CALL FOR PARTICIPATION

The Blizzard Challenge 2013
Evaluating corpus-based speech synthesis on common databases
Simon King, Alan W Black, Keiichi Tokuda and Kishore Prahallad

January 9th 2013

In order to better understand and compare research techniques in building corpus-based speech synthesizers on the same data, the annual Blizzard Challenges 2005–2012 were held. We are now pleased to call for participation in the ninth challenge: Blizzard Challenge 2013. We are also launching the pilot phase of the 2014 challenge. The basic challenge is to take the released speech data, build synthetic voices, and synthesize a prescribed set of test sentences. The output from each synthesizer is then evaluated through extensive listening tests. The following is an outline description of the challenge – for full details and the rules of participation, please see the web pages:

- English

  - Audiobook data kindly provided by The Voice Factory, from a single female speaker, provided as approximately 300 hours of chapter-sized mp3 files, plus approximately 19 hours of non-compressed wav files. The wav files have been segmented into sentences and aligned with the text by Lessac Technologies, Inc.

  - Task 2013-EH1 – build a voice from the provided unsegmented audio; text is not provided, so must be obtained by participants from the web (e.g., Project Gutenberg) and aligned with the audio

  - Task 2013-EH2 – build a voice from the provided segmented audio; the accompanying aligned text may be used, or text may be obtained from the web

- Indian Languages (pilot phase for the 2014 challenge)

  - About 1 hour of speech data in each of four Indian languages (Hindi, Bengali, Kannada and Tamil) recorded by native non-professional speakers in quiet office environments. Text is provided in UTF-8 format. No other information, such as segment labels, is provided.

  - Task 2013-IH1 – Build one voice in each language from the provided data.

Preliminary information about the Blizzard Challenge 2014: About 1 hour of speech data in each of 7–9 Indian languages will be provided.

Blizzard Workshop: The systems will be presented by the participants at the Blizzard Challenge 2013 workshop, which will be a satellite of SSW 2013 in Barcelona, Spain: http://synsig.org/index.php/Speech_Synthesis_Workshop_(SSW)

Website: http://www.synsig.org/index.php/Blizzard_Challenge_2013

Registration: Interested parties should register as soon as possible, by mailing blizzard@festvox.org. They should identify a contact person in their team, and provide email and postal addresses. A registration fee is payable for each system submitted for evaluation – details can be found on the website.

Licenses: the license for the released data can be found via the Blizzard website. Data will be released to each participant once the appropriate license has been agreed to.

Mailing list: it is essential to join the blizzard-discuss mailing list, which is used to send important information to participants. Instructions are on the Blizzard website.

Provisional timeline:

Jan 15	2013	Databases released
Mar	2013	test sentences released to participants
Mar	2013	participants submit synthetic speech
Apr	2013	main evaluation system goes live (standard MOS, SIM, WER tests)
Jun	2013	end of evaluation period
Jul	2013	participants submit their workshop papers
Aug 25-29	2013	Interspeech 2013, Lyon, France
Aug 31 – Sep 2	2013	SSW8, Barcelona, Spain
Sep 3	2013	Blizzard 2013 workshop, Barcelona, Spain

Further Information: For further information please contact blizzard@festvox.org.

A description of the challenge was published at Interspeech 2005:
http://www.festvox.org/blizzard/bc2005/IS051946.PDF