Mobile Input & Output Technologies

Mensch-Maschine-Interaktion 2, WS 2010/2011

Michael Rohs
michael.rohs@ifi.lmu.de
MHCI Lab, LMU München
# Lectures & Exercises

<table>
<thead>
<tr>
<th>Lecture</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12.1.</td>
<td>Mobile Device Platforms</td>
</tr>
<tr>
<td></td>
<td>19.1.</td>
<td>Introduction to Mobile Interaction</td>
</tr>
<tr>
<td></td>
<td>26.1.</td>
<td>Prototyping and Evaluation of Mobile Systems</td>
</tr>
<tr>
<td></td>
<td>326.1.</td>
<td>Prototyping and Evaluation of Mobile Systems</td>
</tr>
<tr>
<td></td>
<td>42.2.</td>
<td>Mobile Input &amp; Output Technologies</td>
</tr>
<tr>
<td></td>
<td>59.2.</td>
<td>Location &amp; Context, UI Design for Small Displays</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>Developing countries + Android-Eclipse</td>
</tr>
<tr>
<td>10.1.</td>
<td></td>
<td>Recipe input</td>
</tr>
<tr>
<td>17.1.</td>
<td></td>
<td>Touch input, gestures</td>
</tr>
<tr>
<td>24.1.</td>
<td></td>
<td>Evaluation of mobile LMU Web portal</td>
</tr>
<tr>
<td>31.1.</td>
<td></td>
<td>Location-based audio</td>
</tr>
</tbody>
</table>
Mobile Text Entry

Partly based on slides by Scott MacKenzie:
Text Entry on Mobile Devices

- Mobile text entry is huge
  - SMS (>2.5 billion users; 4.1 billion SMSs each day, US, 2009)
  - Email, calendars, notes, passwords, etc.

- Small devices require alternative input methods
  - Smaller keyboards, stylus input, finger input, gestures

- Many text entry methods exist
  - Companies are ambitiously searching for improvements

Source: http://digitaldaily.allthingsd.com/20091008/omfg-4-1-billion-text-messages-sent-every-day-in-us/

Key-based

Finger-based

Stylus-based

Tilt-based
Text Entry Speed on Mobile Devices

• Goal: High-speed entry at low error rates
  – Movement minimization
  – Low attention demand
  – Low cognitive demand

• Entry speeds depend on task type and practice

• Typical text entry speeds
  – Handwriting speeds: 13-22 words per minute (wpm)
  – Desktop touch typing: 60+ wpm
  – Soft (on-screen) keyboards:
    40+ wpm after lots of practice,
    typically 18-28 wpm for qwerty,
    5-7 wpm for unfamiliar layout
Keyboard Layouts for Mobile Devices

• Querty variations
  – Querty designed to be slow
  – Prevented typing machines from jamming
    • alternate between sides of the keyboard

Dvorak Keyboard

• **Speed typing by**
  - Maximizing home row (where fingers rest)
  - Alternate hand typing
• Most frequent letters and digraphs easiest to type
Fitaly and Opti Keyboards

- Designed for stylus input on soft (on-screen) keyboards
- Minimizing stylus movement during text entry
- Stylus movement for entering the ten most and least frequent digrams:

Half-Qwerty and ABC Keyboards

- **Half-qwerty**
  - One-handed operation
  - 30 wpm

- **ABC keyboards**
  - Familiar arrangement
  - Non-qwerty shape

Very Small Devices

- 5 keys (e.g., pager)

- 3 keys (e.g., watch)

Keyboards and Ambiguity

- Keyboard miniaturization: Smaller keys, Less keys
- Unambiguous keyboards
  - One key, one character
- Ambiguous keyboards
  - One key, many characters
  - Disambiguation methods (manually driven, semiautomatic)

Ambiguity

- Ambiguity occurs if fewer keys than symbols in the language
- Disambiguation needed to select intended letter from possibilities
- Typical example: Phone keypad

![Keyboard Layout]

Unambiguous Keyboards

• One key, one character

• FasTap keyboard
  – Keys in space between keys
  – 9.3 wpm
Ambiguous Keyboards

- One key, many characters
- Standard 12-button phone keyboard, larger variants

Twiddler, chord keyboard

Nokia N73

Blackberry 7100
Manual Disambiguation

• Consecutive disambiguation
  – Press key, then disambiguate
  – Example: Multitap
    • Disambiguating presses on same key (timeout or timeout kill)

• Concurrent disambiguation
  – Disambiguate while pressing key (via tilting or chord)
  – Example: Tilting
    • Tilt in a certain direction while pressing
  – Example: Chord-keyboard on rear of device
    • Not widely used
Disambiguation by Multitap

RUNNER = 7778866n6633777
RUNNER

SUMMER = 7777886n633777
SUMMER

STONES = 7778666N66337777
STONES

“n” = next character on key

TiltType, Univ. Washington

- Text input method for watches or pagers
- Press and hold button while tilting device
- 9 tilting directions (corners + edges)
- Buttons select to character set

Kurt Partridge et al.: TiltType: Accelerometer-Supported Text Entry for Very Small Devices. UIST 2002 technote
portolano.cs.washington.edu/projects/tilttype
Dictionary-Based Disambiguation (T9)

- Term frequency stored in dictionary

- Most frequent possibility presented first

- “n” = key for next frequent possibility

RUNNER = 786637nn
RUNNE R

SUMMER = 786637
SUMMER

STONES = 786637n
STONE S

Simplified Handwriting: Unistroke

- Single-stroke handwriting recognition
  - Each letter is a single stroke, simple recognition
  - Users have to learn the strokes
  - “Graffiti” intuitive unistroke alphabet (5 min practice: 97% accuracy)

- Slow (15 wpm)
- Users have to attend to and respond to recognition process
- Recognition constrains variability of writing styles
Unipad: Language-Based Acceleration for Unistroke

• Speeding up stylus-based text entry
  – Eyes-free entry possible for unistroke
  – Look at suggestions during eyes-free unistrokes

• Language-based acceleration techniques
  – Word completion list based on corpus (word, frequency)
    • Tap candidate
  – Frequent word prompting ("for", "the", "you", "and", etc.)
    • Tap frequent word
  – Suffix completion based on suffix list ("ing", "ness", "ly", etc.)
    • Top-left to bottom-right stroke, tap suffix

Unipad: Acceleration by Word Completion

• Word completion example
  – User is entering word “hours”
  – State after two strokes (“ho”)

• Experimental interface
  – First line shows text to enter
  – Second line shows text already entered
  – Pad below
    • Entering strokes
    • Word completion list

Unipad: Acceleration by Frequent Word

- Frequent word example
  - User is about to enter “of”
- Pad shows frequent word list
  - User taps “of”

Unipad: Acceleration by Suffix Completion

• Suffix completion example
  – User is entering “parking”
  – State after 4 strokes (“park”)

• Pad shows word completion list
  – User enters top-left to bottom-right stroke to show suffix list

• Pad shows suffix list
  – User taps “ing”

http://www.yorku.ca/mack/nordichi2006.html
Unipad: Performance

- Entry speed >40 wpm possible
  - KSPC ≈ 0.5 (key strokes per character)
- Expert performance simulated on sentence
  “the quick brown fox jumps over the lazy dog” (43 chars)

(27 strokes)

http://www.yorku.ca/mack/nordichi2006.html
**EdgeWrite**

- Provide physical constraints
- Moving stylus along edges and diagonals of square input area
- People with motor impairments
- Input = Sequence of visited corners

**Example: Digits**

QuickWriting: Gesture-Based Input

- Combine visual keyboards with stylus movements
- Following a path through letters of the word to enter
- Motor memory for paths
- Reduced stress and fatigue compared to tapping

- Ken Perlin: Quickwriting: Continuous Stylus-based Text Entry. UIST’98.

Quickwriting, http://mrl.nyu.edu/~perlin/demos/Quikwrite2_0.html
Swype

- Text entry via continuous swipes, lifting between words
- Guesses most likely word from language model
- Manual disambiguation possible
- Example: entering the word “quick”:

  ![Swype keyboard image]

  - World record text message: 26 words typed in 25.94s
Touch Screen Gestures

Source: GestureWorks.com
Which Gestures are These?

• Hint: one is “flick” and one is “drag”

• Relevant gesture parameters
  – Velocity profile
  – Shape
  – Direction
And this one?

• Multi-touch pinch inwards
  – Typically mapped to “zoom out”

• Relevant gesture parameters
  – Number of touch points
  – Shape
  – Direction

• Challenge: finding intuitive mappings
  – Who should do this?
Gesture Usage

• **Letter and digit recognizer**
  – Fixed gesture set
  – E.g., based on neural network classifier
  – Trained on large corpus of collected data

• **User-customizable recognizer**
  – Typically template based
  – Nearest-neighbor matching

• **Usage**
  – **Shortcuts to frequent content**
    • Contacts
    • Applications
    • Functionality: “take me home home”
  – **Gesture location = operand, gesture shape = operation**
    • Annotations, editing marks
Example Application: Gesture Search

• Find items on Android phones
  – Contacts, applications, songs, bookmarks
  – Drawing alphabet gestures
• http://gesturesearch.googlelabs.com

Recognition of Touch Screen Gestures

• Touch screens on many mobile devices
  – Mostly used for tapping (pointing tasks)
  – Suitable for swiping (crossing tasks)
  – Suitable for entering complex gestures

• Gesture recognition challenging
  – Pattern matching, machine learning

• Approaches for simple UI prototyping
  – $1 Recognizer
    • http://depts.washington.edu/aimgroup/proj/dollar/
  – Protractor
    • http://yanglisite.net
Recognition of User-Defined Touch Screen Gestures

• Template-based recognizers
  – Template preserves shape and sequence of training gesture
  – Nearest neighbor approach

• Process
  – Store training samples as templates (multiple templates per gesture)
  – Compare unknown gesture against templates
  – Choose class of most similar template

• Advantages
  – Purely data-driven, customizable (no assumed underlying model)
  – Small number of examples per class sufficient

• Disadvantages
  – Comparison with all templates can be time and space consuming
Template-Based Recognizers

• Templates (4 classes, 3 examples per class)

- check
- “X”
- triangle
- pigtail

• Query gesture
Gesture Set of “$1 Recognizer”

- Unistroke gestures (touch – move – release)

- Dot indicates start point

- http://depts.washington.edu/aimgroup/proj/dollar/
Variability in Raw Input

- Number and distribution of sample points depends on
  - Sampling rate
  - Movement speed and variability
  - Movement amplitude (scale)
  - Initial position and orientation
Preprocessing of Gesture Trace

• Resample to fixed number of points
  – E.g., N = 16 points
  – Linear interpolation
  – Length per step = \( \text{pathLength} / (N-1) \)

• Compute centroid \( c \)

• Translate by \(-c\)
  – Centered at origin

• Normalize \( v \) (to length 1)
  – Treat trace as vector of \( \mathbb{R}^{2N} \):
    \[ v = x_1, y_1, x_2, y_2, \ldots, x_N, y_N \]
Gesture Recognition

• Gesture recognition = search for most similar template
• Preprocessed query gesture $g$ and templates $t_j$
  – Resampled ($N=16$), centroid translated to origin, normalized
• “Most similar” metric?
  – Sum of squared differences between points
    $$\min_{j=1..M} \left\{ \sum_{i=1..2N} (g_i - t_{ji})^2 \right\}$$
  – Scalar product between query gesture and template
    $$\min_{j=1..M} \left\{ \arccos \left( \sum_{i=1..2N} (g_i t_{ji})^2 \right) \right\} \quad \text{or} \quad$$
    $$\max_{j=1..M} \left\{ \sum_{i=1..2N} (g_i t_{ji})^2 \right\}$$
• Remaining variability: rotation (and gesture class)
Optimal Angular Distance

• Orientation of template might be different from query gesture
• Example:

Overlaying query gesture (black) and optimally rotated best-matching template (red):

• How to find the optimal angle?
Finding the Optimal Angular Distance

- Wobbrock et al., UIST’07
  - “Seed and search”: Given query and template, try different orientations and take best one

- Li, “Protractor”, CHI’10
  - Closed form solution!
  - Better speed and performance!

- Closed form solution: Find $\theta$ that optimizes metric
  - Metric: Min. angle between query gesture $g$ and template $t$ in $\mathbb{R}^{2N}$
  - Optimal angle: $\theta = \arg\min_{-\pi \leq \theta \leq \pi} \{ \arccos(g \cdot t(\theta)) \}$
  - Equivalent: Max. scalar product between $g$ and $t$ in $\mathbb{R}^{2N}$
  - Optimal angle: $\theta = \arg\max_{-\pi \leq \theta \leq \pi} \{ g \cdot t(\theta) \}$
Optimal Angular Distance: Closed Form Solution

- Maximize scalar product $g \cdot t(\theta)$
- Find $\theta$ that maximizes scalar product between $g$ and $t$
  \[
  \theta = \arg\max_{-\pi \leq \theta \leq \pi} \{ g \cdot t(\theta) \}
  \]
  \[
  g = x_1, y_1, ..., x_N, y_N
  \]
  \[
  t(0) = x^t_1, y^t_1, ..., x^t_N, y^t_N
  \]
- Rotate each point in $t$ by $\theta$
  \[
  R(\theta) = \begin{bmatrix}
  \cos \theta & -\sin \theta \\
  \sin \theta & \cos \theta
  \end{bmatrix}
  \]
  \[
  x' = x \cos \theta - y \sin \theta
  \]
  \[
  y' = x \sin \theta + y \cos \theta
  \]
  \[
  t(\theta) = x^t_1 \cos \theta - y^t_1 \sin \theta, \quad x^t_1 \sin \theta + y^t_1 \cos \theta, \ldots
  \]
Optimal Angular Distance: Closed Form Solution

• Scalar product $g \cdot t(\theta)$
  
  $= \sum_{i=1}^{N}(x_i x_i^t \cos \theta - y_i y_i^t \sin \theta) + y_i (x_i^t \sin \theta + y_i^t \cos \theta)$
  
  $= \sum_{i=1}^{N}(x_i x_i^t \cos \theta - x_i y_i^t \sin \theta + y_i x_i^t \sin \theta + y_i y_i^t \cos \theta)$
  
  $= \sum_{i=1}^{N}(\cos \theta (x_i x_i^t + y_i y_i^t)) \sin \theta (y_i x_i^t - x_i y_i^t))$
  
  $= \cos \theta \sum_{i=1}^{N}(x_i x_i^t + y_i y_i^t)$ + $\sin \theta \sum_{i=1}^{N}(y_i x_i^t - x_i y_i^t)$
  
  $= a \cos \theta + b \sin \theta$
  
  with $a = \sum_{i=1}^{N}(x_i x_i^t + y_i y_i^t)$
  
  and $b = \sum_{i=1}^{N}(y_i x_i^t - x_i y_i^t)$

• Remaining task: $\theta = \arg\min(a \cos \theta + b \sin \theta) = \arg\min(f(\theta))$
  
  Find extremum of $f$ by deriving $f$ w.r.t. $\theta$ and setting $f'(\theta) = 0$:
  
  $-a \sin \theta + b \cos \theta = 0 \iff a \sin \theta = b \cos \theta$
  
  $\iff \sin \theta / \cos \theta = b / a = \tan \theta$
  
  $\iff \theta = \tan^{-1}(b / a)$
Display and Touch Screen Technologies
Liquid Crystal Display (LCD)

- An LCD cell is a voltage-controlled “light valve”
- Twisted nematic effect
  - Orientation of molecules controls orientation of polarized light
  - Off: Liquid crystal molecules form helix structure, 90° rotation
  - On: Electric field aligns molecules, second polarizer blocks light

![Off state](source: Wikipedia)

![On state](source: Wikipedia)

![LCD pattern](source: Wikipedia)
Liquid Crystal Display (LCD)

Off state

On state

Liquid Crystal Display (LCD)

• Advantages
  – Low power consumption for controlling the twisted nematic effect
    • Low operating voltages (batteries)
    • Now current flow required
  – Cheap
  – Compact: light, small, low depth
  – Flicker free, sharp, undistorted image

• Disadvantages
  – Backlight illumination consumes significant amounts of power
  – Difficult manufacturing process (dead pixels, defective panels)
  – Fixed pixel resolution
  – Limited contrast and viewing angles (early LCDs)
**Touch Screens**

- **Resistive**
  - Suitable for stylus input
- **Capacitive**
  - Direct finger input, e.g., iPhone
- **Surface Acoustic Wave (SAW)**
  - Senses diffraction of waves on surface
- **Frustrated Total Internal Reflection**
  - Jeff Han’s multitouch table
Resistive Touch Screens

1. Polyester film
2. Upper resistive circuit layer
3. Conductive metal coating
4. Lower resistive circuit layer
5. Insulating dots
6. Glass/acrylic substrate
7. Touching the overlay surface causes (2) to touch (4), producing a circuit switch from the activated area
8. Touchscreen controller measures voltages through resistive layers and converts them into the digital X and Y coordinates of the activated area.

(www.fastpoint.com)
Capacitive Touch Screens

• Senses capacitive changes
  – **Only works with finger, not with stylus**

• iPhone
  – **Uses additional grid for better multitouch disambiguation**

(www.unwiredview.com)
Self-Capacitance Touch Screen

- Detects only a single touch point
- Measures capacitance of electrode to ground
- Finger near electrode: human body capacitance changes self-capacitance of electrode

- Materials: copper, indium tin oxide (ITO), printed ink
  - ITO: (almost) transparent capacitive electrodes
- Rows isolated from columns in grid arrangement
- Size and spacing between electrodes determines precision

Gary Barrett and Ryomei Omote: Projected-Capacitive Touch Technology. Information Display 3/10, pp. 16-21
Self-Capacitance Touch Screen

• Scans each electrode individually
• Sensing only

Diagram showing a grid with electrodes labeled X0 to X3 and Y0 to Y3, with touch points at Y1 and Y2.
Mutual Capacitance Screen

- Unlimited number of touch points
- Measures capacitance of intersections of electrodes
- Human-body capacitance changes capacitance of intersections ("steals" charge)

- High resolution
- Less sensitive to EMI than self-capacitance

- Typically 9 columns, 16 rows = 144 electrode intersections
- Interpolation achieves 1024x1024 (10 bit) resolution
Mutual Capacitance Screen

- Senses each pair \((X_i, Y_j)\) of electrodes individually
- Driving and sensing
It’s Easy:

- [http://mediathek.daserste.de/daserste/servlet/content/6099692](http://mediathek.daserste.de/daserste/servlet/content/6099692)
Pico Projectors

- Standalone or integrated in mobile phones
- Interesting for collaborative applications
  - Example: sharing media
- Problems (current technology)
  - Availability of projection space
  - Ambient light
  - Power consumption
  - Focusing
Audio and Haptics

Partly based on slides by Stephen Brewster:


Multimodality

• Involve different senses through different modalities
  – Audio, tactile
  – Suit different users, tasks, and contexts

• Problems of visual modality
  – Screen space small
  – Eyes heavily used when mobile

• Reasons for multimodality
  – Sole use of one modality not effective
  – Particular modality may not always be available all of the time
  – User involved in other tasks → Attention may be occupied
Multimodal Interaction

• Allow people to do everyday tasks while using mobile technology
  – “Eyes-free” or “hands-free”

• Interaction techniques that suit real environments
  – Non-speech audio and tactile feedback
  – Sensors for gestural input
  – Speech input
Non-Speech Audio

• Earcons (Blattner)
  – Musically structured sounds (abstract)

• Auditory Icons (Gaver)
  – Natural, everyday sounds (representational)

• Sonification
  – Visualization using sound
  – Mapping data parameters to audio parameters (abstract)
Earcons

• Structured audio messages based on abstract sounds
  – Created by manipulation of sound properties: timbre, rhythm, pitch, tempo, spatial location (stereo, 3D sound), etc.

• Composed of motives

• Can be compound
  – Sub-units combined to make messages

• Or hierarchical
  – Sounds manipulated to make complex structures

Examples from: [http://www.dcs.gla.ac.uk/~stephen/earconexperiment1/earcon_expts_1.shtml](http://www.dcs.gla.ac.uk/~stephen/earconexperiment1/earcon_expts_1.shtml)
Auditory Icons

• Sounds mapped to interface events by analogy to everyday sound-producing events
  – E.g., selecting — tapping; copying — filling
  – Iconic v. symbolic mapping

• Auditory icons can be parameterized
  – E.g. material for type, loudness for size
  – Multiple layers of information in single sounds
  – Reduces repetition and annoyance

• The SonicFinder
  – Selecting, copying, dragging

Sonification

- Mapping of data values to auditory parameters
- Most commonly x-axis to time, y-axis to pitch

Sonification of Luminance Histograms in Digital Cameras

• Difficult to focus visual attention on subject and technical parameters
  – Exposure, aperture, battery life, image mode, etc.
• Idea: Sonified luminance histogram

• Sonification of remaining memory space

3D Audio Interaction

• Increase the audio display space
• 3D audio
  – “Cocktail party effect”
  – Provides larger display area
  – Monitor more sound sources
• “Audio Windows” (Cohen)
  – Each application gets its own part of the audio space
• Pie Menus (Brewster, CHI’03, Marentakis, CHI’06)
  – Audio items placed around the head


Brewster, Lumsden, Bell, Hall, Tasker. Multimodal 'eyes-free' interaction techniques for wearable devices. CHI '03.

Marentakis, Brewster. Effects of feedback, mobility and index of difficulty on deictic spatial audio target acquisition in the horizontal plane. CHI '06.

Haptics

- Definition: Sense and/or motor activity based in the skin, muscles, joints, and tendons
- Two parts
  - Kinaesthesis: Sense and motor activity based in the muscles, joints, and tendons
  - Touch: Sense based on receptors in the skin
    - Tactile: mechanical simulation of the skin
Why Haptic Interaction?

• Has benefits over visual display
  – Eyes-free

• Has benefits over audio display
  – Personal not public
  – Only the receiver knows there has been a message

• People have a tactile display with them all the time
  – Mobile phone

Tactile Technologies

• Vibration motor with asymmetric weight
  – Eccentricity induces vibrations
  – Speed controls vibration frequency
  – Precision limited (several ms startup time)

Source: Haptics, audio output and sensor input in mobile HCI by Stephen Brewster
Tutorial Mobile HCI 2008.
Design of Tactons

• Tactons = tactile icons
  – Structured, abstract messages that can be used to communicate non-visually (Brown, 2005)
  – Tactile equivalent to Earcons

• Encode information using parameters of cutaneous perception
  – Body location
  – Rhythm
  – Duration
  – Waveform
  – Intensity

Tacton Parameters

• Spatial location (on forearm, waist, hand) very effective
  – Good performance with up to 4 locations
  – Wrist and ankle less effective, especially mobile

Tacton Parameters

• Rhythm very effective
  – Easily identified with three levels

• Waveform
  – Carefully designed sine, square, and sawtooth wave forms very effective (tuned to capabilities of actuator)

• Intensity
  – Two levels
  – Hard to use and may need to be controlled by user

Example: Tactile Button Feedback

• Touchscreen phones have no tactile feedback for buttons
  – More errors typing text and numbers

• Performance comparison of physical buttons, touchscreen, and touchscreen+tactile
  – In lab and on subway

• Touchscreen+tactile as good as physical buttons
  – Touchscreen alone was poor

Brewster, Chohan, Brown: Tactile feedback for mobile interactions. CHI ’07.
Example: Tactile Navigation

- Non-visual interface for GPS + compass
- Belt of 4 actuators
  - Placed north, south, east, west
- Vibrations gave direction and distance

- Users could follow paths accurately without a screen

The End