



ICAO

*International Civil Aviation Organization***SEVENTH MEETING OF THE ASIA/PACIFIC AIR
TRAFFIC MANAGEMENT AUTOMATION SYSTEM
TASK FORCE (ATMAS TF/7)***Bangkok, Thailand 2-4 June 2026*

Agenda Item 4: ATM Automation System Implementation Experience by States

4.6. Development of New Technology

A SAFETY PROTECTION SYSTEM BASED ON SPEECH RECOGNITION

(Presented by China)

SUMMARY

This paper introduces a safety protection system based on speech recognition, which transforms the conversations between controllers and pilots into text. It also compares the controllers' instructions with integrated track information from ATMAS to identify discrepancies, performs alert calculations.

1. INTRODUCTION

1.1 Air-ground communication is a fundamental safety-critical element of civil aviation. The integrity of information exchange between air traffic controllers and pilots is paramount to flight safety. Currently, the monitoring of these communications remains predominantly manual and is subject to human performance limitations, including fatigue, cognitive workload, and environmental factors such as radio interference or linguistic variations.

1.2 At ATMAS TF/6 in 2025, China presented an Information paper titled Research on Speech Recognition Technology in ATMAS, which introduced the application of speech recognition in ATMAS. Since then, further research has been undertaken to explore the integrated processing of voice data with ATMAS surveillance track information. This integration facilitates real-time verification between issued control instructions and actual flight trajectories, thereby enhancing conformance monitoring.

2. DISCUSSION

Hardware Architecture

2.1 The system is engineered with an independent architecture, comprising speech recognition processors, data processors, application servers, and speech recognition clients. The infrastructure further integrates crucial networking and security components, such as switches, firewalls, and network gateways, as depicted in Figure 1.

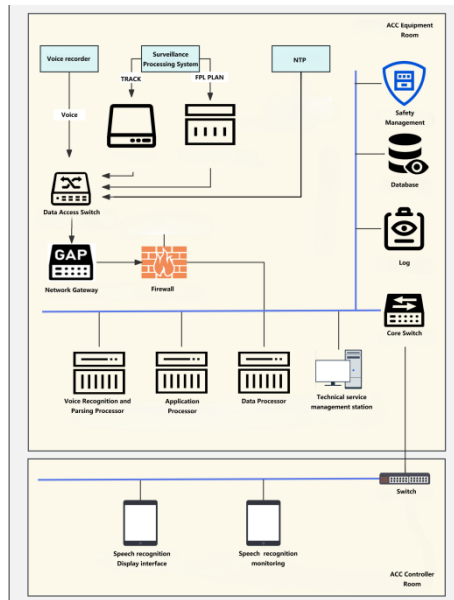


Figure1. System Hardware Architecture

System Algorithm Design

2.2 The Voice Activity Detection (VAD) technique mitigates ambient noise and enhances the speech signal, ensuring effective segmentation of continuous air-ground communications.

2.3 Automatic Speech Recognition (ASR) decodes voice data by the utilization of acoustic and language models, facilitating the conversion of audio inputs into text-based data

2.4 Natural Language Understanding (NLU) acts as the bridge between raw audio transcription and actionable ATM data, enabling the system to understand the underlying operational intent rather than just the literal words.

2.5 The Scenario Reasoning and Alerting enable the automated identification and instantaneous alerting of potential risks.

Functional Module

2.6 The data ingestion and preprocessing module is responsible for the access and preprocessing of multi-source data, including quality analysis, noise reduction and segmentation of voice signals, as well as the parsing of surveillance data.

2.7 Controller Voice Real-Time Monitoring and Alerting Function: Voice information is converted into text content by a speech recognition engine and displayed on the human-machine interface, with critical information highlighted. The alerting function can be configured according to airspace. Key functions include the following:

a. Readback Inconsistency Alert: The system compares the controller instructions with the pilot readback in real time to identify whether critical parameters (e.g., altitude, heading, speed, frequency, etc.) are consistent, as shown in Figure 2. When detect a deviation, the system highlights the inconsistent items on the human-machine interface, as shown in Figure 3, and issues an audible and visual alert to prompt the controller to take action.

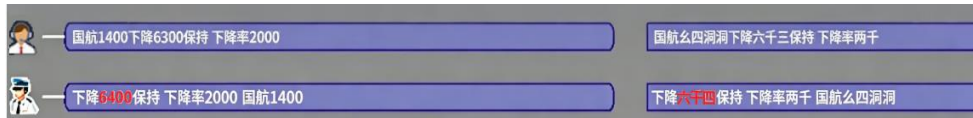


Figure 2 Flight Crew Readback Inconsistency Detection

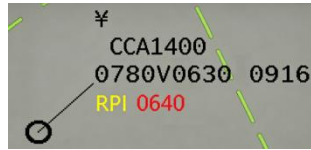


Figure 3 Readback Inconsistency Alert HMI Display

b. Instruction Inconsistency Alert: Once a controller instruction is issued, the system verifies its consistency with the surveillance track. For example, if the controller instruction is inconsistent with the Cleared flight level of the surveillance track (as shown in Figures 4 and 5), or with the aircraft’s heading, the system triggers an inconsistency alert.



Figure 4 Instruction Information Recognition

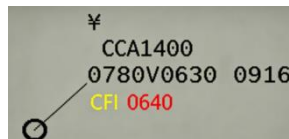


Figure 5 Instruction Altitude Inconsistency Alert

c. Frequency Management Alert: The system tracks the flight handover process and identifies whether the controller has completed the frequency transfer within the offline-configured time. If the transfer is overdue, the system prompts “Handover Incomplete” or “Frequency Transfer Delayed,” assisting the controller in monitoring handover process.

Test Results

2.8 Speech recognition accuracy was evaluated with over 500 hours of a voice recordings from area control center (ACC) controller, covering Chinese, English, and Chinese-English mixed communications.

2.9 The text output generated by the Automatic Speech Recognition (ASR) module was input into the Natural Language Understanding (NLU) module to assess the accuracy of controller instruction. The evaluation results for instruction intent accuracy are summarized in Table 1.

Instruction Feature	Accuracy
Instruction Intent (e.g., Altitude Adjustment, Heading Change)	96.5%
Critical Parameters (e.g., Call Sign, Altitude Value)	97%
Speaker Role Identification (e.g., Controller, Pilot)	98.2%

Table 1 Instruction Feature Accuracy Verification Results

2.10 Alerting Functions: In the readback inconsistency alert test, 500 test commands with readback errors were simulated. The system generated 492 valid alerts, 5 missed alerts and 3 false alerts, achieving an alert accuracy of 98.4% and an average response delay within 1 second. In the instruction inconsistency alert test, 300 false surveillance track samples were simulated. The system generated 287

valid alerts, 6 missed alerts and 7 false alerts. The alert accuracy reached 95.7%, with an average response delay less than 0.8 seconds.

2.11 Conclusions and Prospects

- a) The system realizes accurate voice recognition and can provide early alerts for potential air traffic control operational risks.
- b) In the next step, we will formulate interface specifications between the ATMAS system and the speech recognition system. Text data generated by speech recognition system will be transmitted to ATMAS, enabling the comparison of voice data, track data, and flight plan data, and achieving integrated display of alert information within the ATMAS system.

3. ACTION BY THE MEETING

3.1 The meeting is invited to:

- c) note the information contained in this paper; and
- d) discuss any relevant matter as appropriate
