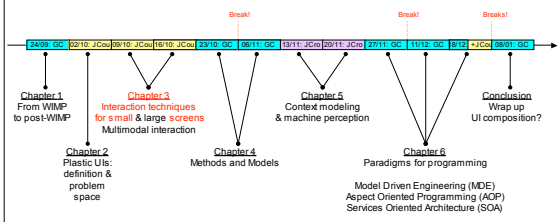
Mobile and Context-aware Interactive Systems

Gaëlle Calvary, Joëlle Coutaz and James Crowley

Université Joseph Fourier (Grenoble I)
ENSIMAG

Laboratoire d'Informatique de Grenoble (LIG)

Outline and schedule



GC: Gaëlle Calvary
JCou: Joëlle Coutaz
JCro: James Crowley

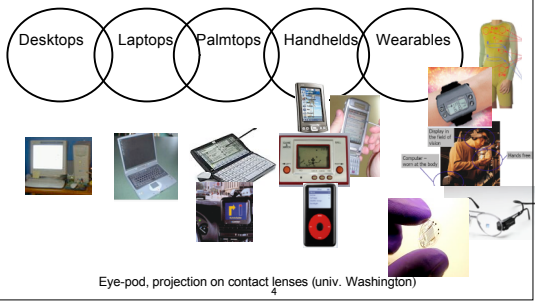
2

Outline of Chapter 3 - part 1 (mobile devices)

1. Introduction
2. Inputs
3. Screen output
4. Novel interaction techniques

3

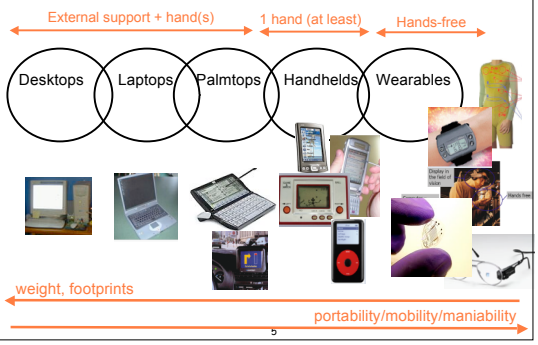
Evolution of personal computers: a continuum



Eye-pod, projection on contact lenses (univ. Washington)

4

Evolution of personal computers: a continuum



5

Handhelds: diversity, but in common …

- Small screens
- Limited input capabilities
- No standards yet
- Recommandation
  - Apply the foundations of HCI - e.g., the method
  - Create without re-inventing the wheel: draw your design on existing solutions (in order to support consistency across applications/brands) and invent new design solutions to overcome current limitations

6

Outline of Chapter 3 - part 1 (mobile devices)

1. Introduction
2. Inputs
3. Screen output
4. Novel interaction techniques

Limited input devices …

- The "choose" task: prefer selection (if the number of options is < 20) to type in (minimizes slips)
- If the number of options is large (>20)
  - Structure the list of options into meaningful sub-lists
    - For example, use the centrality (importance) of the domain-dependent concepts
    - Attention: the user may get lost when navigating within large/deep trees (use the bread-crumb trail technique)

Structured alphab. List
on a very small mobile phone screen

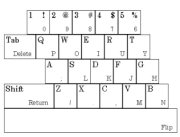
- Improve the techniques for text entry …

The Ipod contacts list with the alphabet as a local navigation bar

Keyboard text entry

- External physical keyboard for PDA's (foldable)
  - Comfortable
  - Inadequate when on the move
  - Keyboard-stylus homing made difficult

- External physical half-keyboard: "Space bar" to switch between the two half-keyboards
  - Minimizes the footprint
  - Adequate for experts with one hand available (e.g., cash-clerks)



Keyboard text entry

- Projected keyboard
  - Technique
    - Laser projection
    - Finger localization with infra-red camera
  - Requires a surface: inadequate when on the move
  - No kinesthetic feedback
  - Sound modality to simulate the key down click

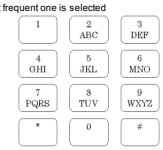
10

Keyboard text entry

- Built-in keyboard, but small (hard for elderlies)

  - 1 character per key, but then the keys are very small (Palmtops
    - indice of difficulty is high (cf. Fitts' Law, last slide)
  - Multiple characters per key
    - Temporal constraints+multiple K (mobile phones)
    - Increases time (cf. KLM)
  - Predictive techniques for word completion (guess words to reduce the number of keys to type in)
    - the T9 system: 1 keystroke per character, 0 used as a separator
    - prediction based on the knowledge of the language used
    - for ambiguities (several words are good candidates), the most frequent one is selected

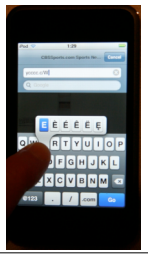
```
843 78425 27696 369 58677 6837 843 5299 364
the quick brown fox jumps over the jazz dog
tie stick crown     lumps muds tie lazy fog
vie                         vie
```



11

Text entry with stylus/finger combined with tactile screen

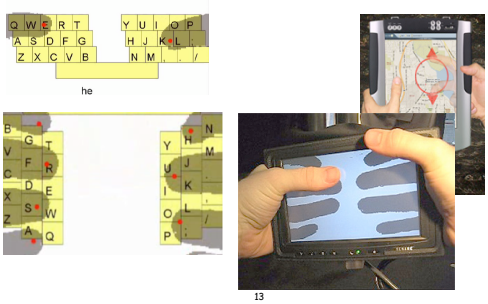
- Soft keyboard
  - Small footprint but uses screen real estate. Semi-transparency can help
  - Looking for the stylus is awkward for quick answers
  - The finger hides the target
    (see "Shift" for improvements, and the Ipod Touch)



12

Text entry: prototype …

- LucidTouch: a see-through mobile device, pseudo-tranparency



13

Text entry: prototype …

- LucidTouch: a see-through mobile device, pseudo-tranparency

14

Text entry using hand-writing recognition

- Cursive hand-writing recognition (as opposed to character recognition)
  - Seminal work: The Apple Newton (1993-1998)
    - Too demanding in terms of computing resources -> too slow
    - Too long, the learning phase -> users would lost patience
  - Available on some smart phones

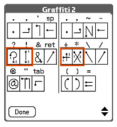
The precursor : the Apple Newton

15

Text entry using hand-writing recognition

- Unistrokes
  - The user is faster than writing the standard characters
  - Robust system recognition
  - Precursor: Xerox PARC Unistrokes (1993)
  - Commercialization: Palm Graffiti

http://www.palm.com/us/products/input/graffiti2.html



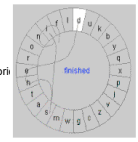
Capitals
To make a capital, just write the letter in the middle of the Graffiti 2 writing area.

Try it! Graffiti 2 Demo

Graffiti 1

Text entry based on gesture and circular menus

- "Guided" gesture
- Principles: minimize trajectories (cf. Fitts'Law)
  - 1 stroke per word (instead of 1 stroke per character)
  - Circular soft-keyboard
  - Characters layout depends on the language: shorter trajectori
- Efficient for experts (those who have remembered the gesture)



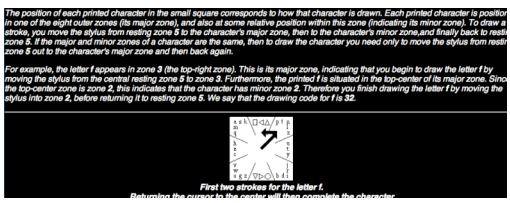
[*Cirrin*, Mankoff-Abowd 98, For writing "finished"]

[*Quickwriting*, Perlin 98, for writing "quik", see details next slide]

17

Text entry based on gesture and circular menus

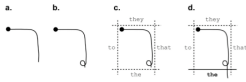
- Quickwriting: detailed explanation (many improvements since then)
  - http://mrl.nyu.edu/~perlin/demos/Quikwrite1.html



18

Text entry based on gesture and circular menus

- Fisch (Fluid in-stroke completion shorthand): from characters to words



In Figure 1a, the user has written a Graffiti unistroke for the letter "t". If the user were to lift his stylus at this point, a "t" would be produced. However, in Figure 1b, the user continues the stroke by making a pigtail loop. In Figure 1c, the system detects this loop and imposes an appropriately sized bounding box around the stroke. The "t" is recognized and the system presents the four most common words beginning with "t" at each side of the box. The sides on which the words appear are fixed such that the same word always appears on the same side for a given stroke, allowing users to reliably enter words in single strokes. In Figure 1d, the user fluidly continues the gesture to perform a crossing task, which selects the word "the" across the penetrated boundary. Studies have shown crossing to be faster than pointing for short-range selection tasks [1]. Thus, when the user lifts, the word "the" is entered. Importantly, the same stroke always produces the word "the".

19

Text entry based on gesture

- SHARK (SHorthand Aided Rapid Keyboarding]

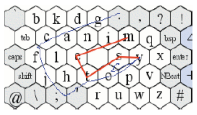


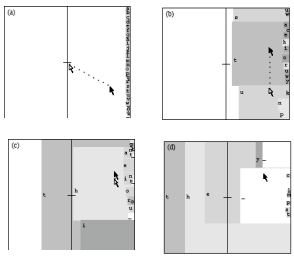
Figure 1. A user has written the word "system" in SHARK^2 on a 57K lexicon. As is evident in the figure, the system can handle gestures that are far away from the ideal sokgraph (marked in bold), both in terms of overall shape and location proximity. Shown is the ATOMIK layout. Sokgraphs can be defined on any layout, such as QWERTY.

[Per-Ola Kristensson, Shumin Zhai, UIST2006]

20

Text entry based on a mix of prediction and pointing

- Dasher: prediction and pointing (pen, mouse, gaze for motor-impaired users)



[*Dasher*, Ward 00, to write "the"]

21

Text entry based on a mix of prediction and pointing

- *Dasher* in action

22

Pointing techniques

- "Offset Cursor" [Potter, Shneiderman 1988)]: how to replace the stylus with the finger without hiding the target (See LucidTouch based on pseudo-transparency as an alternative)
- Selection technique: finger touches the screen, display the cursor at 1/2 inch over the finger, move the finger to bring the cursor over the target, remove finger

Target

Finger hides the target
Cursor is offset from the target

Finger moves to bring the cursor over the target

23

Pointing techniques

- Shift [P. Baudisch 07, MicroSoft]: improves the "offset cursor" technique

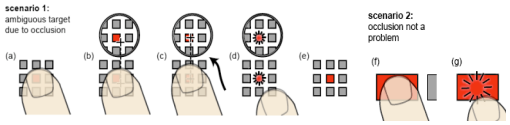


Figure 2. Shift technique walkthrough. (a-e) Scenario 1, ambiguous target selection due to occlusion: (a) on contact, Shift determines if occlusion is a problem for targets under the finger; (b) Shift responds by displaying a callout containing a copy of the occluded area with a pointer showing the finger selection point; (c) keeping the finger on the display, the user makes corrective movements until the pointer is over the target; (d) lifting the finger selects the target; and (e) removes the callout. (f-g) Scenario 2: (f) when occlusion is not a problem (g) Shift does not "escalate" and instead behaves like a regular, unmodified touch screen.

24

Pointing techniques

- Shift [P. Baudisch CHI07, MicroSoft] in action

25

Pointing techniques

- "Jump and Refine" [Hachet et al. Bordeaux, CHI 2007]: based on a two-level resolution grid (cells)
- Adequate for selecting objects in 3D
- But needs extra clicks for notifications (the initial trigger, grid-level change, final selection)

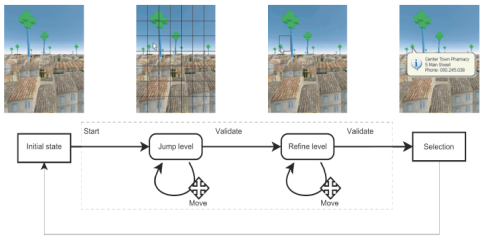


Figure 2. The *Jump and Refine* workflow: the selection process is called by means of an assigned key. In the *jump* level, the cursor is moved from one cell to another. A first "validate" key pressure leads to the *refine* level where the cursor can be accurately positioned. A second pressure generates a "click event" with the position of the cursor as coordinates.

Audio pointing (eyes-free/blind)

Figure 1. Using earPod. (a, b) Sliding the thumb on the circular touchpad allows discovery of menu items; (c) the desired item is selected by lifting the thumb; (d) faster finger motions cause partial playback of audio. Size of the touchpad has been exaggerated for illustration purposes.

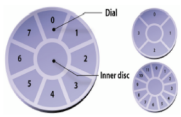


Figure 3. The functional areas of earPod's touchpad. Up to 12 menu items can be mapped to the track. The inner disc is used for canceling a selection.

27

Outline of Chapter 3 - part 1 (mobile devices)

1. Introduction
2. Inputs
3. Screen output
4. Novel interaction techniques

Screen: the starting point (more in Chapter 4)

- Use the task model and the Abstract UI (AUI) that define
  - The overall structure of the Concrete UI (CUI): navigation and the visual structure of the display information space)
  - The domain-dependent concepts that must be made observable
- Identify the central tasks
  - They are the "raison d'être" of the system or they are frequent
  - They make sense on the targeted mobile device
- For each task
  - Identifiy the central domain-dependent concepts
  - i.e. identify those that are necessary (right away observable) from those that are less necessary (possibly browsable)
- Within the CUI, the user must be able to:
  - accomplish the central tasks
  - and observe/manipulate the central domain-dependent concepts
- Then, your CUI can exploit the following interaction techniques …

Control widgets: general rules

- Icons
  - NB: nouns are easier to iconify than verbs
  - Requirement: samll size but recognizable
  - Reuse conventions, no hieroglyph!
  - Avoid text within an icon (next to it, yes)
  - Use contrasted colors: in the sunlight, colors fade away
  - Outline icons with a thin dark line

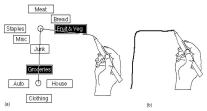
Control widgets: general rules

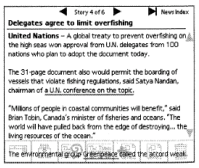
- Location of buttons (to be contextualized)
  - At the bottom of the screen, left justified, if possible (becase at the top, you hide the content of the screen with the hand)
  - Ordering: frequency, logical (semantic) ordering, alphabetical ordering
  - Cancel: on the right

31

How to minimize the footprint of control widgets ?

- Semi-transparency
- Marking menus [Kurtenback 93]
  - Circular shape (15% faster than the linear menu, 8 items at most)
  - The menu pops-up if the cursor stays still for approximately one second
  - If the cursor is moved immediately, then the system interprets the stroke (gesture): appropriate for the expert user





32

Fisheye technique

- Allows the user to see the details without loosing the context of the details (Furnas: seminal work)

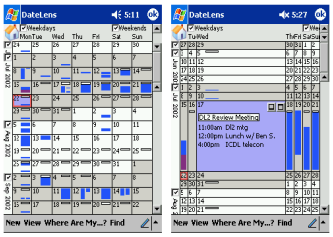
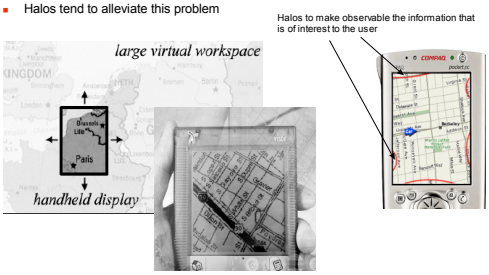


Figure 1: DateLens with the view configured to show 12 weeks (left). The right view shows the result of tapping on July 17th which focuses on that day.

33

The Peep Hole technique

- A peep hole on a large information space [Ka-Ping Yee, CHI2003]
- Problem: loss of the context -
- Halos tend to alleviate this problem



Halos to make observable the information that is of interest to the user

34

Textual summaries of pages

- Source pages are transformed by a proxy into a hierarchy of STU (Semantic Text Unit)
- Technique
  - Lists, paragraphes, images Alt tags
  - Fonts and hierarchical structuration of pages
- The system displays the hierarchy (2 to 3 level deep)
  1 STU per line and the user browses the hierarchy
- Markers are used as hints for navigating
- However, does not preserve the layout of the source page

Figure 2: An STU Progressively Displayed in Three States

STU opened (-)

STU not Opened yet (+)

To display an STU step by step (•)

Buyukkokten 00 - Stanford

35

Graphical summaries (thumbnails)

- Thumbnail version of the source web page: preserves the layout of the source page => from the user's perspective : consistency
- The browser is augmented with new tools

RSVP: Rapid Serial Visualization Presentation for reading unlimited text in a limited space

Browser with new tools

Thumbnail of the page

Extracted elements

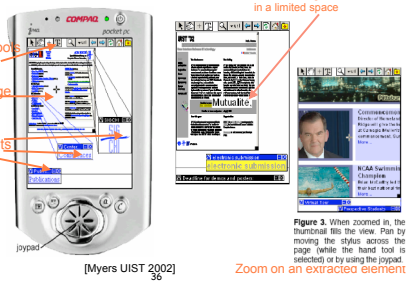


Figure 3. When zoomed in, the thumbnail fills the view. Pan by moving the stylus across the page (while the hand tool is selected) or by using the joypad.

[Myers UIST 2002]

Zoom on an extracted element

36

Graphical summaries (thumbnails)

- RSVP metrics per word for a good pacing
  - Base duration: 150 ms
  - + 25ms per letter
  - +40 ms for a capitalized word
  - 310 ms for a comma
  - 420 ms for a period
  - Other punctuation: 25 to 40 ms

RSVP: Rapid Serial Visualization
Presentation for reading unlimited text
in a limited space

[Myers UIST 2002]

Graphical summaries (thumbnails)

- Problem with thumbnails
  - Text is hard to read
  - Additional articulatory tasks for zooming in and zooming out

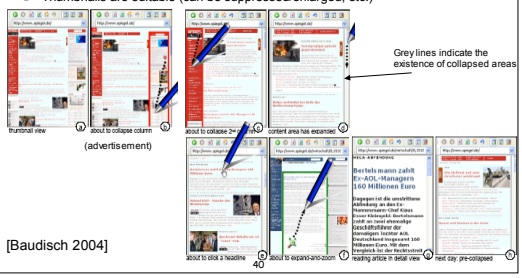
Improved thumbnails

- Summary thumbnails
  - Page layout is preserved
  - Texts are replaced with a summary, thus can be displayed using bigger fonts within their original space
  - Semantic zoom

[Baudisch CHI 2005]

Hybrid technique: thumbnail+editing

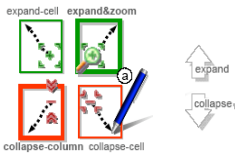
- Collapse-to-Zoom
  - Pages that are to big to fit the screen are presented as a "summary thumbnail"
  - Thumbnails are editable (can be suppressed/enlarged, etc.)



[Baudisch 2004]

Hybrid technique: thumbnail+editing

- Collapse-to-Zoom, interacion technique: marquee menu
  - markee menu // selection of the entity to be manipulated
  - 4 commands



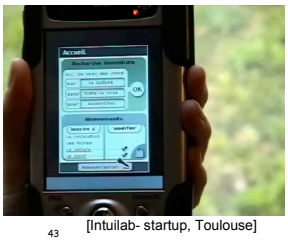
[Baudisch 2004]

Outline of Chapter 3 - part 1 (mobile devices)

1. Introduction
2. Inputs
3. Screen output
4. Novel interaction techniques

Multimodal user interface

- Text input using a combination of speech and stylus selection
  - No keyboard
  - Requires a robust speech recognition system
  - Network bandwidth may be limited: signal preprocessing on the PDA



[Intuilab- startup, Toulouse]

43

Movement-based UI - 3D gesture

- TinyMotion: capture the movement of the mobile phone with the camera
  - Contraints: static background and stable lighting conditions
  - Alternative: accelerometers

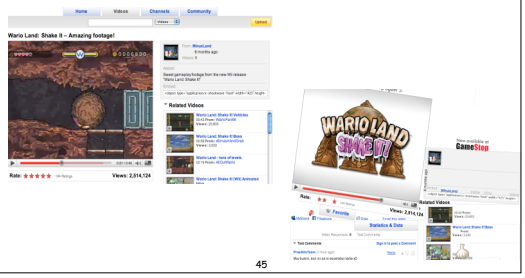
[*TinyMotion*, Wand, Zhai, Canny, UIST2006]

Figure 2: Sample *TinyMotion* applications and games. From left to right, top to bottom – Motion Menu, Image/Map Viewer, Mobile Gesture, Camera Tetris, Camera Snake and Camera BreakOut

44

Movement-based UI - 3D gesture

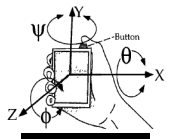
- Ipod: accelerometers and orientation of the screen
- The Wii mote: http://www.youtube.com/experiencewii (implem. flash)



45

Movement-based UI - 3D gesture

- Tilting using accelerometers [Rekimoto 1996, Harrison&Fishkin 1998]
- Pb: legibility of the contrast when tilting forward => need techniques for visual compensation



Photos album

Cylinder-shape menu

46

Movement-based UI and audio&tactile feedback

- Shoogle senses data (and makes sense of them!) [Univ. Glasgow 07]
- Sensors: the Shake includes an accelerometer, gyroscope, magnetometer, vibro-tactile display communicating via bluetooth
- Expresses the existence of messages, remaining resources (e.g., the battery level) using a dynamic model inspired from mechanics (balls anchored with springs bouncing against walls)

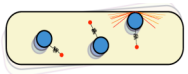


Figure 3. The simulated system. A number of balls, anchored via springs, bounce around within the virtual container. When they impact (as in the top right) sound and vibration are generated.

47

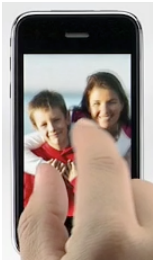
Movement-based UI and audio&tactile feedback

- Shoogle [Univ. Glasgow 07] in action

48

Two-handed interaction

- Ipod touch, iPhone



49

Two-handed interaction

- RodDirect: the stylus as a peripheral device (2 degrees of freedom) + the second hand
- The user rotates and slides the stylus to operate handhelds [Miura UIST2005]

Two-handed interaction

- RodDirect in action

51

actuators and feedback

- Vibro-tactile feedback: rather usual
- Shape-change feedback: the dynamic knob [Deutsche Telecom et Univ. Postdam]

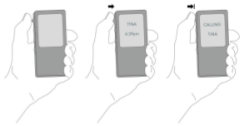


Fig. 2: Checking for missed calls through feeling the extensions on the phone's side

ComTouch [Chang]

52

Distorsion: bendable display

- The Gummi interface: just bend it! (Sony, Schwesig CHI 2004)

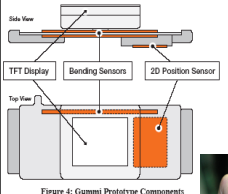
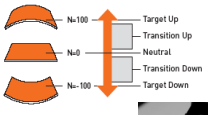


Figure 4: Gummi Prototype Components



53

Distortion: bendable display

- Gummi: zoom-in a subway station

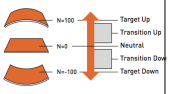


Figure 8: Analog Link

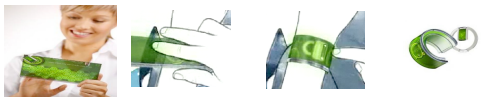
- Alpha-blending of 2 graphics layers

Figure 9: Map Blending

54

Distorsion: the possibilities of nano-technologies

- The Morph: a fully polymorphic phone (Nokia and Univ. de Cambridge)
  - Flexible material, controlable transparency, and self-cleaning
  - On the market in ten years from now



55

Distributed UI: coupling a mobile device with a large screen

- SpotCode (Univ. Cambridge), Bluetooth connection, circular tags, video camera and 3D gesture
  - Browsing, zoom-in, zoom-out

56

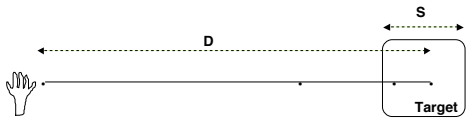
Nouvelles formes d'interaction : Couplage avec grand écran

- Application de SpotCode en action

57

Fitts'Law

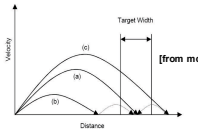
- Measures the duration (time) necessary to select a target of size S located at a distance D from the hand
  - $T = a + I.Log_2 (2D/S)$
  - I denotes the bandwidth (or IP = indice of performance) of the user = time to transmit one bit of information = 100 ms/bit (means)
  - $ID = Log_2 (2D/S)$ is called the indice of difficulty (expressed in bits)



58

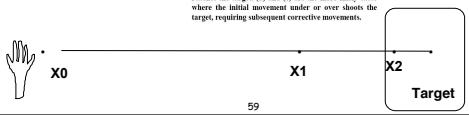
Fitts'Law

- Explanations of Fitts'Law
  - Closed-loop model: a sequence of micro-mouvements towards the target: each movement brings the hand closer to the target, feedback for the next micro-movement (iterative correction)
  - Variable impulse model
    - Initial muscular impulse
    - Followed by docking
  - Mixed model (cf. picture)
    - Optimized initial impulse
    - Subsequent corrective movement



[from mcguffin CHI 2002]

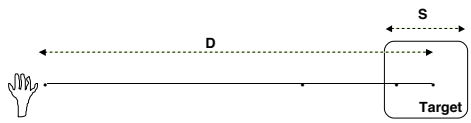
Figure 1. Possible sequence(s) of submovements toward a target as described by the *optimized initial impulse* model [12]. (a) is the case where a single movement reaches the target. (b) and (c) are the more likely cases where the initial movement under or over shoots the target, requiring subsequent corrective movements.



59

Fitts'Law

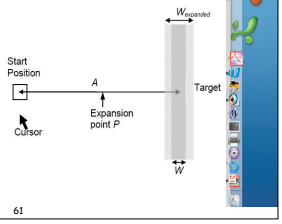
- Application of Fitts'Law: reduce pointing time
  - By increasing the size of the targets
  - By reducing distances



60

Fitts'Law

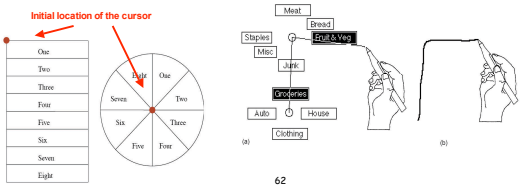
- Increase the size of the targets in the visual space
  - uses pixels =>
    - Exploit widget transparency
    - Change the size of the target in a transient manner (or a variation of it: add hooks to small targets)
- Examples of transient size changes
  - Experiments of McGuffin& Balakrishnan [CHI 2002]
    - Increase the size of the widgets close to the mouse cursor
    - OK for insulated widgets
    - Occlusion if several widgets close to each other
  - Technique fish-eye du dock de MacOS X
    - Pb of the moving target: inside the dock, the widgets are enlarged visually whereas in the motor space, the sizes are kept unchanged => no advantage for selecting + risk of overshooting



61

Fitts'Law

- Reduce the distances in the motor space
- Examples
  - Circular menus / linear menus (8 items) : 15% faster (and less errors) [Callahan 88]
  - Extension: the marking menus [Kurtenback, Buxton 1993]
    - For non-experts: the menu pops up if the cursor stays still for some milliseconds
    - For experts: if the cursor moves right away, the menu does not pop up but the system interprets the gesture



62

Fitts'Law

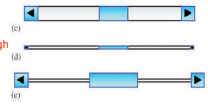
- Reduce the distances in the motor space (continued)
- Other examples
  - Exploitation of the CD ratio (Control to Display ratio)
    - Ratio between the movement of the mouse and that of the cursor (#cm/#pixels)
    - The CD ratio is modified so that an acceleration of the mouse results in a larger movement of the cursor
  - Elimination of empty spaces
    - 75% faster for non-cluttered screens
  - Sticky targets [Cockburn 2003]
    - Contrains the cursor close to a target
    - OK for non-cluttered screens

63

Fitts'Law

- Reduce the targets in the visual space, but enlarge the targets in the motor space
- Example: semantic pointing [R. Blanch 2005]

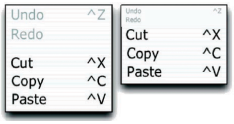
The original scroll bar



The scroll bar in the visual space : 3 pixels wide are enough

Scroll-bar in the motor space: the key targets have the same size in the motor space as in the original visual space

Aletr box : top, original version in the visual space. Bottom, in the motor space



Application to menus: right, the menu in the motor space

64