

技術論文

J. of The Korean Society for Aeronautical and Space Sciences 41(12), 1011-1017(2013)

DOI:http://dx.doi.org/10.5139/JKSAS.2013.41.12.1011

전투기용 음성명령 시스템에 대한 연구

김성우*, 서민기*, 오영환**, 김봉규***

A Study on Cockpit Voice Command System
for Fighter Aircraft

Seongwoo Kim*, Mingi Seo*, Yunghwan Oh** and Bonggyu Kim***

LIG Nex1 Avionics R&D Lab*, Korea Advanced Institute of Science and Technology**,
Agency for Defense Development The 7th R&D Institute***

ABSTRACT

The human voice is the most natural means of communication. The need for speech recognition technology is increasing gradually to increase the ease of human and machine interface. The function of the avionics equipment is getting various and complicated in consequence of the growth of digital technology development, so that the load of pilots in the fighter aircraft must become increased since they don't concentrate only the attack function, but also operate the complicated avionics equipments. Accordingly, if speech recognition technology is applied to the aircraft cockpit as regards the operating the avionics equipments, pilots can spend their time and effort on the mission of fighter aircraft. In this paper, the cockpit voice command system applicable to the fighter aircraft has been developed and the function and the performance of the system verified.

초 록

음성은 사람의 가장 자연스러운 정보 전달 수단이며, 음성인식 기술은 사람이 기계를 사용하는데 있어 편의성을 높이기 위해 필요성이 점차 증대되고 있다. 현대 전투기의 조종석은 디지털 기술의 발달로 인하여 항공전자 장비의 기능이 다양하고 복잡해지고 있으며, 전투기를 조종하여 공격 임무를 수행해야 하는 조종사에게 항공전자 장비의 운용으로 인한 임무 부하량이 증대되기 마련이다. 따라서 음성인식 기술을 이용하여 항공전자 장비를 운용하게 되면, 조종사는 공격 임무에 더 많은 시간과 노력을 할애할 수 있게 된다. 본 연구는 전투기 조종석에 적용 가능한 음성명령 시스템을 개발하고, 검증환경을 구축하여 음성명령 시스템의 기능 및 성능을 검증한 것이다.

Key Words : Speech Recognition(음성 인식), Pattern Matching(패턴 매칭), DVI(Direct Voice Input), Command Syntax(음성명령어), Avionics(항공전자)

1. 서 론

1.1 전투기 음성명령 기술의 개요

HMI(Human Machine Interface)는 현대 전투

기의 기능적인 측면에서 중요한 역할을 담당하고 있다. 항공기 조종석 설계자들은 조종사와 항공기 간의 인터페이스를 향상시키기 위하여 지속적인 연구를 수행하고 있으며, 특히 조종사가 조종

† Received: September 3, 2013 Accepted: November 19 2013

* Corresponding author, E-mail : kim.seongwoo@lignex1.com

<http://journal.ksas.or.kr/>

pISSN 1225-1348 / eISSN 2287-6871

석 외부의 여러 상황에 집중할 수 있도록 임무수행 중 조종석과 조종사를 연결해 주는 인터페이스 수단은 상당히 중요한 요소이다(5).

현대 항공기 조종석은 디지털 기술의 발달로 인하여 항공전자 장비의 기능이 다양하고 복잡해지고 있는 추세이며, 공격 임무를 수행해야 하는 전투기 조종사에게는 항공전자 장비의 운용으로 인하여 조종석 내부에서의 임무부하가 증대되기 마련이다. 따라서 현대 전투기의 경우 음성명령 시스템을 채택하여 항공전자 장비를 운용하게 되면, 조종사는 공격 임무에 집중할 수 있게 되어 임무 부하량 감소의 효과가 있으며, 이로 인하여 전투 비행의 정확성 및 안정성 증대, 상황 판단력 향상, 임무수행 능력 향상에 도움을 줄 수 있다.

항공우주 산업의 선진국에서는 항공기상의 음성인식 제어 기술이 활발히 연구 중에 있으며, 최신 전투기에 해당하는 유럽의 Eurofighter Typhoon에는 이미 음성명령 시스템을 채택하여 운용되고 있으며, 미국의 F-35 Joint Strike Fighter에도 음성명령 시스템이 적용되어 있다.

1.2 전투기 음성명령 기술의 발전과정

전투기의 음성명령 기술 적용은 1980년대부터 시작되었다. 1985년, F-16에서 시도되었던 화자 종속 음성인식의 경우에는 하나의 큰 케이스에 여러 종류의 모듈을 장착하여 수행하였다. 1992년, AV-8B Harrier에서는 임무컴퓨터에 음성인식 모듈을 장착하는 방식이었다.

1999년, Eurofighter Typhoon에서 구현된 음성인식 모듈(SRM)은 인터콤에 해당하는 CAMU (Computing Devices Communication and Audio Management Unit)에 장착되며, 조종사의 voice template을 이용하여 음성인식을 수행한다. 총 100여 단어를 조합하여 음성명령이 구성되며, 특정 조종사의 음성명령만 인식하는 화자종속 시스템이다.

II. 본 론

2.1 개발 목표

현재 운용 중이거나 개발 중인 전투기에 적용되는 음성명령 기술은 적용 화자 기준으로 화자 종속 시스템에 해당한다. 일반적으로 다수의 사용자에게 의하여 운영되는 시스템의 경우에는 화자 독립 시스템으로 구성되어야 하지만, 전투기의 특수성으로 인하여 다수의 사용자를 위한 대중성

보다는 특정 조종사 집단을 위한 정확성이 우선되어야 하기 때문이다.

발음형태로 분류하면, 고립 단어 인식과 연결 단어 인식에 모두 포함되며, 인식 어휘 수에 따라 분류하면 일반적으로 음성명령어에 해당되는 어휘의 수가 수백 단어 이하로 구성이 되므로, 소용량 시스템에 해당된다.

음성인식 기술은 분류 기준에 따라 여러 종류로 나뉜다. 분류된 각 기술은 고유의 장단점을 갖고 있기 때문에 어느 하나가 다른 것보다 우월하다고 할 수 없다. 따라서 음성인식 시스템을 구현하려고 할 때는 적용 분야의 특성과 구현될 시스템의 경제성을 고려하여 적절한 방식을 선택해야 한다.

본 연구는 화자 독립 및 연결 단어 형태의 200 단어 규모에 대하여 인식률 95% 이상의 음성명령 시스템을 개발하여 기능 및 성능을 검증하는 것이다.

2.2 전투기 음성명령 시스템

본 연구에서 개발한 음성명령 시스템은 Fig. 1 과 같이 음성명령 인식 컴퓨터(하드웨어, 운용 프로그램, 음성인식 알고리즘)와 검증 환경(인식 결과 전시용 GUI, 음성명령 입력용 조종사 헤드셋, 음성명령 시작의 기준으로 사용되는 PTT (Push To Talk) 스위치)으로 구성된다.

2.3 음성명령 인식 컴퓨터 하드웨어

음성명령 인식 컴퓨터 하드웨어에는 PowerPC 8640D 기반의 SBC(Single Board Computer)가 사용되며, 운용 프로그램과 음성인식 알고리즘이 SBC에서 수행된다.

아날로그 신호를 디지털 신호로 변환하기 위하여 별도의 IO(Input/Output) 처리 모듈이 사용되며, 조종사에게 sidetone을 제공하기 위하여 AD(Analog to Digital) 변환된 데이터를 그대로 DA(Digital to Analog) 변환하는 기능이 포함된다. 음성신호를 처리하기 위해 필요한 샘플링 주



Fig. 1. Architecture of Voice Command System

과수는 6~16kHz, resolution은 최소 10bit이며(4), 본 연구에서는 획득한 조종석 환경의 음성명령을 분석한 후, 음성신호를 샘플링 주파수 10kHz, resolution 16bit의 디지털 신호로 변환한다.

2.4 음성명령 인식 컴퓨터 운용 프로그램

음성명령 인식 컴퓨터 운용 프로그램은 음성명령 인식 컴퓨터의 SBC에 탑재되어 동작한다. 기능으로는 음성명령 인식 알고리즘의 수행, 음성명령 인식 컴퓨터의 자체점검, 검증환경 운용 컴퓨터와의 연동 및 인터페이스 점검이 있다.

음성명령 인식 컴퓨터 운용 프로그램은 실시간으로 단위 작업들을 반복하여 처리하는 방식의 일반적인 비행운용프로그램의 구조가 아닌 운용자로부터 인식 요청을 받았을 때만 해당 작업을 처리하고 나머지 시간은 자체 점검에 할애하는 이벤트 방식(Event-driven)으로 설계 되었다. 왜냐하면 음성신호 처리에 소요되는 연산시간의 지터(Jitter)가 크기 때문이다(7).

음성명령의 수신 및 알고리즘의 수행, 검증환경 운용 컴퓨터로부터의 요청명령에 대해서는 인터럽트를 수신 대기하는 태스크(Task)를 두고, 음성명령 인식 컴퓨터의 자체점검은 유휴 시간(Idle-time) 동안 지속적으로 수행한다.

이벤트 방식으로 설계된 음성명령 인식 컴퓨터 운용 프로그램으로 인해 음성명령 인식 컴퓨터의 연산 자원을 효율적으로 이용할 수 있다.

2.5 음성인식 알고리즘

2.5.1 음성명령어

전투기 음성명령 시스템을 통하여 조종사는 조종석 내의 시스템 중에서 비행 및 안전 필수 시스템을 제외한 다음과 같은 시스템을 통제할 수 있다. 데이터 입력(통신 주파수, HaveQuick 데이터, IFF 모드 선택, IFF 코드 변경, TACAN 채널, Steerpoint 선택 및 편집, Autopilot 통제, Warning Level 설정 등), MFD 페이지 전환, 레이다 모드 변환, 생존 장비 통제 등이 해당된다.

본 연구에서는 F-15 및 F-16 운영 개념을 분석하여 총 194개의 단어를 추출하여 Fig. 2와 같은 음성명령어를 구성하였다.

음성명령어는 크게 시스템 모드, 통신, 항법, 무장, 생존, 환경제어 등과 관련된 명령어로 분류되며, 공통으로 사용되는 단어인 숫자, 취소 명령 등이 포함된다.

음성명령 시스템을 구현하기 위해서는 음향 모델의 훈련에 사용되는 음성 데이터베이스가 있어야 하며, 좋은 성능의 인식기를 만들기 위해서

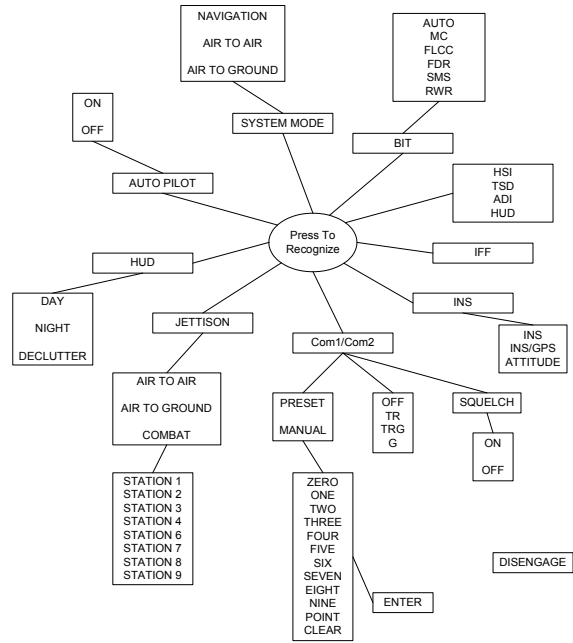


Fig. 2. Command Syntax

는 다양한 화자의 발음이 필요하며 훈련 데이터가 많을수록 높은 인식률을 기대할 수 있다.

본 연구에서는 화자 30명이 194 단어를 5회씩 발화하여 녹음한 후, 상용 사운드 편집기를 이용해서 각 단어 별로 세그먼트이션 하였으며, 녹음시 발화가 잘못 되었거나, 화자의 실수에 의해 누락된 단어 189개를 제외한 총 28,911개의 단어의 음성 데이터베이스를 구축하였다.

단어의 녹음은 테스트 환경과 동일한 환경에서 실시되기 때문에, 지상 모의 시스템 환경에서 헤드셋을 이용하여 녹음을 수행하였다.

2.5.2 연결 단어 인식

연결단어 인식 과정은 Fig. 3과 같이 크게 단어 단위 자동 분할, 특징 추출과 고립단어 인식으로 나뉜다.

기본적으로 입력음성에 단어 단위 자동분할을 수행하여 각 단어의 시작과 끝을 찾고, 각 단어 구간마다 특징추출을 진행하여 추출된 특징을 사용하여 고립단어 인식을 진행하게 된다.

고립단어 인식의 인식결과가 연결되어 명령어

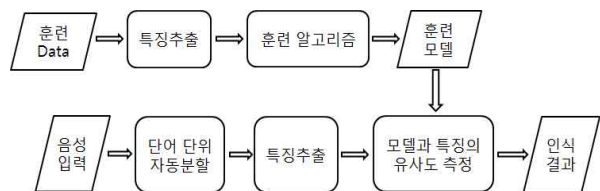


Fig. 3. Flow of Connected Word Recognition

인식결과가 도출된다.

단어단위 자동 분할을 위해서는 음성 에너지와 영 교차율(Zero Crossing Rate, ZCR)을 이용한 묵음구간 추출을 사용하였다.

특징 추출에는 음성인식에서 가장 널리 쓰이고 있는 MFCC (Mel-Frequency Cepstrum Coefficient) 특징을 사용하였으며, 음성인식에 탁월한 성능을 나타내는 HMM(Hidden Markov Model) 알고리즘을 사용하여 고립단어 음성인식 시스템을 개발하였다.

2.5.3 음성인식 후처리

음성명령어는 여러 단어의 조합으로 구성된다. 따라서 음성명령어 인식 시, 발화된 단어 중 한 단어라도 오인식되면 음성명령어의 인식 결과 또한 오인식으로 처리된다. 음성명령어는 일정한 규칙에 따라 단어를 조합하여 구성되므로, 명령어의 구문 정보를 사용하면 오인식된 소수의 단어를 정확한 단어로 교정할 수 있다. 즉, 음성명령어의 구문정보를 사용함으로써, 연산 자원을 절약하면서 효과적이고 경제적으로 인식률을 높일 수 있다. 구문정보를 사용하기 위해서는 음성명령어의 단어 조합 규칙을 수학적으로 표현해야 한다. 본 연구에서는 구문 트리를 사용하여 음성명령어의 단어 조합 규칙을 표현하였다.

2.5.3.1 명령어 구문 트리

명령어의 구문 트리는 음성명령어를 표현하는 효과적인 방법이다. 명령어의 규칙이 주어지면 구문트리는 Table 1 및 Fig. 4와 같이 명령어의 조합으로 구성할 수 있다.

Table 1. Voice Command

명령어 1	JETTISON	AIR TO AIR	STATION 1
명령어 2	JETTISON	AIR TO AIR	STATION 2
명령어 3	JETTISON	AIR TO GROUND	STATION 1
명령어 4	SYSTEM MODE	NAVIGATION	
명령어 5	TACAN	CHANNEL	NUM X

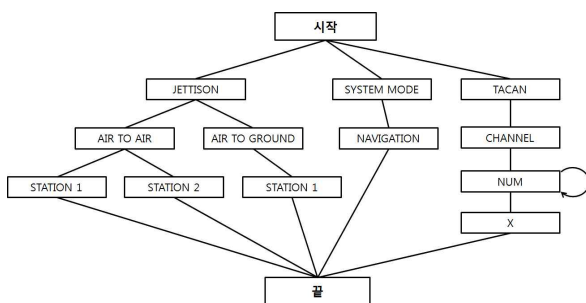


Fig. 4. Grammar Tree of Voice Command

우선 명령어의 시작이 루트 노드가 되고 명령어의 끝이 단말 노드가 된다. 루트 노드와 단말 노드 사이에는 각 단어를 대표하는 노드들이 생성된다.

음성 명령어의 각 단어들은 자신의 바로 앞이나오는 단어의 자식 노드가 되어 트리가 형성된다(2)(3). 명령어의 가장 앞에 있는 단어는 루트 노드의 자식 노드가 되며 명령어의 가장 뒤에 있는 단어는 명령어의 끝을 대표하는 단말 노드의 부모 노드가 된다.

위에 서술된 규칙을 사용하여 Table 1에서 표현되어있는 음성명령어 규칙을 Fig. 4와 같은 구문 트리로 표현할 수 있다.

Figure 4와 같은 구문트리를 사용하면 일련의 단어가 명령어 구문에 부합하는지 판단할 수 있다. 단어열이 루트 노드로부터 시작하여 단말 노드까지 도착하는 경로가 존재할 시 단어열은 명령어 규칙에 부합하게 된다. 구문트리 상의 경로가 없는 단어열은 명령어 규칙에 어긋나므로 오인식 단어의 검출 및 수정을 수행해야 한다.

Figure 4에서 “NUM”은 숫자를 뜻한다. 음성명령어에서 주파수 변환 등 숫자를 사용하는 기능에 관련된 명령어에는 여러 자리의 숫자가 명령어 중간에 삽입된다. 여러 자리의 숫자가 포함된 구문을 표현하기 위해 숫자를 대표하는 노드는 self loop를 구성하게 된다.

본 연구에서 설계한 음성명령어에 사용되는 단어는 200개 미만이며, 전투기의 특성상 간단한 구문을 가지고 있다. 따라서 Fig. 4와 같은 복잡하지 않은 구문 트리를 사용하여 모든 음성명령어를 표현할 수 있다.

2.5.3.2 구문 트리를 사용한 단어 정합

고립단어 인식의 인식결과가 구문에 부합하지 않으면, 구문 트리를 사용하여 오인식된 단어를 검출하고, 명령어의 구성단어를 수정한다.

고립 단어 인식을 진행할 시 각 단어에 대해서 유사도가 높은 순으로 정렬하여 상위에 랭크된 단어를 인식결과 후보로 선정한다. 후보의 개수는 반복된 실험을 통해서 결정한다. 본 연구에서는 각 단어마다 15개의 후보를 선정하였다. 각 후보의 유사도는 단어열의 유사도를 계산할 때 사용된다. 따라서 후보를 선정할 시 각 후보의 유사도도 같이 저장한다. 모든 단어의 후보가 선정된 후, 후보를 사용하여 조합이 가능한 모든 단어열 중, 구문 트리를 통과하면서 유사도가 가장 높은 단어열이 명령어 인식 결과가 된다. 여기서 단어열의 유사도는 각 단어 유사도의 합이

Table 2. Recognition Result of Isolated Word

단어 1		단어 2		단어 3	
후보	유사도	후보	유사도	후보	유사도
JETTISON	-100	EIGHT	-95	STATION 2	-101
SYSTEM MODE	-105	AIR TO AIR	-100	TACAN	-108
SEVEN	-120	NAVIGATION	-101	TWO	-110

Table 3. Flow of Postprocessing

“단어1”의 가능한 조합		“단어1, 단어2”의 가능한 조합		“단어1, 단어2, 단어3”의 가능한 조합	
가능한 조합	유사도	가능한 조합	유사도	가능한 조합	유사도
JETTISON	-100	JETTISON	-200	JETTISON	-301
		AIR TO AIR		AIR TO AIR	
				STATION 2	
SYSTEM MODE	-105	SYSTEM MODE	-206		
		NAVIGATION			

다. Table 2에 고립 단어 인식의 예시를 나타내었다.

위 과정에서 구문 트리를 통과하는 단어열만 고려하기 때문에 구문 규칙에 어긋나는 오인식된 단어는 자동으로 수정된다.

각 단어의 후보로 구성된 모든 가능한 단어열 조합을 고려하게 되면 가능한 조합의 개수가 너무 많아 정합 속도에 영향을 줄 수 있다.

예를 들어, 7개의 단어로 조합된 명령어를 인식할 시, 각 단어마다 5개의 후보를 선정한다면, 모든 가능한 단어열의 개수는 $5^7=78,125$ 개가 된다. 이 모든 단어열이 구문에 부합하는지 판단해야 하고, 각 단어열의 유사도도 계산해야 하기 때문에 많은 계산이 필요하다.

본 연구에서는 효과적인 후처리를 진행하기 위해 단어 발화 순서대로 순차적 단어정합을 진행하는 방법을 사용하였다. Table 3에 후처리의 예시를 나타내었으며, 이 경우의 최종 인식 결과는 JETTISON - AIR TO AIR - STATION 2이다.

순차적 단어 정합 방법에서는, 우선 앞에 발화된 단어들의 조합만 고려하여 단어열 조합을 만들고 이 단어열 중 명령어 구문 트리에 부합할 수 없는 조합은 제거하게 된다. 그리고 뒤에 발화된 단어들에 대해 정합을 진행할 시 앞에서 부적합하다고 판단된 단어열 조합은 더 이상 고려하지 않게 된다. 뒷 단어에 대해 정합을 진행할 시, 앞에서 적합하다고 판단된 조합만이 고려됨으로써, 정합이 진행될수록 가능한 조합의 개수가 줄어들어 계산량을 효과적으로 줄일 수 있게 된다.

2.6 검증환경

음성명령 시스템을 검증하기 위하여 Fig. 5와



Fig. 5. Configuration of Verification Environment

같은 검증환경을 개발하였으며, 검증환경은 음성명령 입력용 조종사 헤드셋, 음성명령 시작 기준으로 사용되는 PTT(Push To Talk) 및 인식결과 전시용 GUI(Graphic User Interface)로 구성된다.

2.6.1 헤드셋 및 PTT

조종사 헤드셋은 음성명령 인식 컴퓨터의 음성 신호(Analog)를 입력하기 위한 장비이며 전투기 조종사의 Oxygen Mask를 모의하기 위해 Ground Crew용 헤드셋을 사용하였다.

PTT 스위치는 Saitek X52 Pro HOTAS를 사용하였으며, 내부를 개조하여 제품에 있는 USB 통신 대신 PTT 신호(Discrete)를 음성명령 인식 컴퓨터로 전송한다.

음성명령 인식 컴퓨터는 PTT 스위치가 눌러진 동안에 입력된 음성 데이터를 디지털화한 후 음성인식 알고리즘을 통해 인식하게 된다. 인식 결과는 검증환경의 운용컴퓨터로 전송되어 인식결과 전시용 GUI에서 확인하게 된다.

2.6.2 인식결과 전시용 GUI

인식결과 전시용 GUI는 Fig. 6과 같이 구성하였으며, 운용컴퓨터(PC)에서 수행된다.

인식결과 전시를 위하여 MFD(Multi Function Display) 형태의 GUI를 중앙에 배치하고 각종 기능들을 주변에 위치시켰다.

인식 결과는 MFD 내부에 전시가 되며, 헤드셋을 통하여 발화한 명령어를 알고리즘에서 인식하게 되면 정인식 여부에 관계없이 MFD 좌측 하단에 인식 결과를 전시하도록 구현하였다. 더불어 정인식의 경우에는 해당 페이지로 전환이 이루어지고 현재 명령어의 색깔이 변경되어 하이 라이트로 표시된다. 반면에 오인식의 경우에는 좌측 하단에 오인식 결과만 전시한다.

음성명령 인식 컴퓨터는 화자가 마지막으로

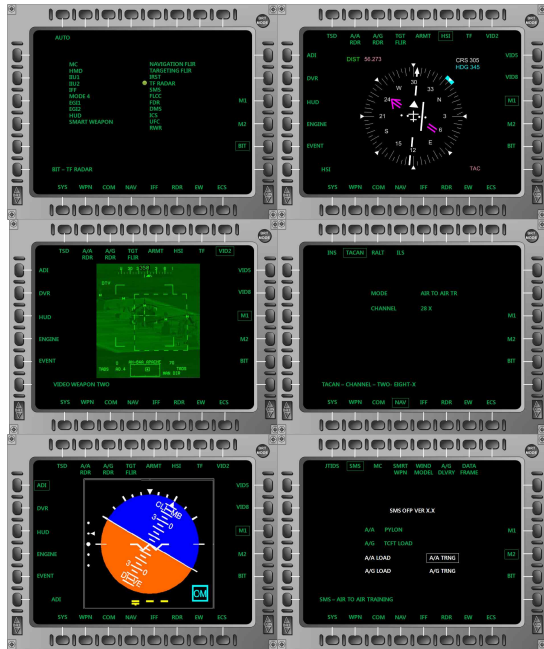
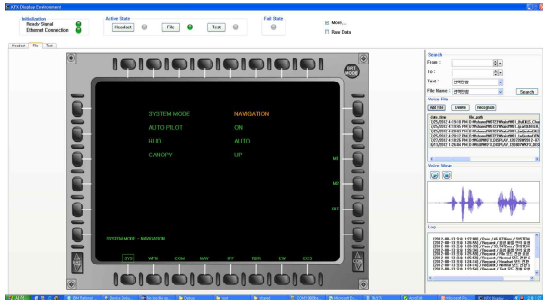


Fig. 6. Graphic User Interface of Recognition Result

발화한 명령어를 음성명령 인식 컴퓨터 내부에 임시 저장하는 기능을 포함한다.

GUI에서는 그 명령어의 전송을 요청하면 음성 데이터가 전시용 PC로 전송되어 그것을 파일로 저장하는 기능을 갖고 있다. 더불어 파일 형태의 명령어를 음성명령 인식 컴퓨터로 전송하여 인식을 요청하고 인식 결과를 재전송 받아서 MFD에 전시하는 기능도 구현하였다. 알고리즘 튜닝에 유용한 기능이다.

또한, 음성명령 인식 컴퓨터로부터 주기적으로 Health 신호(RS-422)를 수신하여 화면에 전시하고 음성데이터 및 결과를 송수신하여 전시하도록 설계하였다.

2.7. 시험절차 및 결과

2.7.1 음성명령어 Test Case

음성명령 시스템을 검증하기 위하여 음성명령어 Test Case를 설계하였다.

Table 4. Test Case of Voice Command

기능	명령어
시스템 모드 설정 - NAV 설정	SYSTEM MODE - NAVIGATION
Master ARM 설정 - SAFE	MASTER ARM - SAFE
U/VHF 기능 선택 - T/R	COM1 - TR
INS 모드 선택 - INS/GPS	INS - INS/GPS
TACAN 채널 설정 - 054 X	TACAN - CHANNEL - ZERO - FIVE - FOUR - X
RADAR ALTIMETER 설정 - STANDBY	RADAR ALTIMETER - STANDBY
ILS 채널 설정 - 165.7	ILS - CHANNEL - ONE - SIX - FIVE - POINT - SEVEN - ENTER
IFF 모드1 코드 변경 - 7890	IFF - MODE1 - CODE - SEVEN - EIGHT - NINE - ZERO - ENTER

음성명령어 중 숫자가 포함된 명령어의 경우에는 숫자 조합을 배제하였고, 총 194개의 단어로 구성 가능한 명령어는 최소 183개이다.

음성명령어 Test Case 중 일부를 Table 4에 나타내었으며, 명령어를 구성하는 단어의 개수는 명령어 별로 1개에서 수개까지가 될 수 있다.

2.7.2 시험절차

알고리즘 성능 시험은 Fig. 5의 검증환경에서 수행되었다. 음성명령 인식 컴퓨터에 전원을 인가하고, 운용컴퓨터에서 GUI 운용 프로그램을 실행한다. 그리고 PTT 스위치의 운용과 함께 음성명령어 Test Case를 발화한다.

음성명령어 Test Case를 이용하여 온라인 음성 명령어 인식 실험을 진행한 결과 명령어 인식률은 Table 5와 같이 산출되었다. 명령어의 모든 단어가 정확히 인식되었으면, 해당 명령어를 정인식으로 판단하였으며, 명령어의 단어 중 한 단어라도 잘못 인식되었을 경우 해당 명령어를 오인식으로 판단하였다.

실험실에서는 잡음이 없는 깨끗한 환경임을 감안하여 임의의 잡음을 추가한 후 음성 인식 단

Table 5. Test Result of Voice Command System

화자	Test Case	인식률	
		후처리 전	후처리 후
1	183개 명령어 x 3회	71.58%	96.90%
2	183개 명령어 x 2회	77.05%	94.81%
평균	183개 명령어 x 5회	73.77%	96.07%

계를 진행하였으며, 훈련 모델을 훈련 시에는 noise adaptive 기법을 적용하여 음성 DB에 동일한 잡음을 추가하고 잡음 제거 알고리즘을 수행한 후 훈련한 훈련 모델을 사용하였다.

2.7.3 시험결과

음성 DB에 포함된 화자 2명이 183개의 Test Case를 5회 발화한 경우 평균 95% 이상의 명령어 인식률을 보였다.

후처리를 진행한 결과, 명령어 인식률이 후처리를 진행하기 전보다 대폭 향상됨을 Table 5에서 확인할 수 있다.

또한, Frame length와 Frame shift를 동일하게 256 sample로 설정하여 Frame을 겹쳐서 추출하지 않았지만, Altitude 및 Attitude와 같은 유사한 단어들을 제외하고는 데이터의 손실로 인한 단어의 오인식은 없음을 확인하였다.

III. 결 론

본 논문에서는 외국의 전투기에 적용한 화자 종속형 방식이 아닌 화자 독립형 방식을 전투기에 적용하기 위한 음성명령 인식 컴퓨터 하드웨어, 운용 프로그램 및 음성인식 알고리즘을 개발하여 지상 모의 검증환경에서 음성명령 시스템의 기능 및 성능을 검증하였다.

적용한 음성인식 알고리즘에서는 음성명령어의 문법 규칙을 고려하여 명령어의 구문트리 정보를 연결 단어 인식에 활용하여 음성명령의 인식률을 향상시켰으며, 음성 신호의 처리 연산량을 감소시키기 위하여 인식된 단어에 순차적 단어 정합 방법을 사용하였다.

음성명령 시스템을 검증하기 위하여 지상 모의 검증환경을 개발하였으며, 음성명령어 Test Case를 이용하여 시스템의 기능 및 성능을 검증하였다.

일반적으로 음성인식 결과 95% 이상의 인식률은 정상적인 시스템 기능 수행이 가능한 수준이

며, 본 연구에서 제시한 시스템은 이것을 만족한다.

향후 실제 전투기에 적용하기 위해서는 적용 전투기의 운용 개념 분석을 통한 음성명령어 도출 및 음성 데이터베이스 구축이 수행되어야 하며, 이기종 전투기에서의 인식률 향상을 위한 알고리즘 튜닝 및 보완이 필요하다.

References

- 1) Lawrence Rabiner, Biing-Hwang Juang, "Fundamentals of speech recognition," Prentice Hall, 1993
- 2) Yves Schabes, Richard C. Waters, "Tree insertion grammar: a cubic-time, parsable formalism that lexicalizes context-free grammar without changing the trees produced", Journal of Computational Linguistics, Vol. 21, No. 4, 1995, pp.479-513
- 3) K. Vijay-Shanker, "Using descriptions of trees in a tree adjoining grammar", Journal of Computational Linguistics, Vol. 18, No. 4, 1992, pp.481-517
- 4) Y. Oh, "Speech Language Information Processing", Hongrung Publishing Company, 1998
- 5) H. Kim, 2003, "The Status and Forecast of Speech Information Processing", The Magazine of the IEEK Vol. 30, No. 7
- 6) S. Kim, C. Shin, B. Kim, "A Study on Fighter Airplane's Voice Command Recognition System Design and Verification Environment", KSAS Fall Conference, 2012, pp.327-331
- 7) M. Seo, C. Shin, B. Kim, "The Structure of Operational Flight Program for voice command recognition based on the event-driven paradigm", KSAS Fall Conference, 2012, pp.1755-1760