just experiments
- Techniques / applications
  - scripted actions, wizards
  - show by example, what can be done
  - guided-tour help & tutorials
- Narrative experience
  - vision: *Computers as Theatre*
  - computer as narrator
  - interactive drama
  - intercative TV
  - interactive entertainment

Evaluation of UI trends

- Key trends to follow up
  - multimodality, new devices
  - gestures – both explicit and implicit
  - ambient environments, especially sound
  - emotions – as context and content
  - soft inference, instead of hard logic
  - narrative/cooperative/negotiating interfaces
- Problematic concepts
  - increasing complexity with rigid conceptual models – as now with PC
  - automatic decision-making – what if it malfunctions?
  - immersive virtual environments – requires extra hardware
  - natural language speech – too easy to not understand

Reference material

- **Books**
  - Baecker & al.: Human-Computer Interaction – Towards the Year 2000
  - Bergman: Information Appliances and Beyond
  - Carroll: Human Computer Interfaces in the New Millennium
  - Glinert: Visual Programming
  - Laurel: Computer as Theatre
  - Norman: Invisible Computer
  - Picard: Affective Computing
  - Wireless World Research Forum (WWRF): Book of Visions 2001
    - http://www.wireless-world-research.org/BoV1.0/BoV/BoV2001v1.0.pdf
- **Bibliographies**
  - *NEC research index*
    http://citeseer.nj.nec.com/HumanComputerInteraction/
  - *VRAIS '96 Tutorial: Annotated Bibliography*
    http://www.hitl.washington.edu/people/grof/VRAIS96/Bib.html
- **Conference proceedings**
  - UIST
  - CHI
  - Siggraph
  - VRAIS – VR'xx
  - ICAD
  - ISWC
  - Ubicomp
  - IUI
  - WWRF
- **Course material & tutorial articles**
  - R.Jacob: non-WIMP user interfaces, CHI'98