SiMPE: Fourth Workshop on Speech in Mobile and Pervasive Environments

A.A. Nanavati   N. Rajput
IBM India Research Lab
4,Block-C, Institutional Area
Vasant Kunj, New Delhi, India.
{namit,rnitendra}@in.ibm.com

A.I. Rudnicky
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA.
air@cs.cmu.edu

M. Turunen
Dept. of Computer Science
University of Tampere
Tampere, Finland.
mturunen@cs.uta.fi

A. L. Kun
Elec. and Comp. Engg.
Univ. of New Hampshire
New Hampshire, USA
andrew.kun@unh.edu

T. Paek   I. Tashev
Microsoft Research
Redmond
Washington, USA
{timpaek,ivantash}@microsoft.com

ABSTRACT

With the proliferation of pervasive devices and the increase in their processing capabilities, client-side speech processing has been emerging as a viable alternative.

SiMPE 2009, the fourth in the series, will continue to explore issues, possibilities, and approaches for enabling speech processing as well as convenient and effective speech and multimodal user interfaces. One of our major goals for SiMPE 2009 is to increase the participation of speech/multimodal HCI designers, and increase their interactions with speech processing experts.

Multimodality got more attention in SiMPE 2008 than it has received in the previous years. In SiMPE 2007 [3], the focus was on *developing regions*. Given the importance of speech in developing regions, SiMPE 2008 had "SiMPE for developing regions" as a topic of interest. We think of this as a key emerging area for mobile speech applications, and will continue this in 2009 as well.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Voice I/O

General Terms

Design, Human Factors, Experimentation, Performance

Keywords

Speech Processing, Mobile Computing, Pervasive Computing

1. BACKGROUND AND MOTIVATION

The proliferation of pervasive devices has stimulated the development of applications that support ubiquitous access via multiple modalities. Since the processing capabilities of pervasive devices differ vastly, device-specific application

adaptation becomes essential. How does one do speech application adaptation for pervasive devices with different resource (such as memory and power) constraints ?

Traditionally, voice-based applications are accessed using dumb telephone devices through Voice Browsers that reside on the server. With the increase in the processing capabilities of pervasive devices, client-side speech processing is emerging as a viable alternative. Client-side speech processing leads to a reduction in server bottlenecks, round trip costs, transmission errors as well as bandwidth and connectivity requirements. The goal therefore is to leverage a client's capabilities to the fullest. Given the variety of pervasive clients, we need a flexible way of *distributing speech processing* between the client and the server based on the client's capabilities and runtime considerations.

Given the penetration of mobile phones in emerging economies (such as Africa, China and India), and that Voice-driven interfaces to applications have been found to have immense appeal for semi-literate and illiterate users [8], the need for designing flexible, adaptive and robust voice UIs is imminent.

Enabling conversational systems on pervasive devices will require new models, algorithms, systems that are robust across a variety of mobile and ubiquitous devices and dynamic and noisy environments. This will require novel methods for evaluating such systems. Several user-interface related questions will have to be answered.

1.1 Goals of SiMPE

SiMPE has only two ambitious goals:

- To provide a platform that brings together researchers from speech processing, algorithm design, application development and UI design to fuel faster growth of this multi-disciplinary area.

- To pose interesting problems to this community that will foster cross-pollination of ideas and hopefully define the course that SiMPE research should take over the coming years.

2. TOPICS OF INTEREST

All areas that enable, optimise or enhance Speech in mobile and pervasive environments and devices. Possible areas include, but are not restricted to:

- Speech interfaces/applications for Developing Regions

- Speech User Interaction in Cars

- Multilingual Speech Recognition

- Robust Speech Recognition in Noisy and Resource-constrained Environments

- Memory/Energy Efficient Algorithms

- Multimodal User Interfaces for Mobile Devices

- Protocols and Standards for Speech Applications

- Distributed Speech Processing

- Mobile Application Adaptation and Learning

- Prototypical System Architectures

- User Modelling

2.1 Seed Questions

1. How to do speech recognition in noisy environments ?

2. What are the usability issues with speech on mobile devices ?

3. Are there efficient ways of enabling voice search on the device?

4. How to make voice UIs flexible and adaptive ?

5. Are there any novel and easier ways to handle multiple languages and dialects ?

6. How do we construct speech systems with small footprints of memory and power consumption ?

7. How can we distribute processing more efficiently given the increased available computing power on handhelds ? How do we trade this off with a remote server to conserve energy ?

8. How do we make such devices adapt automatically to the user, task and environment ?

9. What kind of components and frameworks should be built to enable rapid application creation ?

10. How can we leverage context (such as location) to make more intelligent UIs that reduce the 'cognitive burden' of semi-literate/illiterate users ?

11. How can we design scalable speech applications?

12. How to design speech interfaces that are safe and robust for using in cars?

3. INTENDED AUDIENCE

This multi-disciplinary burgeoning area invites researchers interested in any aspect of the intersection of Speech processing and Mobile computing — speech recognition, speech synthesis, multimodal interfaces, mobile HCI, distributed speech processing, mobile applications, voice user interface design, memory/energy efficient algorithms, UI design — to meet and pave the way forward. We anticipate a good mix of international industrial and academic participation which should lead to lively discussions.

4. AGENDA AND SCHEDULE

Before
We will advertise the workshop on several mailing lists, the SiMPE wiki, as well as the SiMPE website. Over the last two years, we have created a mailing list of about 200 researchers working in this area. The tentative Program Committee has already been decided.

Same as last year, accepted papers will be made available to all participants before the conference, to encourage more lively discussions.

On September 15, 2009
We are proposing to have a full-day workshop. This will be broken into 4 sessions, where three sessions will be based on a theme: core speech processing, applications and usability. The fourth session will be divided into two parts. In the first part, we will review the state-of-the-art of SiMPE. The second part will address methods and pose problems to foster collaboration between the speech community and the UI designers. At the end of each session, there will be a short discussion and a summary statement by the session panel.

After
We will summarise the findings of the workshop on the SiMPE webpage [1] and encourage the participants to contribute to the SiMPE wiki [5]. Depending upon the type of submissions, we will try to identify an appropriate venue for publishing the papers.

5. ORGANISING COMMITTEE

- Amit Anil Nanavati, *IBM Research, India Research Laboratory.*
  namit@in.ibm.com

- Nitendra Rajput, *IBM Research, India Research Laboratory.*
  rnitendra@in.ibm.com

- Alexander I. Rudnicky, *Carnegie Mellon University.*
  air@cs.cmu.edu

- Markku Turunen, *University of Tampere, Finland.*
  mturunen@cs.uta.fi

- Andrew Kun, *University of New Hampshire, USA.*
  andrew.kun@unh.edu

- Tim Paek, *Microsoft Research, USA.*
  timpaek@microsoft.com

- Ivan Tashev, *Microsoft Research, USA.*
  ivantash@microsoft.com

6. PROGRAMME COMMITTEE

- Abeer Alwan, UCLA, USA.

- Alex Acero, Microsoft, USA.

- Matt Jones, Swansea University, UK.

- Anirudha Joshi, IIT Bombay, India.

- Yoon Kim, Novauris Technologies, USA.

- Lars Bo Larsen, Aalborg University, Denmark.

- Gary Marsden, Univ. of Cape Town, South Africa.

- Michael McTear, University of Ulster, Ireland.

- Shrikanth S. Narayanan, USC, USA.

- Sharon Oviatt, Incaa Designs, USA.

- Antonio M. Peinado, Universidad de Granada, Spain.

- Ian Pitt, University College Cork, Ireland.

- Jahanzeb Sherwani, CMU, USA.

- Yaxin Zhang, Nuance Communications, China.

7. REFERENCES

[1] http://research.ihost.com/SiMPE
[2] http://research.ihost.com/SiMPE/2008
[3] http://research.ihost.com/SiMPE/2007
[4] http://research.ihost.com/SiMPE/2006
[5] http://simpe.wikispaces.com
[6] ITU World Telecommunication Indicators Database.
[7] http://www.cs.uta.fi/hci/spi/Jaspis/
[8] T. S. Parikh and E. D. Lazowska. Designing an Architecture for Delivering Mobile Information Services to the Rural Developing World. In Proc. Intl. Conf. on World Wide Web (WWW), May 2006.