Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 89 (2016) 632 – 639

Procedia Computer Science

Twelfth International Multi-Conference on Information Processing-2016 (IMCIP-2016)

Development of Voice Activated Ground Control Station

D. K. Rahul[a,*], S. Veena[b], H. Lokesha[b], S. Vinay[c], B. Pramod Kumar[c], C. M. Ananda[b] and Vinod B. Durdi[a]

[a]*Dayananda Sagar Institute of Technology, Bengaluru, Karnataka, India 560 078*
[b]*CSIR-National Aerospace Laboratories, Bengaluru, Karnataka, India 560 017*
[c]*Karnataka Regional Engineering College, Suratkal, Karnataka, India 575 025*

Abstract

This paper chronicles the development of Automatic Speech Recognition (ASR) system that can be integrated to Ground Control Station (GCS) of MAVs to achieve voice activation. The first part of the paper highlights the nature of aerospace speech signals and hence the issues to be considered while designing a voice activated aerospace application. The second part describes the development and integration of an ASR capability to the GCS.

Peer-review under responsibility of organizing committee of the Organizing Committee of IMCIP-2016

Keywords: Automatic Speech Recognition; Ground Control Station; Hidden Markov Model; Micro Air Vehicle.

1. Introduction

Speech is a natural and intuitive means of communication and hence an efficient alternative for information management. Recent initiatives try to exploit its potential in aerospace domain for control and command operations. Some of the cited advantages are freeing operator's hands and efficient conveying of messages[1].

Effective interpretation of speech by machine is the key to success of speech based applications. Significant advances in the hardware required for development of intelligent search algorithms is able to achieve a high level of performance, making it applicable for many critical applications[2]. Since Aerospace applications require a highly reliable ASR system which is capable of working in adverse and stressful conditions, being transparent to the user, and understanding its users accurately without having to tailor their individual speech and vocabulary to suit the system, the system development becomes a challenging task[3].

A realistic ASR implies the capability of a machine to understand a spoken utterance on any subject by all speakers in different environments, which is quite challenging. The success is determined by the capability of algorithm set employed for capturing the inter and intra-variability of speakers, the nature of the utterance (continuous speech versus isolated words), the vocabulary size and the complexity of the language against different environmental conditions under which the recognition operation is performed[4].

This paper describes the development of ASR algorithms for voice controlled Ground Control Station application.

*Corresponding author. Tel.: +91 9538953933.
E-mail address: rahuldk.dkr@gmail.com

Peer-review under responsibility of organizing committee of the Organizing Committee of IMCIP-2016
doi:10.1016/j.procs.2016.06.026

Nomenclature

MAV – Micro Air Vehicle
GCS – Ground Control Station
ASR – Automatic Speech Recognition
ATC – Air Traffic Controller
SAD – Speech Activity Detection
KWS – Keyword Spotting
MFCC – Mel Frequency Cepstral Coefficients
HMM – Hidden Markov Model
C_h, $\delta^{(1)}$, $\delta^{(2)}$ – Cepstral, Delta and Acceleration coefficients
$\boldsymbol{\pi}$, **A**, **B** – Initial state distribution vector, State Transition probability matrix and Continuous observation probability density function matrix of HMM
C# – C sharp coding language

2. Application of Speech Technology in Aerospace

The applications of ASR in aviation are rapidly evolving and some of them are, Voice Activated GCS[1,3], Training for military (or civilian) Air Traffic Controllers (ATC)[5] and Voice controlled cockpit[6]. Even though ASR applications are quite popular in other domains, developing it for aerospace applications is quite challenging, owing to stringent requirements.

It is very important to understand the nature of speech used in aerospace to achieve effective speech based system design. The speech here is highly structured and an average of 200 vocabularies is sufficient, hence difficulty involved with speech recognition is relatively low[7]. DARPA[8] has classified the speech processing algorithms required for aerospace applications under Speech Activity Detection (SAD), Language Identification (LID), Key Word Spotting (KWS) and Speaker Identification (SID). However, the presence of adverse conditions like high-noise, high-stress, g-force, pressure breathing conditions, noise, wind and gust degrades the performance of the speech recognition systems[7]. Further, the other main factors to be considered are its real time performance, recognition rate and robustness. Javier *et al.*[9] has thrown light on various aspects to be considered in development of robust speech interface for aerial vehicles. Mark *et al.*[1] presents the evaluation results of speech control of UAVs and concludes that speech input was superior to manual operations; however he emphasizes the need for robust speech recognition.

3. Components of Speech Processing System

Figure 1 shows the block diagram of an ASR system. A brief description is given to understand the functionality of each sub block. The main purpose of the front end is efficient feature extraction from voiced part of the speech signal. The speech endpoint detection and silence removal are adapted for dimensionality reduction in speech that facilitates

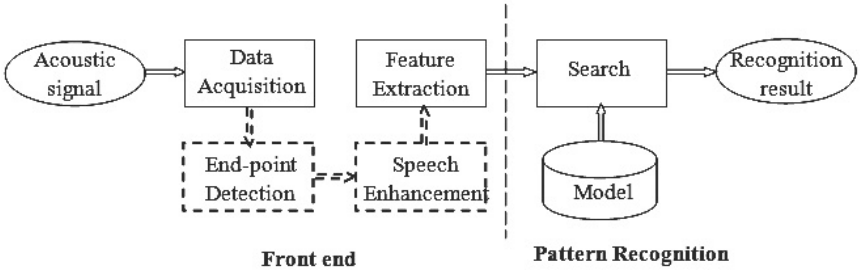

Fig. 1. Blocks of an Automatic Speech Recognition System.