Special Issue on
New Frontiers in Rich Transcription

A rich transcript is a transcript of a recorded event along with metadata to enrich the word stream with useful information such as identifying speakers, sentence units, proper nouns, speaker locations, etc. As the volume of online media increases and additional, layered content extraction technologies are built, rich transcription has become a critical foundation for delivering extracted content to down-stream applications such as spoken document retrieval, summarization, semantic navigation, speech data mining, and others.

The special issue on "New Frontiers in Rich Transcription" will focus on the recent research on technologies that generate rich transcriptions automatically and on its applications. The field of rich transcription draws on expertise from a variety of disciplines including: (a) signal acquistion (recording room design, microphone and camera design, sensor synchronization, etc.), (b) automatic content extraction and supporting technologies (signal processing, room acoustics compensation, spatial and multichannel audio processing, robust speech recognition, speaker recognition/diarization/tracking, spoken language understanding, speech recognition, multimodal information integration from audio and video sensors, etc.), (c) corpora infrastructure (meta-data standards, annotations procedures, etc.), and (d) performance benchmarking (ground truthing, evaluation metrics, etc.) In the end, rich transcriptions serve as enabler of a variety of spoken document applications.

Many large international projects (e.g. the NIST RT evaluations) have been active in the area of rich transcription, engaging in efforts of extracting useful content from a range of media such as broadcast news, conversational telephone speech, multi-party meeting recordings, lecture recordings. The current special issue aims to be one of the first in bringing together the enabling technologies that are critical in rich transcription of media with a large variety of speaker styles, spoken content and acoustic environments. This area has also led to new research directions recently, such as multimodal signal processing or automatic human behavior modeling.

The purpose of this special issue is to present overview papers, recent advances in Rich Transcription research as well as new ideas for the

direction of the field. We encourage submissions about the following and other related topics:

- Robust Automatic Speech Recognition for Rich Transcription
- Speaker Diarization and Localization
- Speaker-attributed-Speech-to-Text
- Data collection and Annotation
- Benchmarking Metrology for Rich Transcription
- Natural language processing for Rich Transcription
- Multimodal Processing for Rich Transcription
- Online Methods for Rich Transcription
- Future Trends in Rich Transcription

Submissions must not have been previously published, with the exception that substantial extensions of conference papers are considered.
Submissions must be made through IEEE's manuscript central at:
http://mc.manuscriptcentral.com/sps-ieee
Selecting the special issue as target.

Timeline:
    EXTENDED Submission deadline: 1 September 2010
    Notification of acceptance: 1 January 2011
    Final manuscript due: 1 July 2011

Guest editors:
    Gerald Friedland, fractor@icsi.berkeley.edu
    Jonathan Fiscus, jfiscus@nist.gov
    Thomas Hain, T.Hain@dcs.shef.ac.uk
    Sadaoki Furui, furui@cs.titech.ac.jp