*d*STOP:
Notification system estimating arrival of destination with multimodal approach

Presented by Team 9
Talibli Farid, Minhyo Jung, and Jisup Lee
Spring 2022, SNU CSE

Outline

- Refined Our Idea
- Key functions
  - Listen to the name of stops (Speech recognition)
  - Where we are (Location by GPS)
  - When the stop happens (IMU sensors)
- Architecture
- Challenges
  - Multimodal approach to combine sensors
- Remained schedule
- Outputs and how to evaluate

Refined our idea

- Problems
  - Although mobile devices have been highly developed in past decades,
  - we still focus on the map applications when we're on a subway or bus to go to somewhere.
  - It would be **useful** if there are more **reliable destination notification system**.
- Current Solutions
  - Map app will help us to find a path or ETA.
  - For example, however, this may not be exact when in the underground. (no gps signal)
- Target Users
  - Who do not want to focus on the map app on a transportation.
  - Newcomers in Korea or SNU (we may narrow down to the SNU for this project)

Refined our idea

- Our Approach is that
- It's a **Mobile Application** which is a notification system estimating arrival to the destination using multimodal approach
  - Multimodal inputs
    - User's destination
    - Ambient sounds
    - Real-time location information (gps, wifi, etc)
    - Accelerometer, etc.
  - Output
    - Alarm at right destination and on right time
  - Furthermore, naturally **pervasive interface** is what we aim for.

Refined our idea

- Scenarios
  - A user enter the destination (as the name of subway station or bus stop) on the app.
  - Then app runs in the background to gather and analysis.
    - ambient sounds → to extract the name of stops
    - location information such as gps, wifi. → to infer exact location
    - patterns of IMU sensors → how many times a vehicle stops
  - Considering overall infos, app can classify the status quo; arrival or not.
  - Finally, the app can notify a user based on those informations and inferences.

Key functions

- We have 3 key functions to recognise the destination
  - Speech Recognition
  - Location Aware
  - Stop Recognition
- Then link them to analysis

Key functions

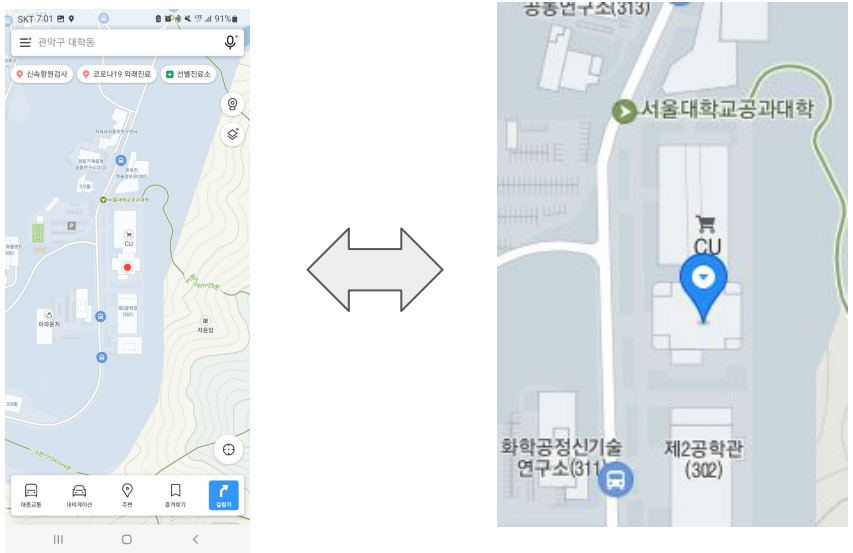
- Speech Recognition
  - Goal of this function is to detect and analyze the sound of station announcement.
  - The system will start a continuous stream of audio to the speech-to-text API.
  - If the text stream returned from the model matches the station the user specified, one of the two criteria for the stop recognition will be fulfilled.
  - One possible API to use is Google Cloud Speech API
  - In case if the API underperforms, we will need to:
    - Gather station announcement data
    - Deploy the model and connect it to the application

Key functions

- Location Aware
  - Goal of this function is tracking the current location
  - There are some open APIS that helps get the location information
    - Google Geolocation API, Kakao map Android API
    - Map view
    - Mark the location
    - Overlay (Lines, Circles, etc …)

Key functions

- Location Aware
  - Simple Example : KakaoMap and marking the location
  - Marked location with API and current location in KakaoMap is the same



Key functions

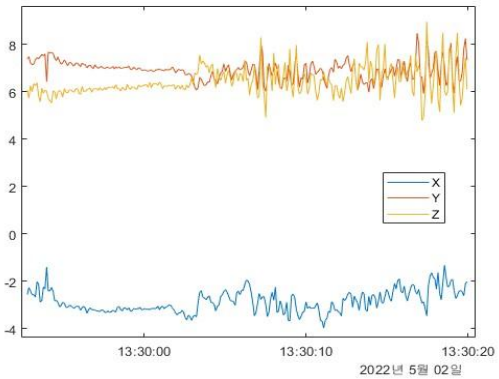
- Stop Recognition
  - Goal of this function is that a device should aware moving status a vehicle.
  - Esp. in case of stops, how many times a vehicle pull up or not.
  - This will help for the system to infer the status of a vehicle in our hypothesis.
  - The recognition is based on an analysis signal stream of IMU sensors.
  - Here are two tasks
    - Gather real-world data
    - Build classify algorithm

Key functions

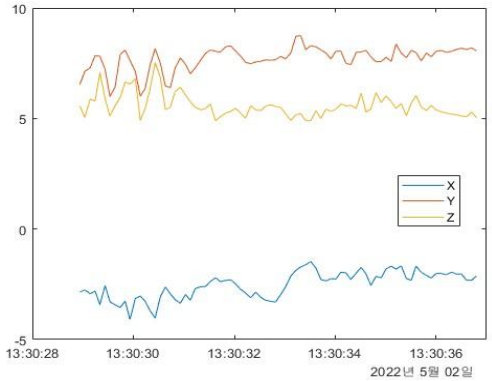
- Stop Recognition - Progress
  - Data Acquisition & Processing (~ing)
    - Gather labelled accelerometer and angular velocity
    - Real-world mobile sensors data
      - iPhone 11 - Matlab mobile (sample rate: 10Hz, duration: ~30 secs.)
      - records acc. and ang. velocity when subway and bus start and stop
  - Data Processing
    - Applying FFT to extract the patterns.

Key functions

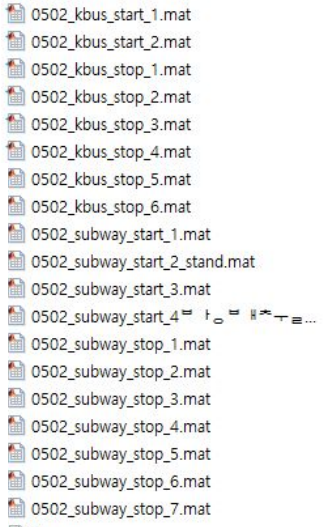
- Stop Recognition - Data
  - Raw examples of start and stop case of Bus-Gwanak 02



Start



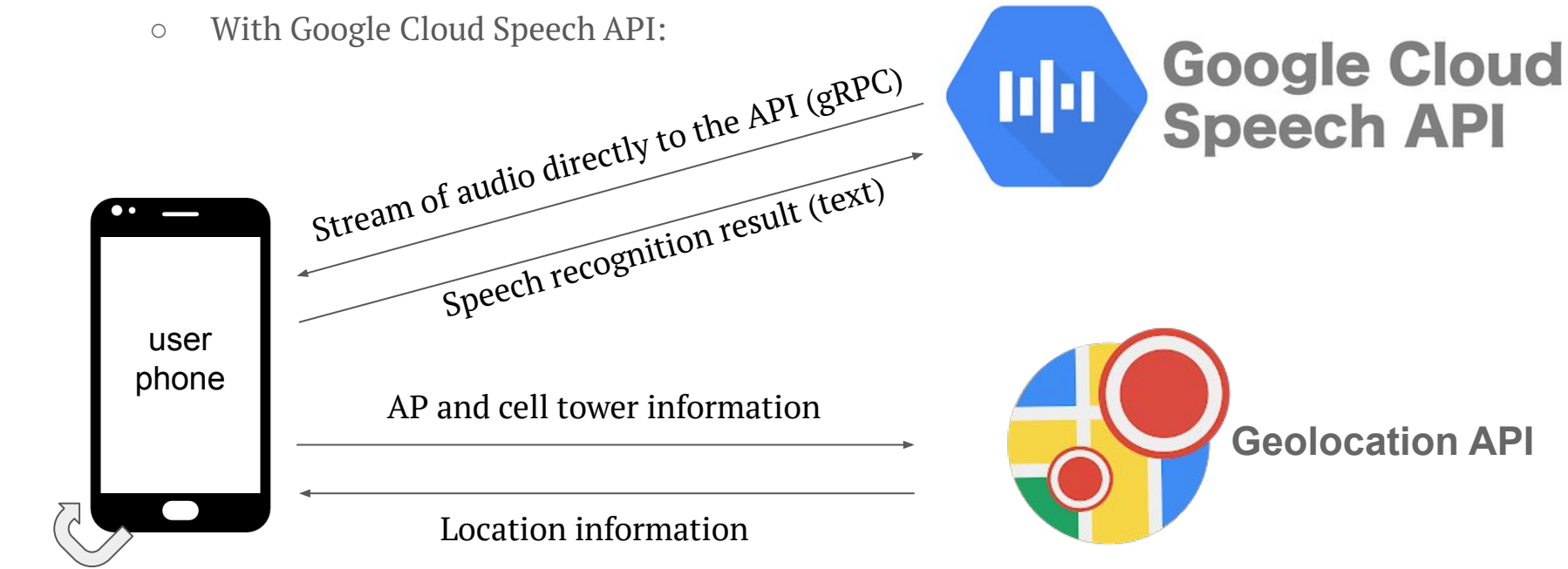
Stop



Part of data list

Architecture

- Basic architecture for our application:
  - With Google Cloud Speech API:



Challenges

- In Architecture
  - Build strong algorithm to combine multimodal approach
- For each key functions
  - Speech recognition
  - Location aware
  - Stop recognition

Challenges

- Speech recognition
  - For Google Cloud Speech API, the accuracy of the speech recognition in realistic setting for a stream might not be as good.
  - In case of a crowded bus, there might be too much noise.
  - In case of making our own model, gathering enough data might be problematic.

Challenges

- Location aware
  - The case when location information providing is interrupted
  - Measure the arrival time

Challenges

- Stop Recognition
  - Data stream from IMU sensor may vary in real-world situations
    - User’s position and moving status.
    - Location of a device
    - Type of a vehicle (Bus and subway shows differently. Stop light in the road)
  - Garbage in, garbage out problem
    - That is; if data is dirty, the inference will be not correct.

Remained schedule

- Detailed plan

No	Title	Owner	Process rate	Phase 1															Phase 2																			Phase 3															Phase 4														
				W4					W5					W6					W7					W8					W9					W10					W11					W12					W13					W14							
				M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F	M	Tu	W	Th	F
1	**Project Design / Class requirements**																																																													
1.1	refine ideas based on a feedback	Minhyo, Jisup, Farid	100%	o	o	o	o	o																																																							
1.2	search necessary techniques	Minhyo, Jisup, Farid	100%						o	o	o	o	o																																																	
2	**Proof of Concept/Preparation Data**																																																												
2.1	collect the announcement data	Farid	50%											o	o	o	o	o																																													
2.2	collect the label of location	Jisup	100%											o	o	o	o	o																																												
2.3	collect the acceleration data	Minhyo	100%											o	o	o	o	o																																												
2.4	test multi-modal approach	Minhyo, Jisup, Farid	80%											o	o	o	o	o																																												
3	**Algorithm Development**																																																												
3.1	Architecture design	Minhyo, Jisup, Farid	80%																o	o	o	o	o	o	o	o	o	o																																		
3.2	Process location information	Jisup	50%																								o	o	o	o	o	o	o	o	o	o	o	o	o	o																							
3.3	Process sound information	Farid	50%																								o	o	o	o	o	o	o	o	o	o	o	o	o	o																							
3.4	Process acceleration data	Minhyo	70%																								o	o	o	o	o	o	o	o	o	o	o	o	o	o																							
4	**Development of Application**																																																												
4.1	Application Architecture Design	Minhyo, Jisup, Farid	80%																o	o	o	o	o	o	o	o	o	o																																		
4.2	Application UX&UI Design	TBD	0%																														o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
4.3	Implementation	TBD	0%																														o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
5	**User Study & Validation**																																																												
5.1	user study	TBD	0%																																																										
5.2	validation	TBD	0%																																																									
5.3	write a report	TBD	0%																																																									

Remained schedule

- What we have done
  - Refine Ideas, search necessary techniques
  - Collect Multi-modal data
  - Architecture design
  - Processing collected data(In progress)
- What we will do
  - Application UX & UI design(All member)
  - Implementation
    - Sound part(Farid)
    - Location part(Jisup)
    - Acceleration part(Minhyo)
  - User Study(All member)

Outcome and how to evaluate it

- Final deliverable
  - A mobile application that:
    - Aids newcomers to SNU in finding their way around campus
    - Allows users to forget about navigation after entering public transportation
    - Is intuitive and easy to use for everyone
- Evaluation
  - Whether the application performs its role with few errors
  - Whether one function can be replaced by other functions when it is stopped
  - User satisfaction

Thank you!

- Q&A