Automatic Speech Recognition (ASR) for Developing Regions

Chuck Wooters
wooters@icsi.berkeley.edu
International Computer Science Institute
Outline

• Overview
• Current Status
• Future Plans
ASR for Developing Regions

Can speech technology be used to improve access to information?

- Literacy issues
- A more “user-friendly” interface
- Hands-busy applications?
- May help reduce cost - no mouse/keyboard needed
Challenges for ASR in Developing Regions

• Low computation & memory requirements (for reduced cost)
• Difficult acoustic conditions: mic placement, multiple mic types, environmental noise
• Scarcity of training data
• Sustainability: Spkr indep. vs Spkr dep
• What is a “useful” application?
• Multiple languages
Current Status

• Software “testbed”
• Tamil Data Collection
• Tamil Digit Recognition
Software Testbed

• Workstation-based platform
  • easy cross-task/cross-language comparison
  • rapid experiment turnaround
  • data: English digits, Tamil commands (new!)

• Isolated-word recognizers
  • Rabin/Sergui (token passing), running on a hardware simulation
  • Standard Viterbi decoder (developed by student Thomas Kuo), centralized memory manager, memory-based file system
  • both compatible with standard toolkit (HTK)
Tamil Data Collection

• Data collected in Berkeley and India
• 4650 Recorded tokens
  • Vocab Size: 31 words
  • # Speakers: 30
  • # Repetitions: 5
  • Two mic types
Tamil Digit Recognition

• No Application, just testing the technology/data
• Simple recognizer
  HTK-based, spkr indep, select one of 11 words
• 22 train spkrs, 8 test spkrs
• Initial results: 2.7% error*

* only one word per utterance = easy task

poojyam onru irandu moonru naangu aindhu aaru ezhu ettu onbadhu
Future Work

• Continue with testbed development
• More experiments with Tamil data (data collection in other languages?)
• Experiments with accuracy vs. speed/memor
• Identification of a useful application
Thanks!

wooters@icsi.berkeley.edu