

Multimedia Content with a Speech Track: ACM Multimedia 2010 Workshop on Searching Spontaneous Conversational Speech

Martha Larson
Delft University of Technology
Delft, Netherlands
m.a.larson@tudelft.nl

Roeland Ordelman
University of Twente
Enschede, Netherlands
& Sound and Vision
Hilversum, Netherlands
rordelman@beeldengeluid.nl

Florian Metze
Carnegie Mellon University
Pittsburgh, PA, USA
fmetze@cs.cmu.edu

Wessel Kraaij
Radboud University
Nijmegen, Netherlands
& TNO, Delft, Netherlands
wessel.kraaij@tno.nl

Franciska de Jong
University of Twente
Enschede, Netherlands
f.m.g.dejong@ewi.utwente.nl

Categories and Subject Descriptors: H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing

General Terms: Algorithms, Design, Languages

Keywords: spoken content, speech, speech recognition, audio-visual retrieval, multimedia access

1. INTRODUCTION

When multimedia content has a speech track, a whole array of techniques involving speech recognition and analysis becomes available for indexing and structuring and can provide users with improved access and search. The set of new domains standing to benefit from these techniques encompasses talkshows, lectures, meetings, interviews, debates, podcasts, call center recordings, cultural heritage archives, social video on the Web and spoken natural language queries. The success that has been achieved in the area of broadcast news retrieval [3] has yet to be extended to providing users with access and retrieval in all domains. New domains are challenging because they often involve unplanned, spontaneously produced speech and/or a conversational setting.

The goal of the SSCS workshop series on *Searching Spontaneous Conversational Speech* is to help push forward research on these new domains by promoting cross-pollination between researchers from the areas of speech recognition, audio processing, multimedia analysis and information retrieval. After the success of the initial SSCS workshop held at SIGIR 2007 [2], a second workshop followed at SIGIR 2008 [4]. In order to actively connect with the multimedia community, SSCS 2009 [5] was held at ACM Multimedia.

2. SPEECH RETRIEVAL CHALLENGES

The challenges to be faced in order to provide users with improved speech-based access to multimedia collections are notoriously recalcitrant, but also well documented in the lit-

erature [1, 7, 6]. SSCS 2010¹ focuses in particular on queries and results whose length is quite short (cf. [1]), on the issue of Out-Of-Vocabulary (OOV) terms (cf. [7]), on techniques for structuring speech media via segmentation (cf. [6]), and on confronting the variation and irregularities characterizing unplanned speech. A session devoted to demonstration of speech retrieval systems is included to keep workshop discussion centered on real-life applications and user needs.

3. REFERENCES

- [1] J. Allan. Perspectives on information retrieval and speech. In A. R. Coden, E. W. Brown, and S. Srinivasan, editors, *Information Retrieval Techniques for Speech Applications*, volume 2273 of *LNCS*, pages 323–326. Springer Berlin Heidelberg, 2002.
- [2] F. de Jong, D. Oard, R. Ordelman, and S. Raaijmakers. Spoken content retrieval: Searching spontaneous conversational speech. *ACM SIGIR Forum*, 41(2):104–108, 2007.
- [3] J. Garofolo, G. Auzanne, and E. Voorhees. The TREC Spoken Document Retrieval track: A success story. In *RIA0*, pages 1–20, 2000.
- [4] J. Kohler, M. Larson, F. de Jong, W. Kraaij, and R. Ordelman. Spoken content retrieval: Searching spontaneous conversational speech. *ACM SIGIR Forum*, 42(2):67–77, 2008.
- [5] M. Larson, R. Ordelman, F. de Jong, J. Kohler, and W. Kraaij. Multimedia with a speech track: Searching spontaneous conversational speech. *ACM SIGIR Forum*, 44(1):75–80, 2009.
- [6] L.-S. Lee and B. Chen. Spoken document understanding and organization. *Signal Processing Magazine, IEEE*, 22(5):42–60, 2005.
- [7] P. J. Moreno, J. M. Van Thong, B. Logan, and G. J. F. Jones. From multimedia retrieval to knowledge management. *Computer*, 35(4):58–66, 2002.

Copyright is held by the author/owner(s).
MM'10, October 25–29, 2010, Firenze, Italy.
ACM 978-1-60558-933-6/10/10.

¹We would like to thank the project *CHoral: Access to oral history* for its support. We also acknowledge the EU NoE PetaMedia, particularly for its support of the first author.