Multimodal Accessibility: Using Computer Vision to improve Speech Recognition in Simon

Yash Shah
First Year B.Tech Student
Dhirubhai Ambani Institute of Information and Communication Technology

June 30th 2012
Current Scenario

- A major obstacle for command and control speech recognition systems is to differentiate commands from background noise.

- Currently Simon listens (take inputs) to whatever sound is made which may include bundles of noise.
How it can be solved?

- It will make lot more sense to activate the recognition only when the user is actively looking at the screen / robot.

- So my project using computer vision determines when to activate / deactivate the recognition using visual cues

- This is strikingly similar to communication between humans.

Yash Shah
Expected Result

- A face detection / recognition context condition that activates or deactivates a microphone / acoustic model in Simon based on the presence of a user.

- Selecting speech models by Recognizing face

- Toggle simon On/Off with gestures

Yash Shah
Thank you

Email: blazonware@gmail.com
Blog: www.yashshah.com/blog
IRC Nick: yashshah (#kdeaccessibility)