Discovering Natural Language Commands in Multimodal Interfaces

Arjun Srinivasan
Mira Dontcheva
Eytan Adar
Seth Walker

Georgia Tech
Adobe Research
University of Michigan
Adobe Research
Speech-enabled multimodal interfaces are becoming popular...
apply a sepia filter on the canvas

Added a medium sepia filter
apply or light morph filter on both people

Added a light morph filter.
What operations can I perform?

How do I ask the system to perform them?
Discoverability

(1) What operations can I perform?
(2) How do I ask the system to perform them?
Discoverability

(1) What operations can I perform?
(2) How do I ask the system to perform them?

2nd most common challenge with Voice User Interfaces

Patterns for How Users Overcome Obstacles in Voice User Interfaces, Myers et al.
CHI 2018
“Hey Siri what can you do”

You can ask me to get directions, make phone calls, and send new texts and emails. I can show you lots more when you can look at your iPhone.

- **Phone**
  - "Call Brian"

- **FaceTime**
  - "FaceTime Lisa"

- **Apps**
  - "Launch Photos"

- **Messages**
  - "Tell Susan I’ll be right there"

- **Calendar**
  - "Set up a meeting at 9"

- **Sports**
  - "Hockey scores from last night"

- **Photos**
Can we leverage multimodal input to enhance discoverability by suggesting contextually-relevant natural language commands?
**Tooltips**

- **Bold (Ctrl+B)**
  - Make your text bold.

- **Font Color**
  - Change the color of your text.
  - *Tell me more*

- **Brush tool (B)**
  - Paints custom brush strokes
  - *Learn how*
Apply a blur effect on the canvas.
apply a blur effect on the canvas

Added a medium blur filter.
apply a blur effect on the canvas

Added a medium blur filter.

Say ...
Set the blur effect value to light
Apply a classic effect on this person
Remove the blur effect
Set label to ___
Apply a morph effect to this image
Set the stroke size to 5
Apply a **sepia** effect on this **person**
Label this as ____
When?

• Onboarding
• During a session
• On failure
When?

- Onboarding
- During a session
- On failure

What?

- Number of commands
- Coverage vs. Relevance
- Complexity
- Phrasing
- Parameters
When?
- Onboarding
- During a session
- On failure

What?
- Number of commands
- Coverage vs. Relevance
- Complexity
- Phrasing
- Parameters

Where?
- Pop-up window
- Tooltips
- Embedded in GUI
- Panels
Three interface variants to present command suggestions

Adaptive

Embedded

Exhaustive
Say...

Select the draw tool
Insert a rectangle
Insert an image
Show objects in the image

 Say ...

Add a classic filter on the canvas
Change border color of all rectangles to green
Create a copy of the image
Remove all rectangles from the canvas

Say...

Set label to...
IMAGE_1

Change border color to...
Black

Change border size to...
1

Add the filter...
Choose a Filter...
Grayscale, Classic, Sepia, Blur, Morph

Applied filters:
Examples of commands you can issue:

**Edit Effects or Filters**
- Apply a classic filter on the canvas
- Add a light sepia effect on all people
- Add a grayscale filter here
- Apply a light morph effect on this

**Remove Effects or Filters**
- Set the value of the sepia effect on the person to light
- Make the blur filter on all regions heavy
- Set grayscale effect to heavy
- Make morph filter medium

**Change Shape Fill Color**
- Change the fill color of all rectangles to green
- Set fill of the rectangle to blue
- Set fill to red
- Change fill color of this rectangle to orange

**Change Shape/Image Border Color**
- Set border color of the rectangle to green
- Change the stroke of all images to blue
- Change stroke color to red
- Set border to gray

**Change Shape/Image Border Thickness**
- Set the stroke size of the rectangle to 7
Examples of commands you can issue:

**Set Label**
- + Set label to ______
- + Label this as ______

**Add Effects or Filters**
- + Apply a classic filter on the canvas
- + Add a light sepia effect on all people
- + Add a grayscale filter here
- + Apply a light morph effect on this

**Edit Effects or Filters**
- + Set the value of the sepia effect on the person to light
- + Make the blur filter on all regions heavy
- + Set grayscale effect to heavy
- + Make morph filter medium

**Remove Effects or Filters**
- + Remove the blur effect from the background image
- + Delete the sepia filter from all cars
- + Remove the classic effect here
- + Delete the grayscale filter

**Change Shape Fill Color**
- + Change the fill color of all rectangles to green
- + Set fill of the rectangle to blue
- + Set fill to red
- + Change fill color of this rectangle to orange

**Change Shape/Image Border Color**
- + Set border color of the rectangle to green
Adaptive

Embedded

Exhaustive

Command Suggestions
Add a name filter on target
Make the name filter strength
Set fill color to color
Change border to color
Set stroke size to size
Make count copies
Remove target
Highlight entities in the image
...
Add a name filter on target
Make the name filter strength
Set fill color to color
Change border to color
Set stroke size to size
Make count copies
Remove target
Highlight entities in the image
...

Change border to color
Set fill color to color
Set stroke size to size
Make count copies
...

color = [red, blue, ...]
size = [1-10]
...
Add a name filter on target
Make the name filter strength
Set fill color to color
Change border to color
Set stroke size to size
Make count copies
Remove target
Highlight entities in the image
...

Change border to color
Set fill color to color
Set stroke size to size
Make count copies
...

color = [red, blue, ...]
size = [1-10]
...

Change border to blue
Set fill color to red
Set stroke size to 10
Make 5 copies
...

Command Templates
Filter, Rank, Parameterize
Examples
Add a name filter on target
Make the name filter strength
Set fill color to color
Change border to color
Set stroke size to size
Make count copies
Remove target
Highlight entities in the image
...

Change border to color
Set fill color to color
Set stroke size to size
Make count copies
...
color = [red, blue, ...]
size = [1-10]
...

Change border to blue
Set fill color to red
Set stroke size to 10
Make 5 copies
...

Command Templates
Filter, Rank, Parameterize
Examples

Details in the paper
Color this green
Add a red stroke
Make 2 copies
Delete
Available Operations

Operation Selection

Phrasing Templates

Target Type
Usage Freq.
Display Freq.

Fill
Border
Copy
Delete
Available Operations

Phrasing Templates

Operation Selection

Target Type

Usage Freq.

Display Freq.

Fill

Border

Copy

Delete

Change color to ___

Color this ___

Set the fill to ___

Change fill of ___ to ___

...
Available Operations

Operation Selection

Target Type
Usage Freq.
Display Freq.

Phrasing Templates

Template Selection

Input Type
Usage Freq.
Display Freq.

Fill
Border
Copy
Delete

Change color to ____
Color this ____
Set the fill to ____
Change fill of ____ to ____
...

...
Available Operations

Operation Selection
- Target Type
- Usage Freq.
- Display Freq.

Phrasing Templates
- Input Type
- Usage Freq.
- Display Freq.

Template Selection & Parameterization
- Target State

Fill
- Border
- Copy
- Delete

Change color to ___
___ = [blue, green, red, ...

Color this ___

Set the fill to ___

Change fill of ___ to ___

...
Available Operations

Operation Selection

Target Type
Usage Freq.
Display Freq.

Phrasing Templates

Template Selection & Parameterization

Input Type
Usage Freq.
Display Freq.

Examples

Fill
Border
Copy
Delete

Change color to ___
___ = [blue, green, red, ...]

Color this ___

Set the fill to ___

Change fill of ___ to ___

...
Available Operations

Operation Selection
- Target Type
- Usage Freq.
- Display Freq.

Phrasing Templates

Template Selection & Parameterization
- Input Type
- Usage Freq.
- Display Freq.

Examples
- Color this green
- Add a red stroke
- Make 2 copies
Evaluation
Evaluation

• Between-subjects online study with **24 participants** on UserTesting.com

• **Platform**: Chrome running on a touch-enabled Microsoft Surface Pro

• **Minimal Training**: Short videos about the basic interface and how to invoke suggestions
  (no details about available operations and speech commands)
Evaluation

• **Task:** Three before-after image editing tasks

• **Duration:** 32 min (avg.)

• **Compensation:** $10
Edit the image on the left to make it look like the image on the right. Note that it is okay if your output does not look exactly the same as the target image below but try to make it look as similar as possible.
Speech Usage Summary

• Total of 834 spoken commands issued (avg. 49) during 17/24 sessions (6 exhaustive, 5 adaptive, 6 embedded)
Speech Command Failures

- 369/834 (44%) spoken commands failed:

<table>
<thead>
<tr>
<th>Error %</th>
<th>Error Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>65%</td>
<td>Speech recognition &amp; recording errors</td>
</tr>
<tr>
<td>18%</td>
<td>Phrasing errors</td>
</tr>
<tr>
<td>7%</td>
<td>Operation-object mapping errors</td>
</tr>
<tr>
<td>5%</td>
<td>Unsupported operations</td>
</tr>
<tr>
<td>5%</td>
<td>Parameter errors</td>
</tr>
</tbody>
</table>
Speech Command Failures

- 369/834 (44%) spoken commands failed:

<table>
<thead>
<tr>
<th>Error %</th>
<th>Error Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>65%</td>
<td>Speech recognition &amp; recording errors</td>
</tr>
<tr>
<td>18%</td>
<td>Phrasing errors</td>
</tr>
<tr>
<td>7%</td>
<td>Operation-object mapping errors</td>
</tr>
<tr>
<td>5%</td>
<td>Unsupported operations</td>
</tr>
<tr>
<td>5%</td>
<td>Parameter errors</td>
</tr>
</tbody>
</table>

I'm not sure about that styling command. Maybe this example command might help: *Set fill to red*
Suggestions encourage and aid natural language interaction

<table>
<thead>
<tr>
<th>Suggestions</th>
<th>Exhaustive (avg.)</th>
<th>Embedded (avg.)</th>
<th>Adaptive (avg.)</th>
<th>Overall (avg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>helped me learn how to talk to the system</td>
<td>4</td>
<td>3.67</td>
<td>4.4</td>
<td>4.02</td>
</tr>
<tr>
<td>encouraged me to talk to the system</td>
<td>3.83</td>
<td>3.67</td>
<td>4.2</td>
<td>3.88</td>
</tr>
</tbody>
</table>

*scores between 1-5
5 is “strongly agree”
Explanations for domain specific commands

- Suggestions do not overcome lack of domain knowledge
Explanations for domain specific commands

• Command suggestions as interactive widgets 🧠
Future work

• Validating framework in other application domains (e.g. data visualization)

• Supporting additional command types (e.g. gesture + speech)
Conclusions

• Contextual command suggestions aid discoverability and encourage natural language interaction

• Direct manipulation can be used to teach natural language interaction
Thank you

- Contextual command suggestions aid discoverability and encourage natural language interaction
- Direct manipulation can be used to teach natural language interaction

Arjun Srinivasan (@10_arjun)
Mira Dontcheva
Eytan Adar
Seth Walker

bit.ly/voice-hints