Echolocate

Ryan Dillon, Ishita Goluguri, Alex Sun

ABOUT

Motivation:

- While some deaf people may learn to read lips, even the most accomplished lip readers are only able to understand 25-30% of the conversation without context.
- Only about 500,000 people in the USA use American Sign Language, making it difficult for people suffering from hearing loss to participate in a typical conversation, especially in group settings.

Target Audience:

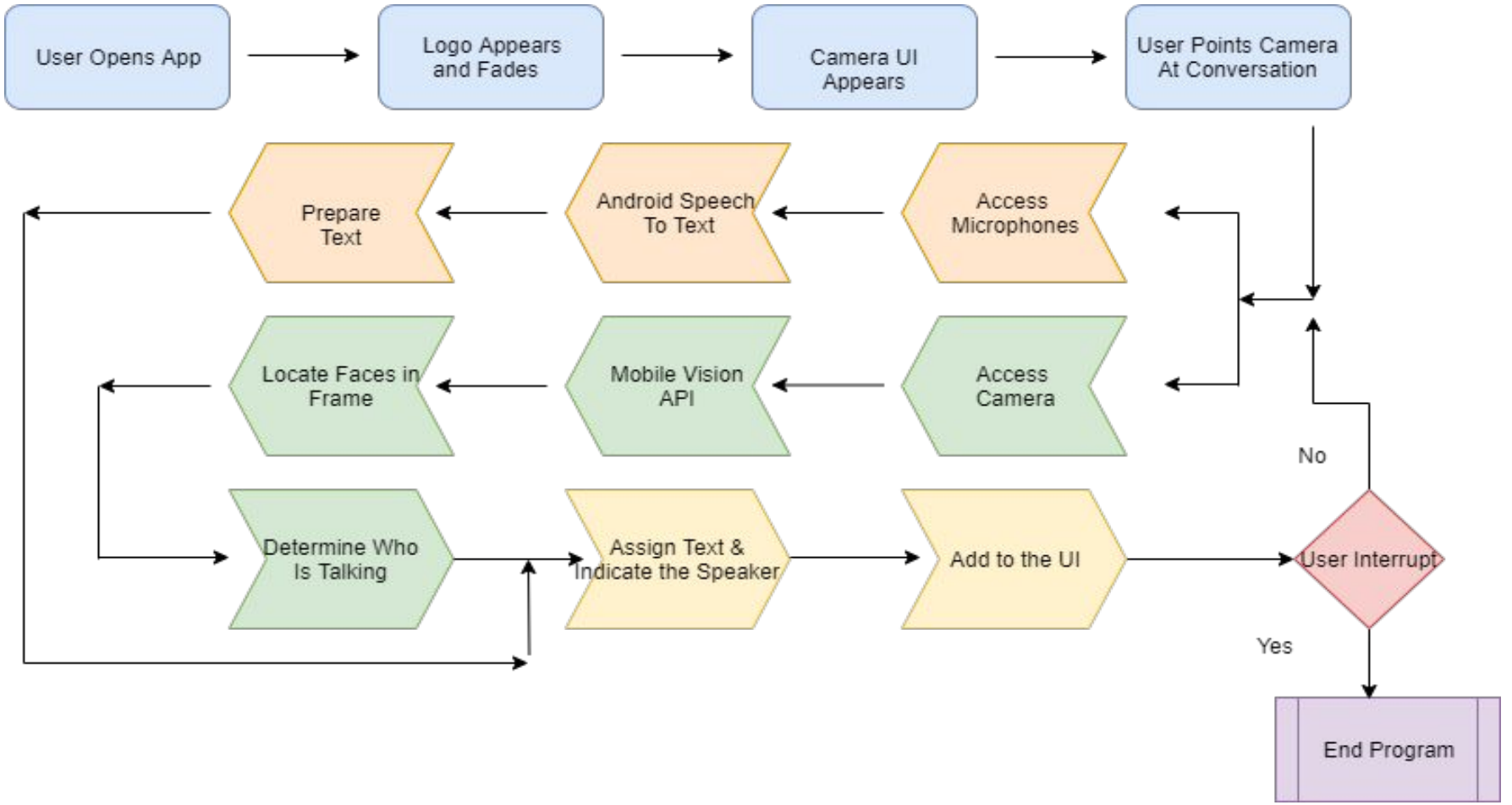
- People that have hearing loss and may have trouble following conversations.

Competitors:

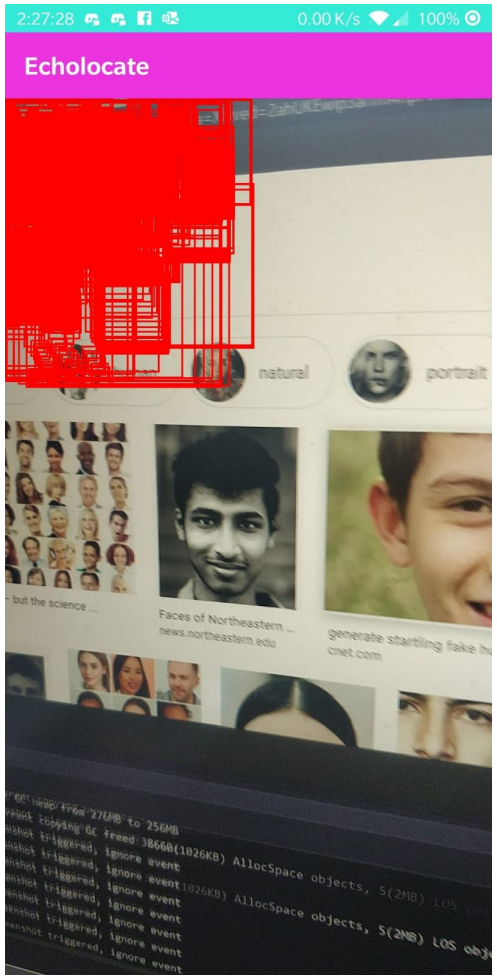
APP DESCRIPTION (& VIDEO)

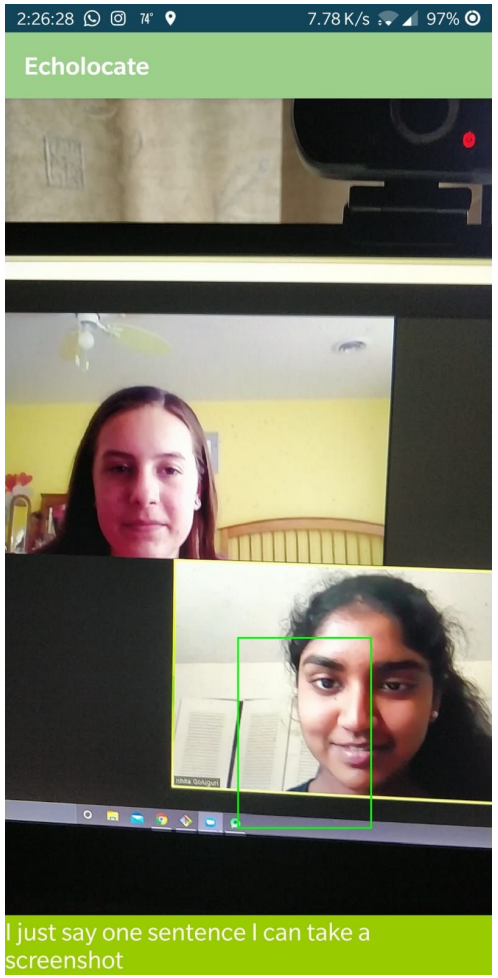
The application identifies who in the group is speaking, and implements speech-to-text translation to display dialogue in a speech bubble. The speaker is highlighted by a green bounding box and the text at the bottom is the speaker's words.

STRUCTURAL FLOW CHART



CHANGES & ITERATIONS





WHO'S SPEAKING ALGORITHM

- Any face within the frame of the camera is detected
- A box is drawn around each person's face and the person's mouth
- The height of both boxes is determined
- The ratio of the mouth height to face height is calculated
- The person with the largest ratio is the one who is talking

TOOLS, TECHNOLOGIES, SERVICES

This application utilizes the following services

- Camera
- Microphone

This application utilizes the following technologies

- Android Speech API
- Android Camera X API
- Firebase ML-Kit Face Detection API

FUTURE EXTENSIONS

- Improve Who's Speaking algorithm
  - Look at multiple frames of each face to determine patterns (requires technological improvements)
- Support rotation (both landscape and portrait)
  - Currently only supports portrait mode
  - Landscape provides more space for text
- Implement Google Cloud Speech-to-Text API
  - Allows live stream conversion
  - Android API is meant for snippets, and stops itself

Google Cloud Platform