



Interaction Design Research
Spring 2017

INPUT MODALITIES

Mohammad Motiei



Definition of an “Input Modality”:

From Wikipedia:

- In the context of **Human-Computer Interaction**, a **modality** is the classification of a single independent channel of sensory **input/output** between a computer and a human.



Goals

- Understanding the importance of input modalities in design
- Finding out about the virtues as well as challenges of combining input modalities through the example of PixelTone which combines speech with direct manipulation for photo editing.
- Learning about 7 principles of interaction design that D. A Norman introduced and investigating them on different input modalities.

PixelTone: A Multimodal Interface for Image Editing

■ Motivations:

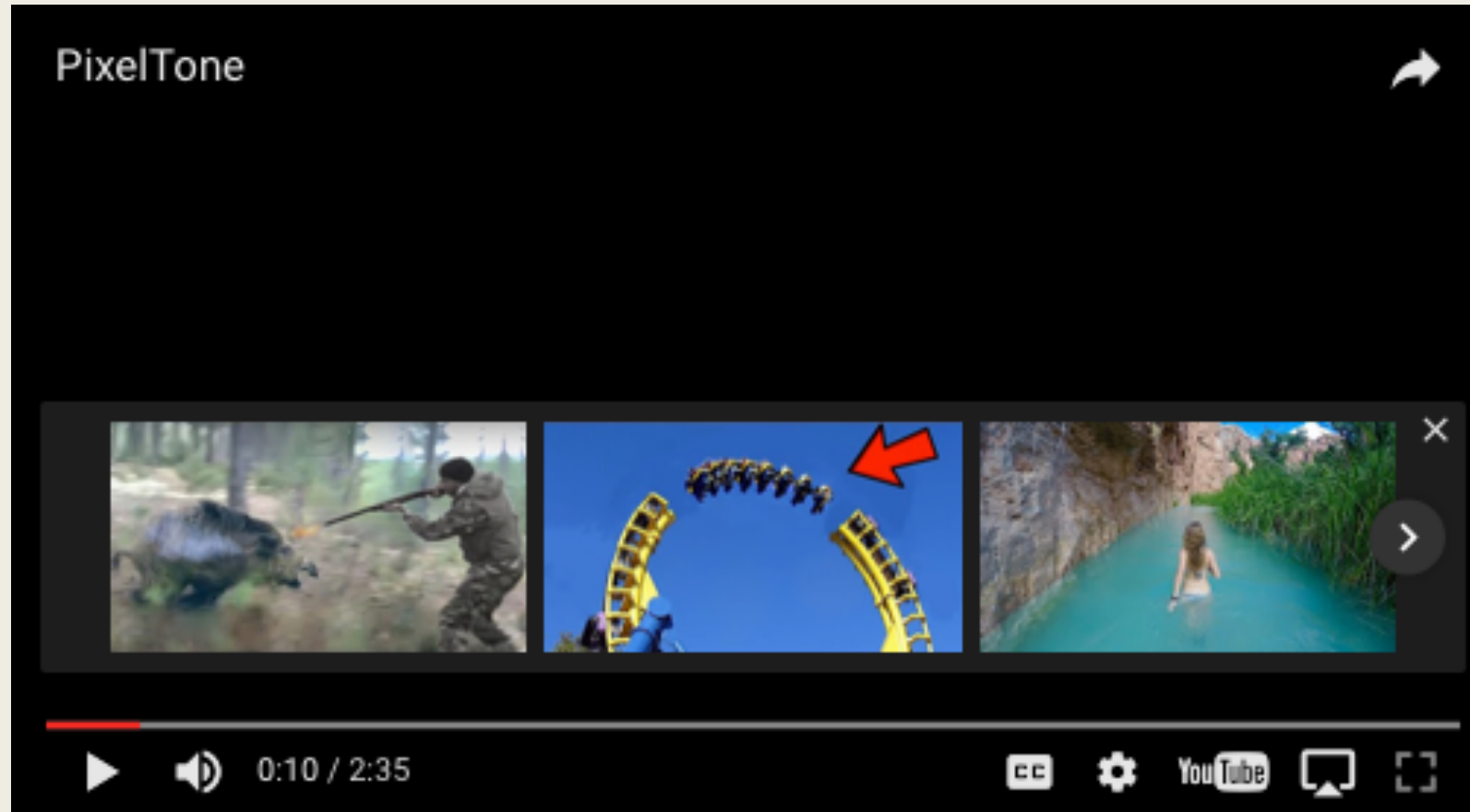
1. The interfaces for photo editing are complex
2. Photo editing applications expect user to learn the language of photo editing
3. Photo editing can be a more difficult task on small, portable devices.

PixelTone:

- In PixelTone, users with a combination of expressing desired changes as well as sketching could edit an image.



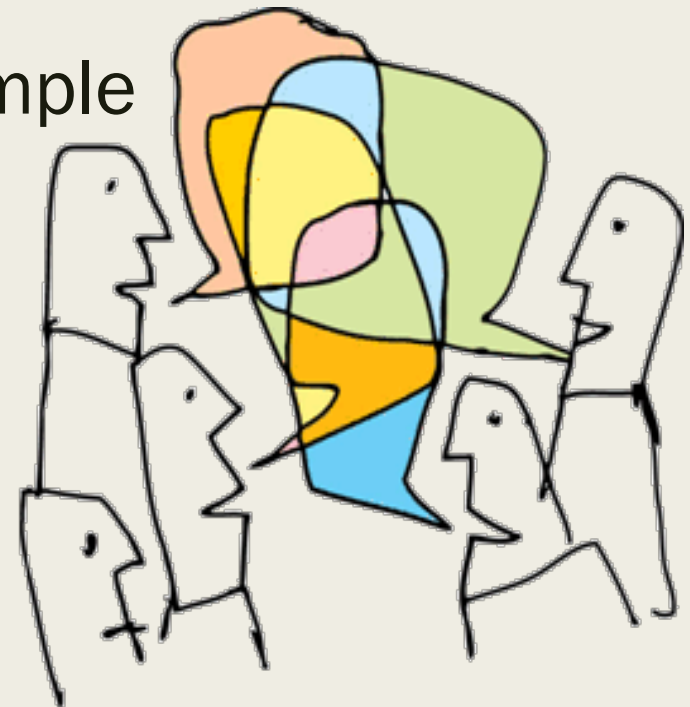
Lets Watch The Demo



Discussion Time!

- Do you think that PixelTone is an example of **ability-based** design?

If the answer is yes, justify it; but if the answer is no, in what way the final product would be different to PixelTone if a designer tries to use ability-based design for it.



Related Work and Other Applications

- Speak'n'Sketch : lets artists issue commands, such as “group”, “rotate”, or “thicker brush” as they are sketching.
- Pausch and Leatherby showed that adding voice to a drawing application reduces time to completion by up to 56%.
- The new feature on Instagram, in which you can record a video while the recording voice would be the music that you are listening to through your handsfree.

Discussion Time!

- What are the contexts in which adding an additional modality is beneficial, when would it be not desirable?



Understanding Natural Language In Editing Process.

- Professional Annotation



Understanding Natural Language In Editing Process.

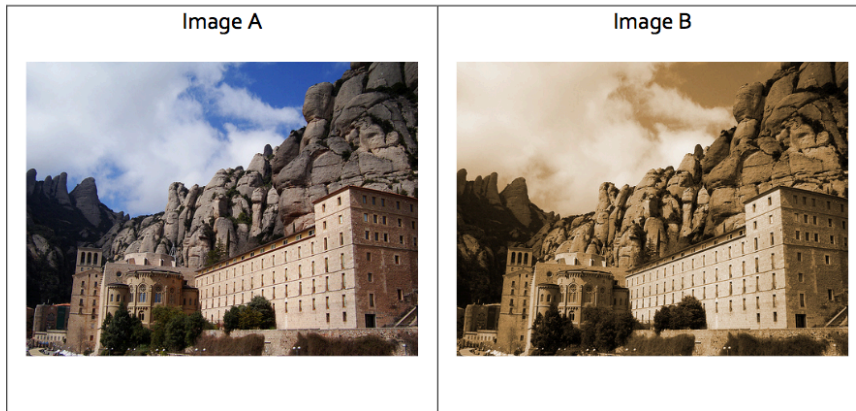
■ Crowdsourcing

- Used Amazon Mechanical Turk
- 10 individual images
- 14 image pairs
- Collected 10 responses for each task (240 response)
- Used the same image set as the lab study

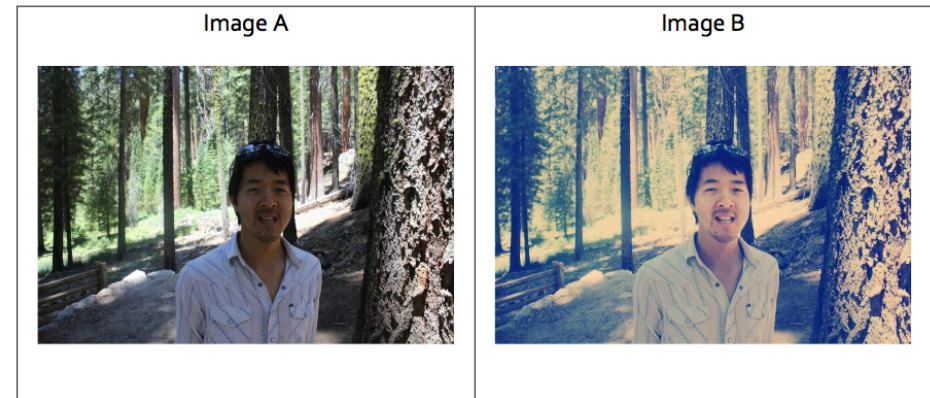
Understanding Natural Language In Editing Process.

- Crowdsourcing (transforming)

Easy #03



Medium #01



Challenging #01



Understanding Natural Language In Editing Process.

- Crowdsourcing (improvement task)

Easy #05

Image A



Medium #03

Image A



Challenging #03

Image A



Understanding Natural Language In Editing Process.

■ Crowdsourcing

A few commands that gathered from the experiment:

- Either add blue tint or reduce the green in the center of the image
- Reduce shadows
- Add shadows
- Use better function
- Remove the grainy texture
- Reduce glare from windows
- Reduce the graininess of the image
- Make the faces sharper
- Zoom in
- Zoom out
- Add a yellowish haze to the image

Design Guidelines

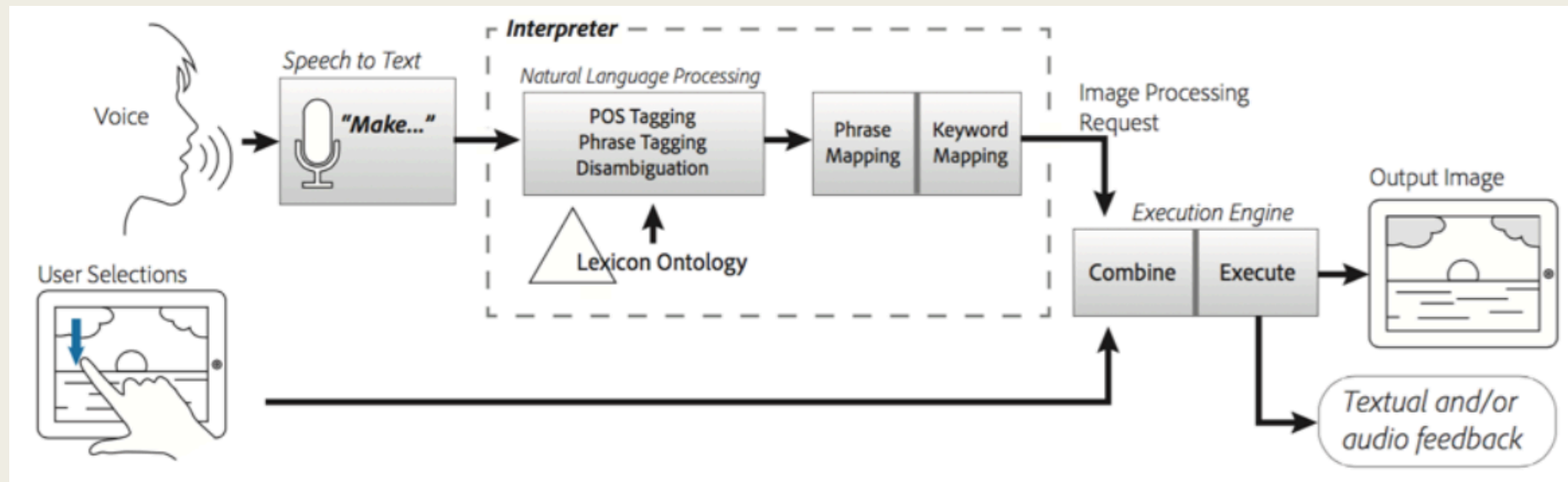
for incorporating natural language into image editing interfaces

- Incorporate object references.
- Support declarative and imperative sentences.
- Guide users with available commands.
- Allow localized image edits.



PixelTone Framework

- Speech Recognition
- Speech Interpretation
- Execution Engine



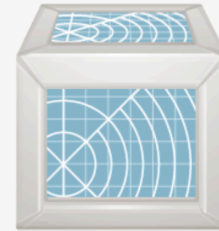
PixelTone Framework

■ Speech Recognition

- Local Speech Recognition

OpenEars® - iPhone Voice Recognition and Text-To-Speech

OpenEars: free speech recognition and speech
synthesis for the iPhone



- Remote Speech Recognition



PixelTone Framework

■ Speech Interpretation

“Make the shadows on the left slightly brighter”

Two level tag hierarchy for parsing phrases:

1- Phrase Level → Verb, Noun, and Adjective

Make: VX the shadows on the left: NX slightly brighter: AX

2-Word Level → Penn Treebank tags

slightly → Adverb (VB) | brighter → Adjective (JJR)

Penn Treebank Tags

- | | | | |
|----------|--|-----------|---------------------------------------|
| 1. CC | Coordinating conjunction | 19. PRP\$ | Possessive pronoun |
| 2. CD | Cardinal number | 20. RB | Adverb |
| 3. DT | Determiner | 21. RBR | Adverb, comparative |
| 4. EX | Existential <i>there</i> | 22. RBS | Adverb, superlative |
| 5. FW | Foreign word | 23. RP | Particle |
| 6. IN | Preposition or subordinating conjunction | 24. SYM | Symbol |
| 7. JJ | Adjective | 25. TO | <i>to</i> |
| 8. JJR | Adjective, comparative | 26. UH | Interjection |
| 9. JJS | Adjective, superlative | 27. VB | Verb, base form |
| 10. LS | List item marker | 28. VBD | Verb, past tense |
| 11. MD | Modal | 29. VBG | Verb, gerund or present participle |
| 12. NN | Noun, singular or mass | 30. VBN | Verb, past participle |
| 13. NNS | Noun, plural | 31. VBP | Verb, non-3rd person singular present |
| 14. NNP | Proper noun, singular | 32. VBZ | Verb, 3rd person singular present |
| 15. NNPS | Proper noun, plural | 33. WDT | Wh-determiner |
| 16. PDT | Predeterminer | 34. WP | Wh-pronoun |
| 17. POS | Possessive ending | 35. WP\$ | Possessive wh-pronoun |
| 18. PRP | Personal pronoun | 36. WRB | Wh-adverb |

PixelTone Framework

■ Supported Operations



Original



Exposure



Auto-Color



Brighten



Darken



Black & White



Posterize



Soft Focus



Contrast



Hue



Vibrance



Saturation



Blur



Vignette



Sepia



Warmth



Coolness



Green Tint



Magenta Tint



Sharpen



Pixellate



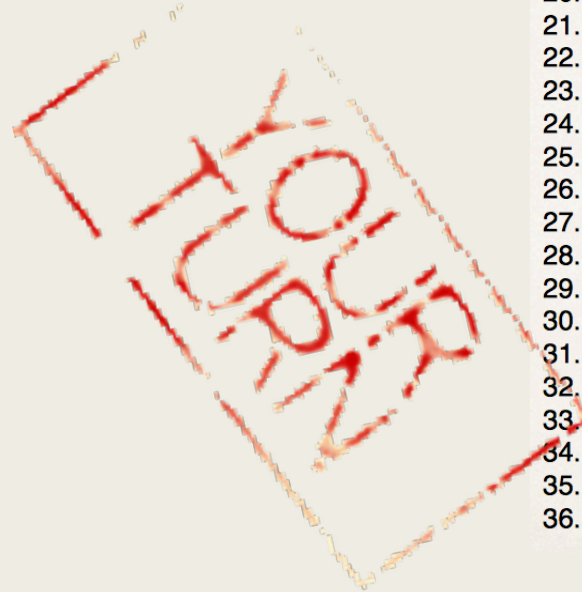
Vintage



Lomo

A Few Examples:

- Adjust white balance for entire image
- *Make the shadows warmer on the right*
- *Add slight amber light to circled areas*
- Make the image not so blurry



1. CC Coordinating conjunction
2. CD Cardinal number
3. DT Determiner
4. EX Existential *there*
5. FW Foreign word
6. IN Preposition or subordinating conjunction
7. JJ Adjective
8. JJR Adjective, comparative
9. JJS Adjective, superlative
10. LS List item marker
11. MD Modal
12. NN Noun, singular or mass
13. NNS Noun, plural
14. NNP Proper noun, singular
15. NNPS Proper noun, plural
16. PDT Predeterminer
17. POS Possessive ending
18. PRP Personal pronoun
19. PRP\$ Possessive pronoun
20. RB Adverb
21. RBR Adverb, comparative
22. RBS Adverb, superlative
23. RP Particle
24. SYM Symbol
25. TO *to*
26. UH Interjection
27. VB Verb, base form
28. VBD Verb, past tense
29. VBG Verb, gerund or present participle
30. VBN Verb, past participle
31. VBP Verb, non-3rd person singular present
32. VBZ Verb, 3rd person singular present
33. WDT Wh-determiner
34. WP Wh-pronoun
35. WP\$ Possessive wh-pronoun
36. WRB Wh-adverb

PixelTone Framework

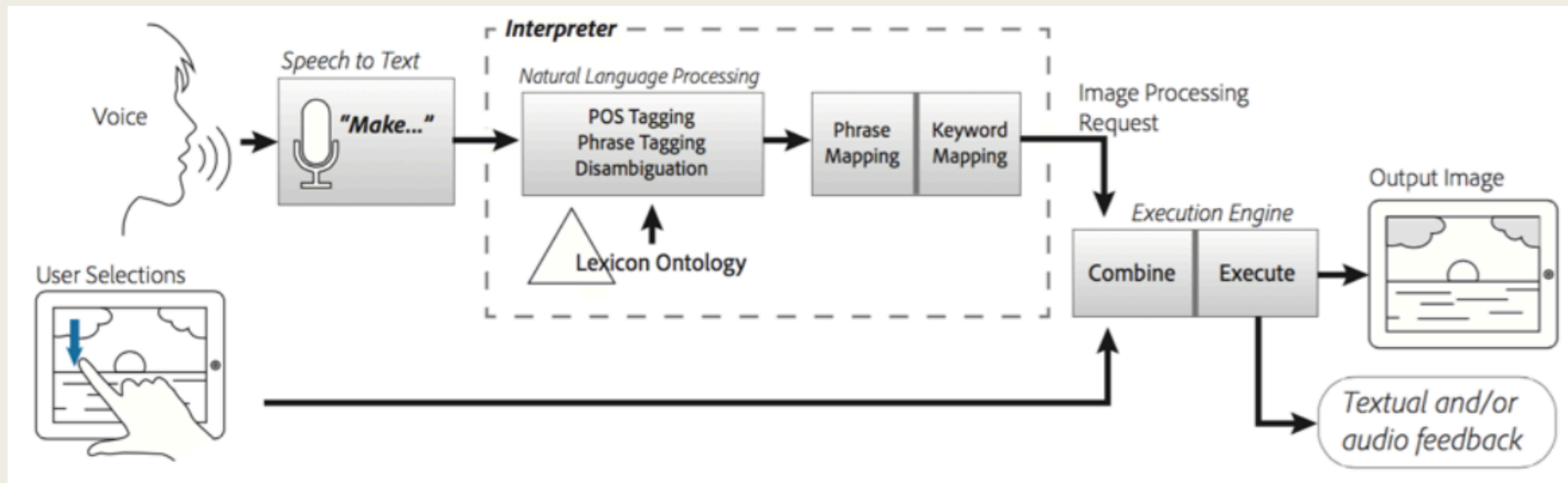
■ Execution Engine

Process command

Combine with direct manipulation

Localize

Blend multiple masks



Discussion Time!

- What would be the feature or change, that by adding it to PixelTone, you could improve the application and turn it to a better and more useful version?



Experiments And Results

Easy #01

Image A



Image B



Duration: 2 mins

"Please transform this image (pointing to Image A) to this image (pointing to Image B)"

Consequent Priming:

Effects applied to image: 1) Black & White and 2) Contrast

Challenging #02

Image A



Image B



Duration: 4 mins

"Please transform this image (pointing to Image A) to this image (pointing to Image B)"

Consequent Priming:

Effects applied to image: 1) Vibrance, 2) Hue, 3) Contrast, 4) Local blurring on edges

Experiments And Results

Easy #06

Image A



Duration: 2 mins

"Please improve this image"

Consequent Priming:
This child needs to be: darkened

Medium #04

Image A



Duration: 3 mins

"Please improve this image"

Consequent Priming:
This image needs improvement on: 1) saturation, 2) vibrance, 3) contrast

Challenging #04

Image A



Duration: 4 mins

"Please improve this image"

Consequent Priming:
This image needs improvement on: 1) vibrance, 2) saturation, 3) contrast, 4) coolness

Experiments And Results.

- **Quantitative Results :**

Success rate, users preference, number and complexity of utterances different between users

- **Qualitative Results :**

1- Users use the speech interface when they have a good idea of what they want to do.

2- Users use the gallery mode when they want to explore options and compare different effects.

3- Users use direct manipulation to fine-tune and explore.

4- Non-native English speakers with accents used speech interaction much less.

From Your Commentaries:

- “Feedback is very important when analyzing how well something is designed. The interaction layer allows users to view what markings they have made and what potential effect those markings could have” Calvin
- “This paper is really useful not in the system that it built, but in the contextualization of the different interactions it affords.” Ariel
- “A concern with the evaluation is that a feeling of "newness" of a prototype can affect people's preferences initially.” Kandarp

Gestural Interfaces: A Step Backward in Usability

■ Why are we having trouble?

1. The lack of established guide-lines for gestural control
2. The misguided insistence by companies to ignore established conventions
3. The developer community's apparent ignorance of long history of HCI

Gestural Interfaces: A Step Backward in Usability

■ Principles of Interaction Design:

1. Visibility
2. Feedback
3. Consistency
4. Non-destructive operations
5. Discoverability
6. Scalability
7. Reliability

Principles of Interaction Design

■ Visibility

Nonexistent Signifiers: Deleting a message/email/event in iPhone.

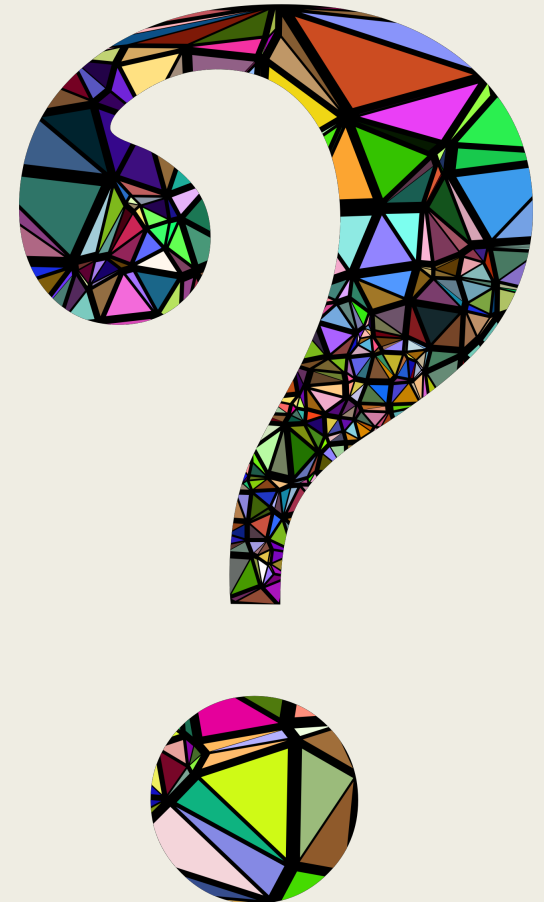
Misleading Signifiers: Four bottoms of control in Android. (back, menu, home, search)

Question

- Could you bring other examples of problems in current gestural interfaces and specify to which principle it belongs?

- Principles of Interaction Design:

1. Visibility
2. Feedback
3. Consistency
4. Non-destructive operations
5. Discoverability
6. Scalability
7. Reliability



Principles of Interaction Design

- **Feedback**

Multiple ways of returning to a previous screen.

Principles of Interaction Design

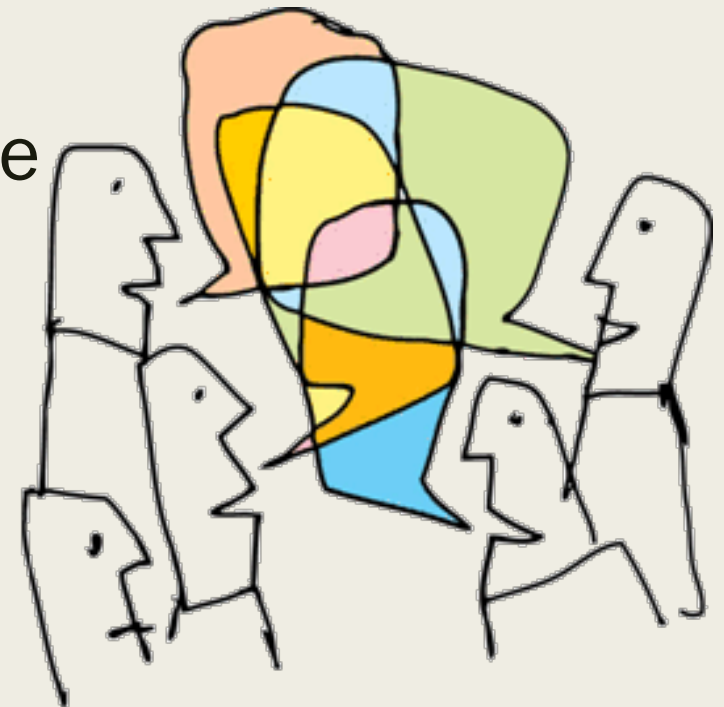
- Consistency and Standards

Some applications allows pinching to change image scale, others use plus and minus boxes.

Companies wish to protect their intellectual properties

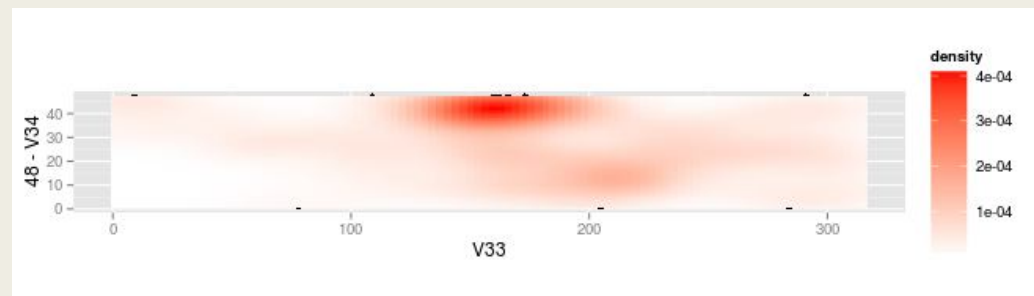
Discussion Time!

- Why or why not we should define a standard for gestures with the same action across different brands, applications, and platforms?



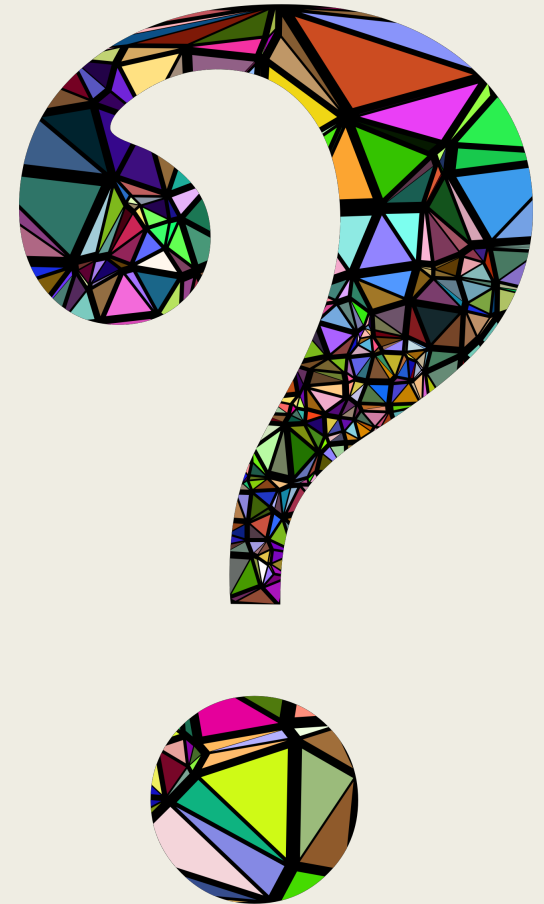
Principles of Interaction Design

■ Reliability



Question

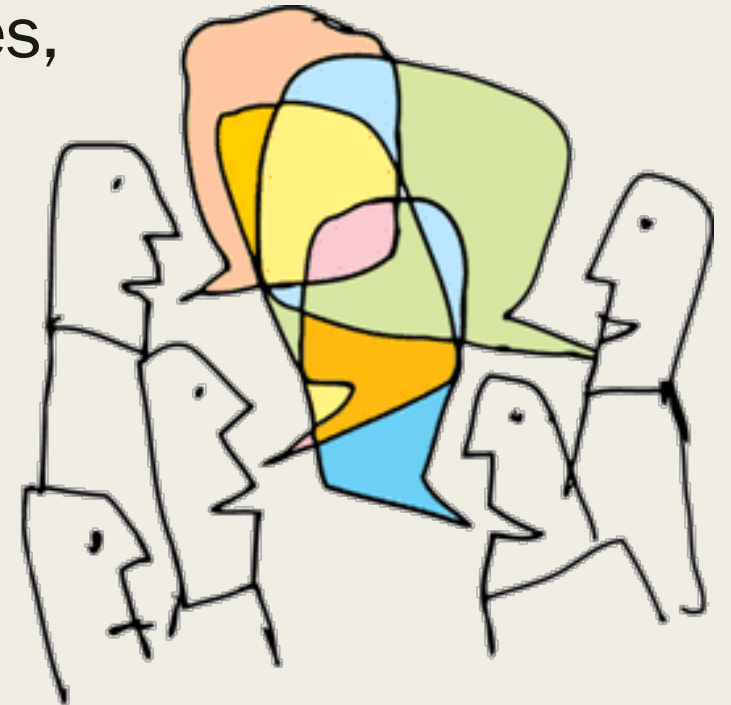
- Why we were not successful in defining a unit standard for gestures with the same action across different brands, applications, and platforms?



Again Discussion!

By considering PixelTone discuss following questions:

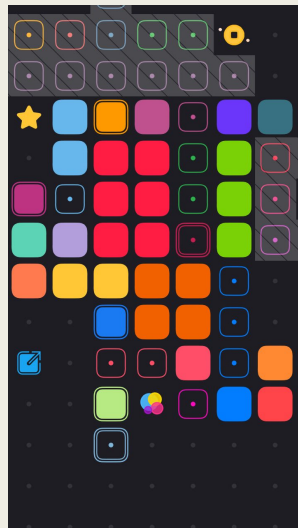
- What are the problems or shortages from the point of view of each of the 7 principles, considering voice input?
- What are the problems or shortages in each of the 7 principles, considering gestural input?
- Does combination of voice input and gestural input solve any of those problems? If yes, how?



Novel Interaction Methods

- We can tilt, shake, rotate and touch, poke and probe.

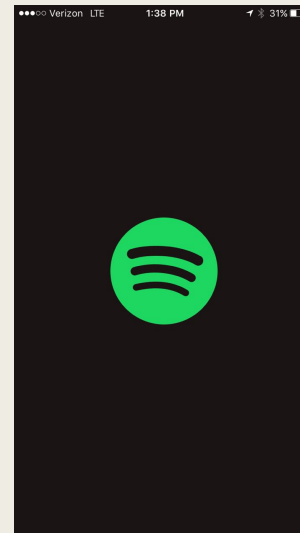
Black Box



SnapChat



Spotify



Thank
You!