Voice Recognition For Image Retrieval on Android Platforms

Sean Samenfeld-Specht (sean.samenfeldspecht@west.cmu.edu)
Instructor: Ted Selker (ted.selker@west.cmu.edu)

Application Architecture

User speaks, using previously stored tag
RecognizerIntent translates speech to text
Tag looked up in the database
ContentResolver used to lookup image
Image retrieved

Storage

User speaks, using descriptive word or phrase
RecognizerIntent translates speech to text
Tag stored in application database

Retrieval

Application Architecture

User selects image

Inspiration

- Smartphones store more and more media in the form of pictures and video
- Retrieving images or other media is always a search and find process
- Voice recognition is now readily available on the Android platform

Solution

- Store descriptive words and phrases for media that can be used as “breadcrumbs”
- Command pattern used for TagAction, allowing for easy extension to other media
- Recent images shown in custom scroll view for easy tagging

Roadmap

- Implement VerbalTags service so other apps can add “teachable” voice recognition easily
- Allow for tagging multiple images with the same tag (i.e., gallery mode)
- Store tags in the cloud
- Add more supported media (video, app launching)

References