## Visualization for the Blind

Richard Anderson,
Ken Fishkin,
Songyun Hahn



## My homework

- "Correlation does not imply causality"
  - Or, "*post hoc ergo propter hoc*"
    - "after this therefore because of this"
- Sorry about running out of time on space shuttle graphs



## Administrivia

- Keep working towards your project proposals
  - I've already heard 4 – way to go!



## Administrivia(2)

- Have you had a conference or journal submission reviewed?
- Have you been a reviewer?

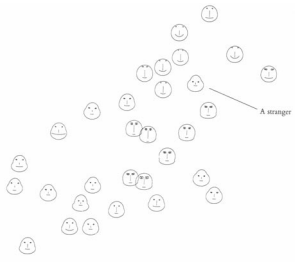


## We've seen many visualization techniques



## But what if you're blind?

## What about this?






## What about this?



## You get the idea

- 1.3 million legally blind in US
- 68% unemployment rate
  - WWW has been a *negative*
- Criminally overlooked area



## Another motivation

- "But even well-meaning Web site builders ask: How can I justify the extra cost for such a small percentage of the public?" (Steven Pemberton, ACM Interactions, Feb 2003, p. 44)



## Another motivation

- "The answer is: Google. … It is clear that at least half the visitors find the sites via Google. And what Google sees is exactly what a blind person sees. Google is a blind user – a billionaire blind user, with millions of friends who listen to its every word. If a blind user can't see your site, neither can Google, and you site will suffer." (Pemberton)



## Mynatt paper

- Really two papers in one – presents both "Mercator" and "GUIB".
- Two different systems for providing access to graphical applications
- Fairly old (predates Web), foundational

## Mercator vs. GUIB

- Both assume app is "black box"
- GUIB relies on *tactile* output – takes screen *pixels* and maps them to a 2D dynamic braille display
- Mercator relies on *audio* output – takes screen *objects* and maps them to audio.
- (some of this is due to Europe vs. US)
- 4 design goals



## Goal: Coherence with Visual

- "An overriding concern … is maintaining coherent, parallel … interfaces". Why?
- Collaboration ("primary reason")
  - "to support discussion"
  - "to support simultaneous interaction"
- Training



## Goal: Information packing

- Need to support random-access of an information-rich environment (e.g. mouse-click anywhere on a map)



## Goal: convey visual information

- Not enough to just translate words
- 1 step up: buttons, menus
- Next step up: sliders, scrollbars
- Uh-oh: random graphics



## Goal: Interaction

- Can only rely on keyboard, and audio output (not even audio input).
- How do you do WIMP without P?



## Goal: wide scope

- Treat apps as "black boxes", so can handle multiple apps
- Both monitor GUI painting calls and build a model of the screen
  - GUIB in Windows/X
  - Mercator in X
- Imposes a severe limit on how smart they can be

## GUIB

- "GUIB continues the use of the spatial metaphor" – a line is a line.
- Map is maintained from NV display to V display, on a per-pixel basis.



## GUIB: text

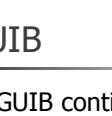
- Caret is a *conceptual* object, but is mapped *physically*

| interaction object | example | braille-based presentation | example |
|---|---|---|---|
| caret, mouse pointer | text | one braille character | t■xt ■ |
| text, text attributes | sample | braille, attributes through dots 7 and 8 or on request | sample |



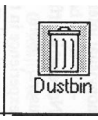
## GUIB: window

- User feels the menu bar

| | | | |
|---|---|---|---|
| window |  | window frame in braille<br>name is spoken<br>pop-up, move, size by sound | +[-] Notepad-(Unt [↓][↑]<br>\|F\|File E\|Edit S\|Search<br>+---------- |



## GUIB: icon

| | | | |
|---|---|---|---|
| icon |  | name in braille | [Icon] Dustbin |





## GUIB: menu

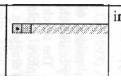
- Note how selection is shown

| | | | |
|---|---|---|---|
| menu |  | all items in braille<br>vertical or horizontal layout | F\|File O\|Options W\|Wi<br>\|*A■Auto Arrang<br>M\|Minimize on<br>*S\|Save Settin<br>+---------- |

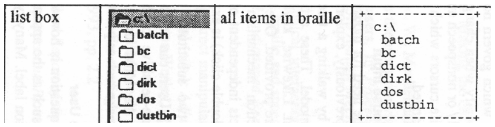


## GUIB: scroll bar

- works well here

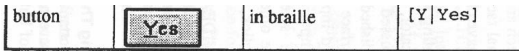
| | | | |
|---|---|---|---|
| scroll bar |  | in braille | ■-□----------■ |

## GUIB: list box






## GUIB: button






## Summary:

- + maps a 2D display to a 2D display, can do it "all at once"
- - adds junk user doesn't care about
- - requires 2D braille display



## Mercator

- Instead of 2D display, uses sound
- Sound is 1D, what do you do?
- Have a focus at any given point, describe what's around the focus
- Use other dimensions of sound (pitch, "earcons" to annotate)



## Mercator: text

| | | | | |
|---|---|---|---|---|
| caret, mouse pointer | text | text around caret is spoken | "e" | audio "click" at caret |
| text, text attributes | sample | text is spoken, attributes are verbalized (??) | "sample" | pitch of speech is modified for attributes |

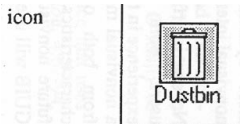
- Uses pitch, "earcon" to show attributes/context



## Mercator: window

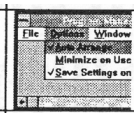

- Uses earcons – you get the idea

## Mercator: icon

| | | |
|---|---|---|
| icon   | "dustbin" | auditory icons |

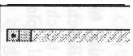


## Mercator: menu

| | | |
|---|---|---|
| menu   | "auto arrange" | auditory icon for menu-button, pitch is modified relative to location in menu |



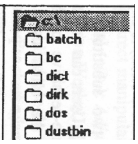
## Mercator: scroll

| | | |
|---|---|---|
| scroll bar  | "slider at zero percent" | auditory icon for scollbar, location conveyed with pitch |

- Requires fine pitch discrimination



## Mercator: list box

| | | |
|---|---|---|
| list box   | "c colon backslash" | auditory icon for list, pitch is modified relative to location in list |



## Mercator: button

| | | |
|---|---|---|
| button   | "yes" | auditory icon for push button |

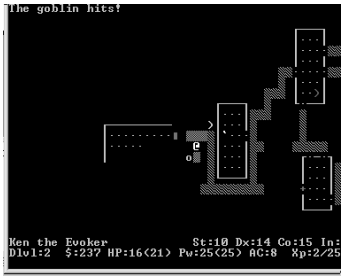


## Questioning the assumptions

- Should Visual/Non-Visual UIs necessarily have a 1-1 map in concepts and metaphors?
- For example, GUIB approach is rather like "curses":

## Counter-Example

- Nethack, 2D:






## Counter-Example:



## Nethack vs. Diablo

- They didn't keep the same metaphors, why should we?



## Yu paper

- Previous works focused on (largely) textual data, in a 2D arrangement. What about visualizations?
- Multivis (http://www.multivis.org) project looking at bar charts, graphs, pie charts.
- You were assigned 1 paper of a "suite"

GIST
GLASGOW INTERACTIVE SYSTEMS GROUP



## Basic concepts

- Like earlier papers, use haptic or audio
- Go beyond them by *combining* them
- Ditch "Coherency" constraint
- Semi-real user studies to test (more on this later)



## Haptic

- Use (a) "phantom" (3D force-feedback bat)
  - (http://www.sensable.com/haptics/products/phantom.htm l)

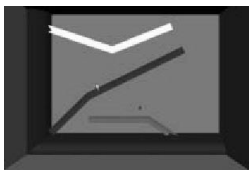
## Haptic

- Or more commonly Logitech force-feedback mouse ($60 vs $10K)



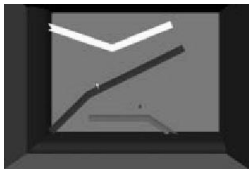
## Haptic

- Make a "sculpture" out of visualization, with "grooves" felt by force-feedback.





## Line graph





## Bar Chart

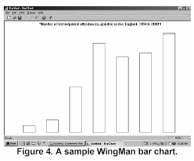


Figure 4. A sample WingMan bar chart.

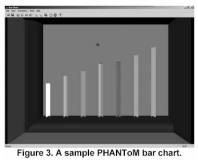
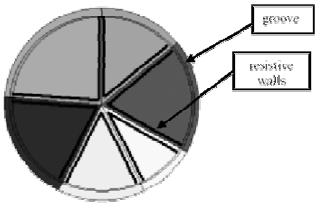


Figure 3. A sample PHANToM bar chart.



## Pie Graph






## Note

- 3rd dimension not really used with phantom
- Also mentioned that felt limited, you mainly "poke" to sense the outline of the shape.

## Audio

- Line: MIDI, high Y → high pitch
  - -- assumes lines with unique Y for each X
- Bar: MIDI, high Y → high pitch
  - Later work: high X → right ear
- Pie: MIDI, high % of pie → high pitch
- In all cases, sound starts/stop when region entered/left



## Non-speech vs. speech

- In later paper, also tried speech sounds (voice "reads" value on graph)
- Found non-speech significantly better



## They tested it!

- Most tests were of undergrads wearing blindfolds



## Digression

- Example of why user testing is hard, and rarely done, or done well
- Particular issue in this environment
  - Hard to find users
  - Human subjects approval takes a *long* time



## Digression(2)

- To be rigorous in the testing:
- 3 types of graphs
- * 4 conditions = 12 scenarios
- * n >= 5 (at least)
- → 60 people



## More on the testing

- Blindfolding questionable: very different set of haptic/audio skills, different expectations, etc.
- However, they did do some testing with blind as well.
- Surprisingly (to me), found only quantitative, not qualitative, differences in the two user groups

## The Results

- Haptic is better than nothing
  - Little things make a big difference: separating bars hurts
- Audio is better than haptic
- Haptic + Audio is better than either



## This just in

- Recent Roth/Kamel/Petrucci/Pun paper ("A Comparison of Three Nonvisual Methods for Presenting Scientific Graphs", J. Visual Impairment and Blindness, June 2002, **96**(6)) found similar results:
  - Haptic + Audio better then either
  - Presenting continuous graphs
  - Did real user testing: sighted and blind had similar qualitative results



## Automatic tactilization of graphical images

**- Brief introduction of the project based on NSF proposal -**

**January 22, 2003**
**Sangyun Hahn**



## Background

- Graphical representation for the blind
  - Audio description
  - Braille graphics
  - Haptics
  - Braille is the best modality for image comprehenssion (Skiff, 2002)



## Production process of braille graphics

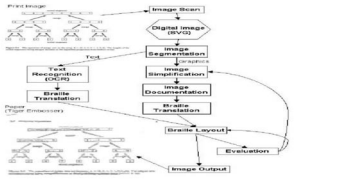



Figure 1. Process for converting a print graphical image into Braille. The Braille image was not optimized to improve legibility or comprehension.



## Tiger embosser

- The first braille printer that supports Windows-style WYSIWYG printing
- Can print braille characters and graphics together
- Print with highest resolution (20 dpi) among currently available embossers
- Can emboss dots with different heights.

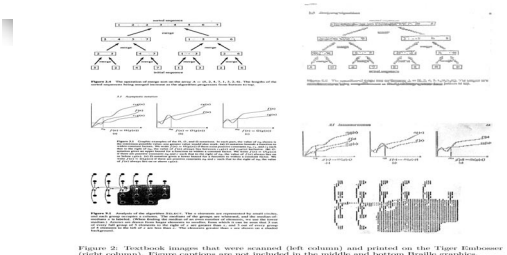


Figure 2: Textbook images that were scanned (left column) and printed on the Tiger Embosser (right column). Figure captions are not included in the middle and bottom Braille graphics.



## Project description

- Study various aspects of braille forms of graphical images influencing comprehensibility of blind users
- Develop image processing and layout algorithms to produce desired braille forms.
- Develop a model for assessing the quality of generated layouts
- Develop a tool to automatically convert or interactively edit graphical images to braille graphics.



## Tactual perception

- Study factors that affect comprehensibility and efficiency of tactual perception.
  - resolution
  - variable dot heights
  - use of different texture
  - ways of representing colors
  - ways of placing labels
  - use of legends



## Automatic translation

- Image processing
  - edge detection, resolution reduction, de-noising, segmentation, filtering
- Automatic classification of image type
- Label placement
  - decide association between graphical elements and labels, and use of legends.



## Quality prediction model

- Methods for predicting processing time
  - GOMS (Goals, Operators, Methods, and selection) analytical model
  - Simulation methods based on computational cognitive architecture
  - Statistical prediction model (Ivory 2002)
  - Reverse-engineering methods



## Tactile Image Editor

- Provide both image and text editing
- Support both automatic translation
- and interactive editing
- Quality prediction
- Online image translation

## Thanks!

- Thank you Sangyun



## Project

- Explore ways to present a scatterplot, Tukey bars, a hyperbolic tree, or any of the other visualization techniques we've discussed in this class that are not addressed by the Glasgow group



## For next time

- Go to http://bailando.sims.berkeley.edu/infovis.htm
  - **Either** read Kai-Ping Yee, Danyel Fisher, Rachna Damija, and Marti Hearst. "Animated Exploration of Graphs with Radial Layout" in *IEEE Infovis Symposium*, San Deigo, CA, October 2001
  - **or** View the video
- Go to http://graphics.stanford.edu/papers/polaris/ and read Chris Stolte, Diane Tang and Pat Hanrahan. "Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases" *IEEE Transactions on Visualization and Computer Graphics, Vol. 8, No. 1*, January 2002, in your favorite format.