IAF SPACE OPERATIONS SYMPOSIUM (B6)
Interactive Presentations - IAF SPACE OPERATIONS SYMPOSIUM (IP)

Author: Dr. Cesar Augusto Guzman Alvarez
VisionSpace Technologies, Germany, research@cguz.org

Dr. William Jones
Astroscale Ltd, United Kingdom, william.jones@community.isunet.edu
Mr. Tiago Nogueira
VisionSpace Technologies, Germany, tiago.nogueira@visionspace.com
Mr. Peter Collins
ESA - European Space Agency, Germany, petercollins00@googlemail.com
Mr. James Eggleston
European Space Agency (ESA/ESOC), Germany, James.Eggleston@esa.int
Mr. João Guerreiro
VisionSpace Technologies, Germany, joao.guerreiro@visionspace.com
Mr. Pedro Silva
1D works UG, Germany, pedro@1d.works

SPEECH TO TEXT FOR AUTOMATIC TRANSCRIPTION AND INDEXING OF VOICE LOOPS AT ESOC

Abstract

Communications in spacecraft operations are most often coordinated through voice loops; discrete audio channels through which information for specific missions or domains is exchanged. The content of voice loops provide great insight into the proceedings of spacecraft operations, however the current paradigm at ESOC is to store them as raw audio files, making the information contained within difficult to extract.

This activity develops an Automatic Speech Recognition (ASR) system that improves the current system by indexing transcriptions of audio files in a searchable database. Storing transcriptions of audio communications provides numerous benefits such as support for error analysis and quality assurance processes, as text is easily searchable with keywords and/or datetimes, contrasting with substantial effort required when listening to every workload when considering time required to manually type out logs, which is especially useful during intensive operational periods.

The current ASR prototype is composed of three main elements: 1) a client Raspberry Pi interfacing with the ESOC voice loop system which records and sends audio to 2) a backend server which stores and transcribes the audio files using IBM's Watson Speech To Text (STT) engine, also hosting 3) the frontend dashboard UI that allows operators to view, recording during the estimated time period of the problem. ASR also reduces operator search, and edit transcriptions.

Due to the highly specific and technical language domain of spacecraft operations (procedure codes, unique mission names, component names etc), as well as a mix of accents present in the international environment of ESOC, substantial customisation of the IBM STT system is required in order to generate accurate transcriptions. This customisation is achieved by passing as training input; out-of-vocabulary words/phrases, ground truth manual transcriptions, and real-world audio recordings of the operator's communications. This allows the STT engine to recognise and understand words/phrases outside the typical conversational domain. For the two operational domains currently served by the ASR system, the

word error rates obtained range from 0.10 to 0.21, representing transcriptions that are understandable to a human reader. Future work includes full integration to the voice-loop infrastructure, bypassing the Raspberry Pi, in order to enhance scalability to multiple missions whilst reducing operational complexity.