



(19)대한민국특허청(KR)  
(12) 등록특허공보(B1)

(51) 。 Int. Cl. G10L 15/06 (2006.01)	(45) 공고일자 (11) 등록번호 (24) 등록일자	2007년01월12일 10-0668297 2007년01월08일
---	-------------------------------------	--

(21) 출원번호 (22) 출원일자 심사청구일자	10-2002-0087943 2002년12월31일 2003년12월31일	(65) 공개번호 (43) 공개일자	10-2004-0061659 2004년07월07일
----------------------------------	---	------------------------	--------------------------------

(73) 특허권자	삼성전자주식회사 경기도 수원시 영통구 매탄동 416
(72) 발명자	정승녕 서울특별시양천구목6동신시가지아파트206동101호  유명현 경기도성남시분당구장안동81-9  김재우 경기도용인시기홍읍영덕리15번지신일아파트109동2006호  박준아 서울특별시마포구신공덕동삼성아파트108동1603호
(74) 대리인	리앤목특허법인 이해영

심사관 : 경연정

전체 청구항 수 : 총 26 항

(54) 음성인식방법 및 장치

(57) 요약

음성인식결과 복수개의 대안으로 이루어진 목록에 대한 사용자의 선택여부에 따라 최종 인식단어를 결정하는 프로세스를 적응적으로 변경시킴으로써 음성인식에 따른 작업 성공율을 향상시킬 수 있는 음성인식방법 및 장치가 개시된다. 음성인식방법은 (a) 사용자가 발화한 음성을 입력하는 단계, (b) 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 단계, 및 (c) 상기 소정 갯수의 인식단어가 소정의 순서에 따라서 나열된 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 단계를 포함한다.

대표도

도 3

## 특허청구의 범위

### 청구항 1.

- (a) 사용자가 발화한 음성을 입력하는 단계;
- (b) 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 단계;
- (c) 상기 소정 갯수의 인식단어가 소정의 순서에 따라서 나열된 대안목록을 디스플레이시키는 단계;
- (d) 상기 디스플레이된 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 있는지를 판단하는 단계;
- (e) 상기 대기시간동안 상기 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 단계; 및
- (f) 상기 사용자의 숙련도에 따라 상기 대기시간을 재조정하는 단계를 포함하는 음성인식방법.

### 청구항 2.

제1 항에 있어서, 상기 방법은

- (g) 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 상기 대기시간을 재조정하는 단계를 더 포함하는 음성인식방법.

### 청구항 3.

제1 항에 있어서, 상기 방법은

- (g) 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 단계를 더 포함하는 음성인식방법.

### 청구항 4.

제1 항 내지 제3 항 중 어느 한 항에 있어서, 상기 (c) 단계는

- (c1) 상기 음성인식결과에 따른 첫 번째대안 인식단어와 최종 인식단어의 조합을 이용하여 오류패턴을 업데이트시키는 단계; 및
- (c2) 상기 오류패턴을 이용하여 상기 대안목록의 나열순서를 재조정하는 단계를 더 포함하는 음성인식방법.

### 청구항 5.

제4 항에 있어서, 상기 (c1) 단계는

- (c11) 상기 음성인식결과에 따른 첫 번째대안 인식단어와 최종 인식단어의 조합이 소정의 오류패턴 데이터베이스에 포함되어 있는 경우 적어도 하나의 발화특징 차이값을 산출하는 단계;

(c12) 상기 (c11) 단계에서 얻어진 발화특징 차이값과 소정의 제1 문턱치를 비교하는 단계; 및

(c13) 상기 (c12) 단계에서의 비교결과, 상기 발화특징 차이값이 제1 문턱치보다 작은 경우 현재 입력된 음성의 발화특징을 포함하여 각 발화특징의 평균치를 산출하여 발화성향을 업데이트시키고, 해당 패턴의 히스토리의 값을 1 증가시켜 히스토리를 업데이트시키는 단계를 포함하는 음성인식방법.

## 청구항 6.

제4 항에 있어서, 상기 (c2) 단계는

(c21) 상기 음성인식결과에 따른 첫 번째대안 단어와 두 번째대안 단어의 조합과 첫 번째대안 단어와 세 번째대안 단어의 조합 중 적어도 하나가 소정의 오류패턴 데이터베이스에 포함되어 있는 경우, 각 조합에 대하여 적어도 하나의 발화특징 차이값을 산출하는 단계;

(c22) 상기 (c21) 단계에서 얻어진 각 조합의 발화특징 차이값을 소정의 제2 문턱치와 비교하는 단계; 및

(c23) 상기 (c22) 단계에서의 비교결과, 각 조합의 상기 발화특징 차이값이 제2 문턱치보다 작은 경우, 해당 조합에 포함된 대안 단어의 스코어를 수정하는 단계를 포함하는 음성인식방법.

## 청구항 7.

제6 항에 있어서, 상기 수정된 스코어는 원래의 스코어에 소정의 가중치와 해당 패턴의 히스토리를 곱한 값을 더함으로써 산출되는 음성인식방법.

## 청구항 8.

삭제

## 청구항 9.

제1 항에 있어서, 상기 (f) 단계는

(f1) 소정의 숙련도 데이터베이스에 저장된 소정 선택시간에서 현재의 최종인식단어 결정에 소요된 시간을 뺀 값으로 선택시간 차이값을 산출하는 단계;

(f2) 상기 (f1) 단계에서 얻어진 선택시간 차이값을 소정의 제3 문턱치와 비교하는 단계;

(f3) 상기 (f2) 단계에서의 비교결과, 선택시간 차이값이 제3 문턱치보다 큰 경우 선택시간을 수정하는 단계;

(f4) 상기 (f2) 단계에서의 비교결과, 선택시간 차이값이 제3 문턱치보다 작거나 같은 경우 선택시간 차이값을 소정의 여유시간과 비교하는 단계;

(f5) 상기 (f4) 단계에서의 비교결과, 선택시간 차이값이 여유시간보다 작은 경우 선택시간을 수정하는 단계; 및

(f6) 상기 (f3) 단계 또는 상기 (f5) 단계에서 수정된 선택시간에 소정의 여유시간을 가산함으로써 사용자의 대기시간을 산출하는 단계를 포함하는 음성인식방법.

## 청구항 10.

제9 항에 있어서, 상기 (f3) 단계에서는 상기 숙련도 데이터베이스에 저장되어 있던 선택시간에서 선택시간 차이값에 소정의 가중치를 곱한 값을 감산함으로써 수정된 선택시간을 산출하는 음성인식방법.

### 청구항 11.

제9 항에 있어서, 상기 (f5) 단계에서는 상기 숙련도 데이터베이스에 저장되어 있던 선택시간에 소정의 부가시간을 가산함으로써 수정된 선택시간을 산출하는 음성인식방법.

### 청구항 12.

제1 항 내지 제3 항 중 어느 한 항에 있어서, 상기 대기시간은 상기 대안목록에 나열된 모든 대안단어에 대하여 동일한 값으로 부여되는 음성인식방법.

### 청구항 13.

제1 항 내지 제3 항 중 어느 한 항에 있어서, 상기 대기시간은 상기 대안목록에 나열된 대안단어의 순서에 따라서 차등적인 값으로 부여되는 음성인식방법.

### 청구항 14.

사용자가 발화한 음성을 입력하고, 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하고, 상기 소정 갯수의 인식단어가 소정의 순서에 따라서 나열된 대안목록을 디스플레이하고, 상기 음성인식결과에 따른 첫 번째 대안 인식단어와 최종 인식단어의 조합을 이용하여 오류패턴을 업데이트시키고, 상기 오류패턴을 이용하여 상기 대안목록의 나열순서를 재조정하는 제1 프로그램; 및

소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 있는지를 판단하고, 상기 소정의 대기시간 동안 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 제2 프로그램을 기록한 컴퓨터가 읽을 수 있는 기록매체.

### 청구항 15.

제14 항에 있어서, 상기 제2 프로그램은 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우 상기 대기시간을 재조정하는 다음, 재조정된 대기시간동안 재차 사용자의 선택변경이 있는지를 판단하여 사용자의 선택변경이 없는 경우 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 기능을 더 구비하는 컴퓨터가 읽을 수 있는 기록매체.

### 청구항 16.

제14 항에 있어서, 상기 제2 프로그램은 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 기능을 더 구비하는 컴퓨터가 읽을 수 있는 기록매체.

### 청구항 17.

사용자가 발화한 음성을 입력하는 음성입력부;

상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 음성인식부; 및

상기 소정 갯수의 인식단어가 소정의 순서에 따라서 나열된 대안목록을 디스플레이시키고, 상기 디스플레이된 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 있는지를 판단하여 상기 대기시간동안 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하고, 상기 음성인식결과에 따른 첫 번째대안 인식단어와 최종 인식 단어의 조합을 이용하여 오류패턴을 업데이트시키고, 상기 업데이트된 오류패턴을 이용하여 상기 대안목록의 나열순서를 재조정하는 후처리부를 포함하는 음성인식장치.

## 청구항 18.

제17 항에 있어서, 상기 후처리부는

상기 음성인식부에서 생성된 소정 갯수의 인식단어를 소정의 순서에 따라서 나열된 대안목록이 포함된 그래픽 사용자 인터페이스용 윈도우를 생성하는 윈도우 생성부;

상기 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 현재 커서가 이동되어 있는 대안단어를 최종 인식단어로 결정하기까지의 대기시간을 설정하는 대기시간 설정부; 및

소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 첫번째 대안단어를 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 상기 대기시간을 재조정하여 재조정된 대기시간동안 재차 사용자의 선택변경이 없는 경우 사용자가 선택한 대안단어를 최종 인식 단어로 결정하는 최종 인식단어 결정부를 구비하는 음성인식장치.

## 청구항 19.

제17 항에 있어서, 상기 후처리부는

상기 음성인식부에서 생성된 소정 갯수의 인식단어를 소정의 순서에 따라서 나열된 대안목록이 포함된 그래픽 사용자 인터페이스용 윈도우를 생성하는 윈도우 생성부;

상기 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 현재 커서가 이동되어 있는 대안단어를 최종 인식단어로 결정하기까지의 대기시간을 설정하는 대기시간 설정부; 및

소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 첫번째 대안단어를 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 최종 인식단어 결정부를 구비하는 음성인식장치.

## 청구항 20.

제18 항 또는 제19 항에 있어서, 상기 후처리부는

상기 음성인식부에서 첫번째 대안으로 채택되었던 인식단어, 상기 최종 인식단어 결정부로부터 제공되는 최종인식단어, 적어도 하나 이상의 사용자 발화특징, 발화성향, 및 히스토리를 데이터베이스화한 오류패턴 데이터베이스; 및

상기 음성인식부로부터 제공되는 음성인식결과 및 스코어를 입력으로 하고, 상기 오류패턴 데이터베이스를 참조하여 상기 오류패턴에 해당하는 인식단어의 스코어를 조정하고, 상기 대안목록의 나열순서를 변경하여 상기 윈도우 생성부로 제공하는 오류패턴 관리부를 더 포함하는 음성인식장치.

**청구항 21.**

제18 항 또는 제19 항에 있어서, 상기 후처리부는

속련도에 따라서 서로 다른 선택시간을 대응시켜 데이터베이스화한 속련도 데이터베이스; 및

상기 속련도 데이터베이스에 저장된 속련도에 따른 선택시간에 소정의 여유시간을 부가한 값을 대기시간으로 재조정하여 상기 대기시간 설정부로 공급하는 속련도 관리부를 더 포함하는 음성인식장치.

**청구항 22.**

제20 항에 있어서, 상기 후처리부는

속련도에 따라서 서로 다른 선택시간을 대응시켜 데이터베이스화한 속련도 데이터베이스; 및

상기 속련도 데이터베이스에 저장된 속련도에 따른 선택시간에 소정의 여유시간을 부가한 값을 대기시간으로 재조정하여 상기 대기시간 설정부로 공급하는 속련도 관리부를 더 포함하는 음성인식장치.

**청구항 23.**

제18 항 또는 제19 항에 있어서, 상기 대기시간은 사용자의 속련도에 따라 결정되는 음성인식장치.

**청구항 24.**

제18 항에 있어서, 상기 재조정된 대기시간은 상기 대안목록에 나열된 모든 대안단어에 대하여 동일한 값으로 부여되는 음성인식장치.

**청구항 25.**

제18 항에 있어서, 상기 재조정된 대기시간은 상기 대안목록에 나열된 대안단어의 순서에 따라서 차등적인 값으로 부여되는 음성인식장치.

**청구항 26.**

- (a) 사용자가 발화한 음성을 입력하는 단계;
- (b) 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 단계;
- (c) 상기 음성인식결과에 따른 첫 번째대안 인식단어와 최종 인식단어의 조합을 이용하여 업데이트되는 오류패턴에 기초하여 조정되는 순서에 따라서, 상기 소정 갯수의 인식단어들이 나열된 대안목록을 디스플레이시키는 단계; 및
- (d) 상기 디스플레이된 대안목록에 대한 사용자의 선택에 따라서 최종 인식단어를 결정하는 단계를 포함하는 음성인식방법.

**청구항 27.**

- (a) 사용자가 발화한 음성을 입력하는 단계;
- (b) 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 단계;
- (c) 상기 음성인식결과에 따른 첫 번째대안 인식단어와 최종 인식단어의 조합을 이용하여 업데이트되는 오류패턴에 기초하여 조정되는 순서에 따라서, 상기 소정 갯수의 인식단어들이 나열된 대안목록을 디스플레이시키는 단계;
- (d) 상기 디스플레이된 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 있는지를 판단하는 단계; 및
- (e) 상기 대기시간동안 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 단계를 포함하는 음성인식방법.

명세서

**발명의 상세한 설명**

**발명의 목적**

**발명이 속하는 기술 및 그 분야의 종래기술**

본 발명은 음성인식에 관한 것으로서, 특히 음성인식결과 복수개의 대안으로 이루어진 목록에 대한 사용자의 선택여부에 따라 최종 인식단어를 결정하는 프로세스를 적응적으로 변경시킴으로써 음성인식에 따른 작업 성공율을 향상시킬 수 있는 음성인식방법 및 장치에 관한 것이다.

음성인식(Speech Recognition) 기술이란 인간의 음성을 컴퓨터가 분석해 이를 인식 또는 이해하는 기술을 말하는데, 발음에 따라 입 모양과 혀의 위치 변화로 특정한 주파수를 갖는 인간의 음성을 이용, 발생된 음성을 전기신호로 변환한 후 음성신호의 주파수 특성을 추출해 발음을 인식하는 기술이다. 최근에는 이와 같은 음성인식 기술이 전화 다이얼링, 장난감 제어, 어학학습 또는 가전기기 제어 등과 같은 다양한 분야에 응용되고 있다.

그런데, 음성인식 기술이 아무리 발전하였다 하더라도 실제의 음성인식 환경이 주변 소음 등으로 매우 열악하여 현재의 기술로는 아직 100 % 인식성공율을 보장할 수 없기 때문에 음성인식에 의해 수행되는 작업의 오류가 빈번히 발생하는 경향이 있다. 이러한 작업의 오류 발생율을 줄이기 위한 방편으로서, 음성인식기의 인식결과에 대하여 사용자에게 확인을 요구하거나, 음성인식기의 인식결과에 따라서 복수개의 대안으로 이루어지는 목록을 사용자에게 제시함으로써, 사용자의 확인 또는 선택에 따라 최종 인식단어를 결정하는 방법들이 사용되고 있다.

이와 같은 방법들과 관련된 종래 기술로는 USP 4866778호, USP 5027406호, USP 5884258호, USP 6314397호, USP 6347296호 등이 있다. USP 4866778호는 가장 효율적으로 검색된 유력한 대안을 제시하고, 유력한 대안이 틀린 경우 다음 대안을 제시하는 방식에 의해 올바른 인식결과를 찾는 기술이다. 이 기술에 따르면 사용자는 시스템에 의해 제시되는 일련의 예/아니오 질문에 일일이 응답해야 하며, 항상 다음 질문에서 어떤 단어가 나올지 모르는 상태가 유지되기 때문에 비효율적이다. USP 5027406호와 USP 5884258호는 음성인식 결과에 대한 대안을 나열하고, 나열된 대안에 대하여 그래픽 사용자 인터페이스(Graphic User Interface) 또는 음성에 의한 사용자의 선택에 따라서 인식결과를 결정하는 기술이다. 이 기술에 따르면, 사용자는 발화 이후에 어떠한 경우라도 올바른 대안을 선택하는 별도의 조작을 취해야 하기 때문에 반복적인 작업에 의해 쉽게 불편함과 피로를 느끼게 된다. USP 6314397호는 최선의 인식결과를 바탕으로 사용자 발화를 일단 문자로 변환하고, 변환된 문자를 사용자가 검토하는 과정에서 이전에 고려되었던 인식결과에 대한 대안을 조회 및 선택하여 수정하는 기술이다. 이 기술은 원활한 작업 방식을 제안하고 있으나, 사용자가 실시간으로 이러한 시스템을 사용할 경우 인식결과를 눈으로 보면서 동시에 문장을 생성한다는 점을 감안하면, 잘못된 인식결과를 일단 그대로 입력하는 것은 불합리하다. USP 6347296호는 일련의 음성인식 작업 중 특정 발화에 대한 불명확한 인식결과를 확정하는데 있어서, 그 이후의 발화에 대한 인식결과를 참조하여 대안들 중에서 자동적으로 선택하는 기술이다. 이 기술에 따르면, 이후의 발화에 대한 인식 결과가 불명확하거나 완벽한 언어모델이 구비되어 있지 않은 실제 상황에서는 오류를 연발할 가능성이 크다.

상기한 바와 같이 종래에는 사용자의 음성에 대하여 정확한 인식결과가 도출되더라도 적어도 1 회의 사용자 확인 또는 선택과 같은 부가작업이 수행되어야 하며, 확인작업이 행해지지 않을 경우에는 최종 인식단어를 결정하는데 소요되는 시간이 무제한적으로 길어지게 되는 단점이 있다.

**발명이 이루고자 하는 기술적 과제**

따라서 본 발명이 이루고자 하는 기술적 과제는 음성인식결과 복수개의 대안으로 이루어진 목록에 대하여 사용자의 선택이 없는 경우 일정한 대기시간 경과 후 첫번째 대안단어를 최종 인식단어로 결정하고, 사용자의 선택이 있는 경우 선택된 대안단어를 최종 인식단어로 결정하거나, 재조정된 대기시간 경과 후 선택된 대안단어를 최종 인식단어로 결정하는 음성인식방법을 제공하는데 있다.

본 발명이 이루고자 하는 다른 기술적 과제는 상기 음성인식방법을 실현하는데 가장 적합한 장치를 제공하는데 있다.

상기 기술적 과제를 달성하기 위하여 본 발명에 따른 음성인식방법은 (a) 사용자가 발화한 음성을 입력하는 단계; (b) 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 단계; 및 (c) 상기 소정 갯수의 인식단어를 소정의 순서에 따라서 나열한 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 단계를 포함한다.

상기 음성인식방법은 (d) 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 상기 대기시간을 재조정하여 상기 (b) 단계로 복귀하는 단계를 더 포함하거나, (d) 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 단계를 더 포함하는 것이 바람직하다.

상기 다른 기술적 과제를 달성하기 위하여 본 발명에 따른 음성인식장치는 사용자가 발화한 음성을 입력하는 음성입력부; 상기 입력되는 음성을 인식하여 유사도 순서에 따라 소정 갯수의 인식단어를 생성하는 음성인식부; 및 상기 소정 갯수의 인식단어를 소정의 순서에 따라서 나열된 대안목록에 대하여 소정의 대기시간동안 사용자의 선택변경이 없는 경우 현재 커서가 위치한 대안단어를 최종 인식단어로 결정하는 후처리부를 포함한다.

또한, 상기 후처리부는 상기 음성인식부에서 생성된 소정 갯수의 인식단어를 소정의 순서에 따라서 나열된 대안목록이 포함된 그래픽 사용자 인터페이스용 윈도우를 생성하는 윈도우 생성부; 상기 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 현재 커서가 이동되어 있는 대안단어를 최종 인식단어로 결정하기까지의 대기시간을 설정하는 대기시간 설정부; 및 소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 첫번째 대안단어를 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 상기 대기시간을 재조정하여 재조정된 대기시간동안 재차 사용자의 선택변경이 없는 경우 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 최종 인식단어 결정부를 구비한다.

또한, 상기 후처리부는 상기 음성인식부에서 생성된 소정 갯수의 인식단어를 소정의 순서에 따라서 나열된 대안목록이 포함된 그래픽 사용자 인터페이스용 윈도우를 생성하는 윈도우 생성부; 상기 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 현재 커서가 이동되어 있는 대안단어를 최종 인식단어로 결정하기까지의 대기시간을 설정하는 대기시간 설정부; 및 소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 첫번째 대안단어를 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안단어를 최종 인식단어로 결정하는 최종 인식단어 결정부를 구비한다.

**발명의 구성**

이하, 본 발명의 실시예에 대하여 첨부된 도면들을 참조하여 상세하게 설명하기로 한다.

도 1은 본 발명의 일 실시예에 따른 음성인식장치의 구성을 나타낸 블럭도로서, 음성입력부(11), 음성인식부(13) 및 후처리부(15)로 이루어진다.

도 1을 참조하면, 음성입력부(11)는 마이크로폰 등으로 이루어지며, 사용자가 발성한 음성을 입력하고, 입력된 음성에 대하여 불필요한 잡음신호를 제거하고, 소정 레벨로 증폭시켜 음성인식부(13)로 제공한다.



음성인식부(13)는 음성입력부(11)로부터 입력되는 음성의 시작점 및 끝점을 검출하여 입력된 음성의 앞뒤에 있는 묵음구간을 제외한 음성구간에서의 특징 데이터를 추출한 후 실시간으로 벡터 양자화를 수행한다. 이후, 특징데이터를 이용하여 데이터베이스에 등록된 단어들에 대하여 유사도가 가장 유사한 단어들을 선정하는 비터비 탐색을 수행한다. 이때, 은닉 마르코프 모델(Hidden Markov Model, HMM)을 이용할 수 있는데, 현재 입력된 음성의 특징데이터와 인식대상 후보단어들에 대하여 미리 훈련하여 구축한 HMM의 특징데이터를 비교하고, 특징데이터들의 차이를 이용하여 가장 유사한 후보단어를 결정한다. 음성인식부(13)에서 비터비 검색이 종료된 후 인식결과로서, 유사도 순서에 따라서 입력음성과 가장 유사한 소정 갯수, 여기서는 3개의 단어를 결정하여 후처리부(15)로 제공한다.

후처리부(15)는 음성인식부(13)의 인식결과를 입력으로 하여 인식결과를 문자신호로 변환하고, 변환된 문자신호를 유사도 순서에 따라서 나열한 그래픽 사용자 인터페이스용 윈도우를 생성한다. 그래픽 사용자 인터페이스용 윈도우의 일례로는 도 9에 도시된 바와 같은 형태를 가지며, 윈도우(91)에는 "첫번째 대안, 여기서는 '탐색기'를 실행 중입니다"라는 메시지가 나타나는 영역(92), 타임 벨트(93)가 표시되는 영역(93), 및 대안 목록이 나타나는 영역(94)을 포함한다. 생성된 윈도우(91)는 미리 설정된 대기시간에 해당하는 타임 벨트(93)가 모두 소진될 때까지 화면상에 디스플레이되고, 대기시간 이내에 사용자로부터 별도의 키 또는 버튼입력이 없는 경우 첫번째 대안을 최종 인식단어로 결정한다. 한편, 대기시간 이내에 사용자로부터 대안 선택용 키 또는 버튼이 입력되면 후술하는 도 7 또는 도 8에 도시된 과정을 통하여 최종 인식단어를 결정한다.

도 2는 도 1에 있어서 후처리부(17)의 세부적인 구성을 나타낸 블럭도로서, 대기시간 설정부(21), 숙련도 관리부(22), 숙련도 데이터베이스(DB)(23), 윈도우 생성부(24), 오류패턴 관리부(25), 오류패턴 데이터베이스(DB)(26) 및 최종 인식단어 결정부(27)로 이루어진다.

도 2를 참조하면, 대기시간 설정부(21)는 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 현재 커서가 이동되어 있는 대안을 최종 인식단어로 결정하기까지의 대기시간을 설정한다. 대기시간은 그래픽 사용자 인터페이스용 윈도우(91)에서 타임 벨트(93)로 표시되며, 대안 목록 중 모든 대안에 대하여 대기시간을 동일하게 부여하거나, 유사도가 높은 대안에서부터 낮은 대안 순서로 대기시간을 차등적으로 부여할 수 있다. 또한, 모든 사용자에 대하여 일률적으로 동일한 대기시간을 부여하거나, 음성인식장치를 다루는 사용자의 숙련도에 따라서 음성인식장치가 장착된 기기의 사용자별로 서로 다른 대기시간을 부여할 수도 있다. 대기시간 설정부(21)에서 설정된 대기시간은 대기시간 설정부(21)에 입력된 음성인식부(13)의 음성인식결과와 함께 윈도우 생성부(24)로 제공된다.

숙련도 관리부(22)는 숙련도 DB(23)에 저장된 숙련도에 따른 선택시간에 소정의 여유시간을 부가한 값을 대기시간으로 재조정하여 대기시간 설정부(21)로 공급한다. 숙련도 관리부(22)는 후술하는 도 6에 도시된 과정을 통하여 대기시간을 재조정한다. 여기서, 대기시간은 모든 대안에 동일한게 적용되는 대기시간 또는 유사도 순서에 따라 차등적으로 부여되는 대기시간을 지칭한다.

숙련도 DB(23)는 숙련도에 따라서 서로 다른 선택시간을 대응시켜 데이터베이스화한 것이다. 여기서, 숙련도는 그래픽 사용자 인터페이스용 윈도우가 디스플레이된 후 최종 인식단어를 결정할 때까지 소요되는 선택시간에 반비례하는 변수로서, 소정 횟수에 대하여 소요된 선택시간을 평균한 값을 사용자의 숙련도로 결정한다.

윈도우 생성부(24)는 도 9에 도시된 바와 같이 현재 진행상황을 알려주는 메시지 영역(92), 대기시간 설정부(21)에서 설정한 대기시간에 해당하는 타임벨트(93) 및 음성인식부(13)의 음성인식결과를 유사도 순서에 따라서 나열한 대안목록(93)을 포함하는 그래픽 사용자 인터페이스용 윈도우(91)를 생성한다. 이때, 대안목록(93)의 대안 나열순서는 유사도 뿐만 아니라 이전의 음성인식 히스토리에서 나타난 오류패턴을 반영하여 결정할 수 있다.

오류패턴 관리부(25)는 음성인식부(13)에 의해 첫번째 대안으로 채택되었던 인식단어와 최종 인식단어 결정부(27)로부터 제공되는 최종 인식단어를 입력으로 하고, 해당 조합이 오류패턴 DB(26)에 저장되어 있는 경우 음성인식부(13)를 통해 윈도우 생성부(24)에 제공된 인식결과에 따른 스코어를 조정하여 윈도우 생성부(24)로 제공한다. 그러면, 윈도우 생성부(24)에서는 조정된 스코어를 기준으로 하여 대안목록(94)의 나열순서를 변경시키게 된다. 예를 들면, 첫번째 대안으로 채택된 인식단어가 "우회진"이고, 최종 결정된 인식단어가 "우리집"인 경우 "우리집"에 소정의 가중치를 부여한다. 그 결과, 음성인식부(13)에서의 "우회진"이 첫번째 대안으로 결정되더라도 윈도우 생성부(24)에서는 "우리집"을 "우회진"보다 높은 순위에 나열할 수 있다.

오류패턴 DB(26)는 음성인식부(13)에 의해 첫번째 대안으로 채택되었던 인식단어와 최종적으로 결정된 인식단어가 서로 다른 경우 이를 오류패턴이라 하며, 오류패턴 테이블의 일례는 도 4에 도시된 바와 같이 인식결과 1순위 단어(41), 최종인식단어(42), 사용자 발화특징 1 내지 n(43), 발화성향(44), 및 오류발생횟수 즉, 히스토리(n,45)로 이루어진다.

최종 인식단어 결정부(27)는 윈도우 생성부(24)에서 생성된 그래픽 사용자 인터페이스용 윈도우(91)에 나타난 대안목록(94)에 대하여 타임벨트(93)로 표시되는 대기시간동안 사용자의 부가작업이 행해지는지 여부에 따라서 현재 커서가 위치한 대안을 최종 인식단어로 결정한다. 즉, 윈도우(91)가 디스플레이된 후 대기시간 이내에 사용자로부터 별도의 키 또는 버튼입력이 없는 경우 현재 커서가 위치한 첫번째 대안을 최종 인식단어로 결정한다. 한편, 대기시간 이내에 사용자로부터 대안 선택용 키 또는 버튼이 입력되면 후술하는 도 7 또는 도 8에 도시된 과정을 통하여 최종 인식단어를 결정한다.

도 3은 도 2에 있어서 오류패턴 관리부(25)에서 오류패턴 DB(24)를 업데이트시키는 과정을 설명하는 흐름도이다.

도 3을 참조하면, 31 단계에서는 최종 인식단어 결정부(27)로부터 제공되는 첫번째대안 인식단어와 최종인식단어의 조합이 오류패턴 DB(26)에 포함되어 있는지를 판단하고, 첫번째대안 인식단어와 최종인식단어의 조합이 오류패턴 DB(26)에 포함되어 있지 않는 경우 본 흐름도를 종료한다.

32 단계에서는 상기 31 단계에서의 판단결과, 첫번째대안 인식단어와 최종인식단어의 조합이 오류패턴 DB(26)에 포함되어 있는 경우 발화특징 차이값을 산출한다. 발화특징 차이값은 해당 오류패턴에 대하여 오류패턴 DB(26)에 저장된 사용자 발화특징 1 내지 n(43)과 현재 입력된 음성의 각 발화특징 간의 차이에 대하여 절대값을 취한 값을 모두 더한 값을 의미한다.

33 단계에서는 상기 32 단계에서 구해진 발화특징 차이값과 업데이트를 위한 소정의 기준값인 제1 문턱치를 비교하고, 발화특징 차이값이 제1 문턱치보다 크거나 같은 경우에는 본 흐름도를 종료한다. 34 단계에서는 상기 33 단계에서의 비교결과, 발화특징 차이값이 제1 문턱치보다 작은 경우, 즉, 즉, 오류가 해당 오류패턴과 같은 이유, 예를 들면 감기에 걸려서, 아침에 목이 갈라져서, 주위가 시끄러워서 등으로 발생하였다고 판단될 경우에는 첫번째대안 인식단어와 최종인식단어가 동일한지를 판단한다. 여기서, 제1 문턱치는 실험적으로 또는 시뮬레이션을 통하여 최적의 값으로 설정될 수 있다.

35 단계에서는 상기 34 단계에서의 판단결과, 첫번째대안 인식단어와 이에 대응하는 최종인식단어가 서로 다른 경우 현재 음성의 발화특징 1 내지 n을 포함하여 발화특징 1 내지 n의 각 평균치를 산출하여 발화성향(44)을 업데이트시키고, 36 단계에서는 히스토리(n)의 값을 1 증가시켜 히스토리(45)를 업데이트시킨다.

37 단계에서는 상기 34 단계에서의 판단결과, 첫번째대안 인식단어와 최종인식단어가 동일한 경우 히스토리(n)이 '0'보다 큰지를 판단하고, 히스토리(n)이 '0'보다 작거나 같은 경우 본 흐름도를 종료한다. 38 단계에서는 상기 37 단계에서의 판단결과, 히스토리(n)이 '0'보다 큰 경우 히스토리(n)의 값을 1 감소시켜 히스토리(45)를 업데이트시킨다.

도 5는 도 2에 있어서 오류패턴 관리부(25)에서 대안 나열순서를 변경하는 과정을 설명하는 흐름도이다.

도 5를 참조하면, 51 단계에서는 음성인식부(13)를 통해 윈도우 생성부(24)에 제공된 다음 표 1에 도시된 바와 같은 인식결과 및 스코어를 참조하여 첫번째대안(1순위) 인식단어와 두번째대안(2순위) 인식단어의 조합이나 첫번째대안(1순위) 인식단어와 세번째대안(3순위) 인식단어의 조합이 오류패턴 DB(26)에 포함되어 있는지를 판단하고, 이러한 조합들이 오류패턴 DB(26)에 포함되어 있지 않는 경우 본 흐름도를 종료한다. 이때, 인식결과를 3순위까지 보여주는 경우를 예로 든다.

**[표 1]**

인식결과	스코어
황길두	10
홍기수	9
홍길동	8

52 단계에서는 상기 51 단계에서의 판단결과, 이들 조합 중 어느 하나가 오류패턴 DB(26)에 포함되어 있는 경우 각 조합에 대하여 발화특징 1 내지 n의 차이값을 산출한다. 도 3에서와 마찬가지로, 발화특징 차이값은 오류패턴 DB(26)에 저장된 사용자 발화특징 1 내지 n(43)과 현재 입력된 음성의 각 발화특징 간의 차이에 대하여 절대값을 취한 값을 모두 더한 값을 의미한다.

53 단계에서는 상기 52 단계에서 산출된 발화특징 차이값을 순서변경을 위한 소정의 기준제2 문턱치와 비교하고, 발화특징 차이값이 제2 문턱치보다 크거나 같은 경우 즉, 오류가 해당 오류패턴과 같은 이유로 발생하지 아니하였다고 판단될 경우 본 흐름도를 종료한다. 여기서, 제2 문턱치는 실험적으로 또는 시뮬레이션을 통하여 최적의 값으로 설정될 수 있다. 54 단계에서는 상기 53 단계에서의 비교결과, 발화특징 차이값이 제2 문턱치보다 작은 경우 즉, 오류가 해당 오류패턴과 같은 이유로 발생하였다고 판단될 경우 해당 대안의 스코어를 수정한다. 예를 들어, 오류패턴 DB(26)에 도 4에서와 같은 오류패턴 테이블이 저장되어 있고, 즉, 첫 번째대안 인식단어와 최종인식단어로서 세 번째대안 인식단어의 조합이 오류패턴으로 포함되어 있고, 가중치를 0.4 로 설정한 경우, 상기 표 1의 인식결과 및 스코어는 다음 표 2와 같이 변경되어진다. 여기서, 수정된 스코어(9.2)은 원래의 스코어(8)에 가중치(0.4)와 히스토리(3)를 곱한 값을 더함으로써 산출되어진다.

[표 2]

인식결과	스코어
황길두	10
홍길동	9.2
홍기수	9

한편, 도 3 내지 도 5에 적용된 사용자의 발화특징 1 내지 n(43)을 살펴보면, 음성인식부(13)에서 음성을 분석하면서 생성되는 정보 즉, 일부는 음성인식의 결과관정에 사용되고, 일부는 참조자료로만 사용되는 정보를 사용하거나, 별도의 방법으로 측정된 다음과 같은 정보를 사용한다.

첫째, 해당 갯수의 음절을 발화하는데 걸리는 시간으로 발화속도를 정의하고, 둘째, 하드웨어에 지정되어 있는 마이크 볼륨에 비하여 지나치게 작거나 큰 목소리 때문에 생기는 오류의 원인인 목소리 크기를 정의한다. 예를 들면 작은 목소리의 경우 소음에 묻히고, 큰 목소리의 경우 하드웨어가 받아들이지 못하는 부분이 생겨 신호의 왜곡을 발생시킨다. 셋째, 음성 입력이 없는 상황 또는 음절간의 공백시에 측정되는 해당 상황하에서의 기본 소음의 정도로서, 음성 대 소음비율(SNR)을 정의한다. 넷째, 감기, 아침에 발생하는 발성기관의 이상으로 인한 음성의 일부 변이 유무를 알 수 있는 특정상황에서의 목소리 갈라짐 상태를 정의한다. 그외 다양한 발화특징이 사용될 수 있다.

도 6은 도 2에 있어서 숙련도 관리부(22)에서 대기시간을 조정하는 과정을 설명하는 흐름도이다.

도 6을 참조하면, 61 단계에서는 숙련도 DB(23)에 저장된 숙련도별 선택시간 중 소정의 선택시간에서 현재의 최종인식단어 결정에 소요된 시간을 뺀 값으로 선택시간 차이값을 산출한다.

62 단계에서는 상기 61 단계에서 구해진 선택시간 차이값을 대기시간 변경을 위한 제3 문턱치와 비교하고, 선택시간 차이값이 제3 문턱치보다 큰 경우, 즉 주어진 선택시간이 사용자가 선택을 결정할 수 있는 시간에 비해 충분히 길다고 판단될 경우에는 선택시간을 수정한다(63 단계). 여기서, 제3 문턱치는 실험적으로 또는 시뮬레이션을 통하여 최적의 값으로 설정될 수 있다. 상기 63 단계에서 수정된 선택시간은 숙련도 DB(23)에 저장되어 있던 선택시간에서 선택시간 차이값에 소정의 가중치를 곱한 값을 감산함으로써 산출되어진다. 예를 들어, 숙련도 DB(23)에 저장되어 있던 선택시간을 0.8 초, 선택시간 차이값이 0.1 초, 가중치를 0.1 로 하는 경우 수정된 선택시간은 0.79 초가 된다. 수정된 선택시간은 숙련도 DB(23)에 저장되어 해당 사용자의 선택시간을 업데이트시킨다.

64 단계에서는 상기 62 단계에서의 비교결과, 선택시간 차이값이 제3 문턱치보다 작거나 같은 경우 즉, 사용자의 최종선택이 선택시간 종료이후 시스템의 타임아웃(timeout)에 의해 정해졌을 경우 선택시간 차이값을 소정의 여유시간과 비교하고, 선택시간 차이값이 여유시간보다 크거나 같은 경우 본 흐름도를 종료한다.

65 단계에서는 상기 64 단계에서의 비교결과, 선택시간 차이값이 여유시간보다 작은 경우 선택시간을 수정한다. 상기 65 단계에서 수정된 선택시간은 숙련도 DB(23)에 저장되어 있던 선택시간에 소정의 부가시간을 가산함으로써 산출되어진다. 예를 들어, 숙련도 DB(23)에 저장되어 있던 선택시간을 0.8 초, 부가시간을 0.02 초로 하는 경우 수정된 선택시간은 0.82 초가 된다. 수정된 선택시간은 숙련도 DB(23)에 저장되어 해당 사용자의 선택시간을 업데이트시킨다. 부가시간은 인식오류의 가능성이 있을 경우 다음 사용시의 잠재적인 오류를 방지하기 위하여 부가하는 시간으로서, 여기서는 0.02 초를 예로 든 것이다.

66 단계에서는 상기 63 단계 또는 상기 65 단계에서 수정된 선택시간에 소정의 여유시간을 가산함으로써 해당 사용자의 대기시간을 산출하여 대기시간 설정부(21)로 제공한다. 여유시간은 사용자의 의도와 상관없이 선택되지 않도록 하기 위하여 여유를 갖는 시간으로서, 여기서는 0.3 초를 예로 든 것이다.

도 7은 본 발명에 따른 음성인식방법의 제1 실시예를 설명하는 흐름도로서, 대안목록 디스플레이단계(71 단계), 사용자 선택변경이 없는 경우에 대한 처리단계(72, 73 및 78 단계) 및 사용자 선택변경이 있는 경우에 대한 처리단계(74 내지 78 단계)로 이루어진다.

도 7을 참조하면, 71 단계에서는 음성인식부(13)의 음성인식결과에 따른 대안목록(94)을 포함한 윈도우(91)를 디스플레이한다. 본 발명에서는 윈도우(91)가 디스플레이되는 순간에는 대안목록(94)에서 커서가 항상 첫번째 대안에 위치하는 것으로 설정하고, 윈도우(91)가 디스플레이되는 순간부터 타임벨트(93)가 진행된다. 72 단계에서는 사용자로부터 별도의 키 또는 버튼 입력이 없이 대기시간 설정부(21)에서 설정된 초기 대기시간이 경과하였는지를 판단한다.

73 단계에서는 상기 72 단계에서의 판단결과, 초기 대기시간이 경과한 경우 현재 커서가 위치한 첫번째 대안을 최종 인식단어로 결정하고, 인식단어에 해당하는 기능을 수행한다(78 단계). 한편, 74 단계에서는 상기 72 단계에서의 판단결과, 초기 대기시간이 경과하지 않은 경우 사용자로부터 별도의 키 또는 버튼 입력에 의해 선택변경이 행해졌는지를 판단한다.

75 단계에서는 상기 74 단계에서의 판단결과, 사용자의 선택변경이 행해진 경우, 대기시간을 재조정한다. 이때, 재조정되는 대기시간은 초기 대기시간과 동일하거나, 대안 나열 순서에 따라서 서로 다를 수 있다. 한편, 상기 74 단계에서의 판단결과, 사용자 선택변경이 행해지지 않은 경우 76 단계로 이행한다. 예를 들면, 도 9에 있어서 '탄성주기'로 사용자가 선택을 변경한 경우, 윈도우(91)의 메시지 영역(92)에는 "탄성주기를 실행 중입니다"라는 메시지와 함께 재조정된 대기시간에 따른 타임 벨트(93)가 작동한다.

76 단계에서는 상기 75 단계에서 재조정된 대기시간 또는 초기 대기시간이 경과하였는지를 판단하고, 재조정된 대기시간 또는 초기 대기시간이 경과하지 않은 경우 상기 74 단계로 복귀하여 재차 사용자의 선택변경이 행해졌는지를 판단한다. 한편, 상기 76 단계에서의 판단결과, 재조정된 대기시간 또는 초기 대기시간이 경과한 경우 사용자의 선택변경에 의해 현재 커서가 위치한 대안을 최종 인식단어로 결정하고, 인식단어에 해당하는 기능을 수행한다(78 단계).

도 8은 본 발명에 따른 음성인식방법의 제2 실시예를 설명하는 흐름도로서, 대안목록 디스플레이단계(81 단계), 사용자 선택변경이 없는 경우에 대한 처리단계(82 내지 86 단계) 및 사용자 선택변경이 있는 경우에 대한 처리단계(84 내지 86 단계)로 이루어진다.

도 8을 참조하면, 81 단계에서는 음성인식부(13)의 음성인식결과에 따른 대안목록(94)을 포함한 윈도우(91)를 디스플레이한다. 윈도우(91)가 디스플레이되는 순간부터 타임벨트(93)가 진행된다. 82 단계에서는 사용자로부터 별도의 키 또는 버튼 입력이 없이 대기시간 설정부(21)에서 설정된 초기 대기시간이 경과하였는지를 판단한다.

83 단계에서는 상기 82 단계에서의 판단결과, 초기 대기시간이 경과한 경우 현재 커서가 위치한 첫번째 대안을 최종 인식단어로 결정하고, 인식단어에 해당하는 기능을 수행한다(86 단계). 한편, 84 단계에서는 상기 82 단계에서의 판단결과, 초기 대기시간이 경과하지 않은 경우 사용자로부터 별도의 키 또는 버튼 입력에 의해 선택변경이 행해졌는지를 판단한다.

85 단계에서는 상기 84 단계에서의 판단결과, 사용자의 선택변경이 행해진 경우, 사용자의 선택변경에 의해 현재 커서가 위치한 대안을 최종 인식단어로 결정하고, 인식단어에 해당하는 기능을 수행한다(86 단계). 한편, 상기 84 단계에서의 판단결과, 사용자 선택변경이 행해지지 않은 경우 82 단계로 복귀한다.

다음, 다양한 인식환경하에서 작업성공률 및 부가작업 횟수에 대하여 기존의 음성인식방법과 본 발명에 따른 음성인식방법을 비교한 결과는 다음 표 3에 나타난 바와 같다.

[표 3]

대안의 제시방식	90% 인식환경				70% 인식환경			
	부가작업 0회	부가작업 1회	부가작업 2회	계	부가작업 0회	부가작업 1회	부가작업 2회	계
기존방식1	90%	0%	0%	90%	70%	0%	0%	70%
기존방식2	0%	90%	0%	90%	0%	70%	0%	70%
기존방식3	0%	99.9%	0%	99.9%	0%	97.3%	0%	97.3%
본 발명	90%	9%	0.9%	99.9%	70%	21%	6.3%	97.3%

표 3을 참조하면, 기존방식 1은 대안을 제시하지 않는 방식이고, 기존방식 2는 최선의 대안에 대하여 사용자가 가부를 결정하는 방식이고, 기존방식 3은 인식결과에 대한 일련의 대안 목록 중 사용자가 하나를 선택하는 방식이다. 표 1에 나타난 데이터는 90% 인식환경은 일반 사무실 소음 환경, 70% 인식환경은 차량의 고속도로 운행소음 환경을 말하며, 인식대상 단어목록이 무한대이고 목록에 포함된 단어들간의 유사도가 서로 동일하다는 가정하에 산출된 것이다. 표 1에 따르면, 본 발명에 의한 음성인식방법을 적용하는 경우, 부가작업이 행해지지 않을 경우의 초기 작업성공율에 덧붙여 부가작업을 거듭 수행할수록 최종 작업성공율을 극대화시킬 수 있음을 알 수 있다.

한편, 상기한 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 예를 들면, 음성인식방법은 사용자가 발화한 음성에 대한 음성인식결과, 소정 갯수의 단어를 소정의 순서에 따라서 나열된 대안목록을 디스플레이하는 제1 프로그램, 및 소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 있었는지를 판단하고, 상기 소정의 대기시간동안 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 대안을 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우 상기 대기시간을 재조정하는 다음, 재조정된 대기시간동안 재차 사용자의 선택변경이 있는지를 판단하여 사용자의 선택변경이 없는 경우 사용자가 선택한 대안을 최종 인식단어로 결정하는 제2 프로그램을 기록한 컴퓨터가 읽을 수 있는 기록매체로 구현가능하다. 여기서, 제2 프로그램은 소정의 대기시간동안 상기 디스플레이된 대안목록에 대한 사용자의 선택변경이 있었는지를 판단하고, 상기 소정의 대기시간동안 사용자의 선택변경이 없는 경우, 상기 대안목록 중 현재 커서가 위치한 대안을 최종 인식단어로 결정하고, 상기 소정의 대기시간동안 사용자의 선택변경이 있는 경우, 사용자가 선택한 대안을 최종 인식단어로 결정하는 기능을 구비하는 것으로 대체가능하다.

컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플라피디스크, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

그리고 본 발명을 구현하기 위한 기능적인(functional) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.

삭제

또한, 본 발명에 따른 음성인식방법 및 장치는 개인용 컴퓨터, 휴대폰이나 PDA(Personal Digital Assistant) 등과 같은 개인용 이동통신기기 등의 다양한 플랫폼에 적용되어, 음성인식에 따른 작업성공율을 향상시킬 수 있다.

**발명의 효과**

상술한 바와 같이 본 발명에 따르면, 열악한 음성인식 환경에서 사용자의 부가작업과 심리적인 압박을 최소화시키면서도 음성명령을 통한 작업의 최종 성공율을 극대화함으로써 작업의 효율성을 향상시킬 수 있다.

또한 본 발명에 따르면, 일정한 대기시간 동안 사용자의 선택변경이 행해지지 않는 경우 자동적으로 후속 작업을 수행함으로써 음성인식을 위해 요구되는 사용자의 버튼 조작 횟수를 최소화시킬 수 있어 사용자에게 편의를 제공하여 시스템에 대한 만족도를 증대시킬 수 있을 뿐 아니라, 사용자별로 적응적으로 조정된 대기시간을 할당함으로써 음성인식에 의해 실행되는 작업의 처리속도를 단축할 수 있다.

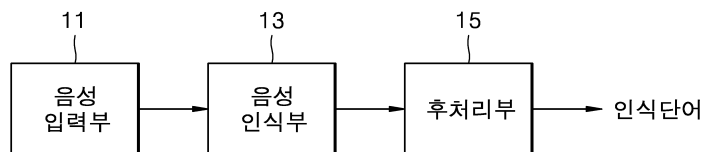
이상 도면과 명세서에서 최적 실시예들이 개시되었다. 여기서 특정한 용어들이 사용되었으나, 이는 단지 본 발명을 설명하기 위한 목적에서 사용된 것이지 의미 한정이나 특허청구범위에 기재된 본 발명의 범위를 제한하기 위하여 사용된 것은 아니다. 그러므로 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 특허청구범위의 기술적 사상에 의해 정해져야 할 것이다.

**도면의 간단한 설명**

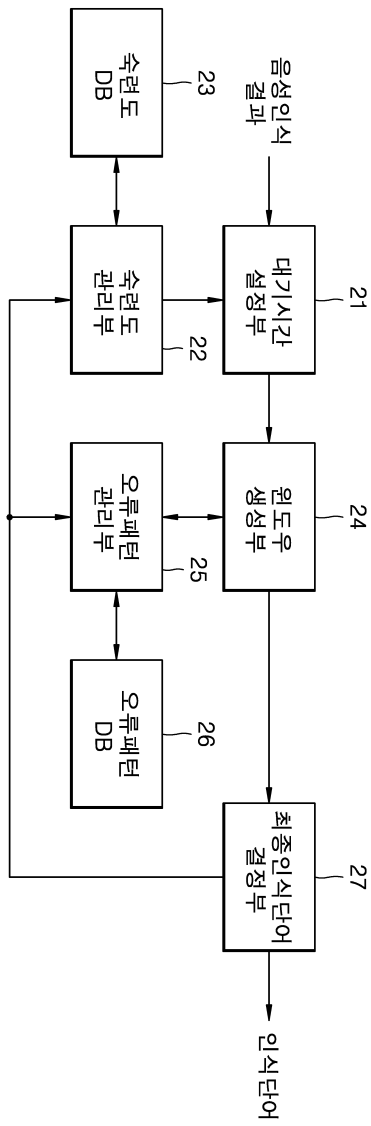
- 도 1은 본 발명의 일실시예에 따른 음성인식장치의 구성을 나타낸 블록도,
- 도 2는 도 1에 있어서 후처리부의 세부적인 구성을 나타낸 블록도,
- 도 3은 도 2에 있어서 오류패턴 관리부에서 오류패턴 DB를 업데이트시키는 과정을 설명하는 흐름도,
- 도 4는 도 2에 있어서 오류패턴 DB의 일예를 나타낸 테이블,
- 도 5는 도 2에 있어서 오류패턴 관리부에서 대안 나열순서를 변경하는 과정을 설명하는 흐름도,
- 도 6은 도 2에 있어서 숙련도 관리부에서 대기시간을 조정하는 과정을 설명하는 흐름도,
- 도 7은 본 발명에 따른 음성인식방법의 제1 실시예를 설명하는 흐름도,
- 도 8은 본 발명에 따른 음성인식방법의 제2 실시예를 설명하는 흐름도, 및
- 도 9는 본 발명에 따라서 생성되는 그래픽 사용자 인터페이스의 일예를 보여주는 도면이다.

**도면**

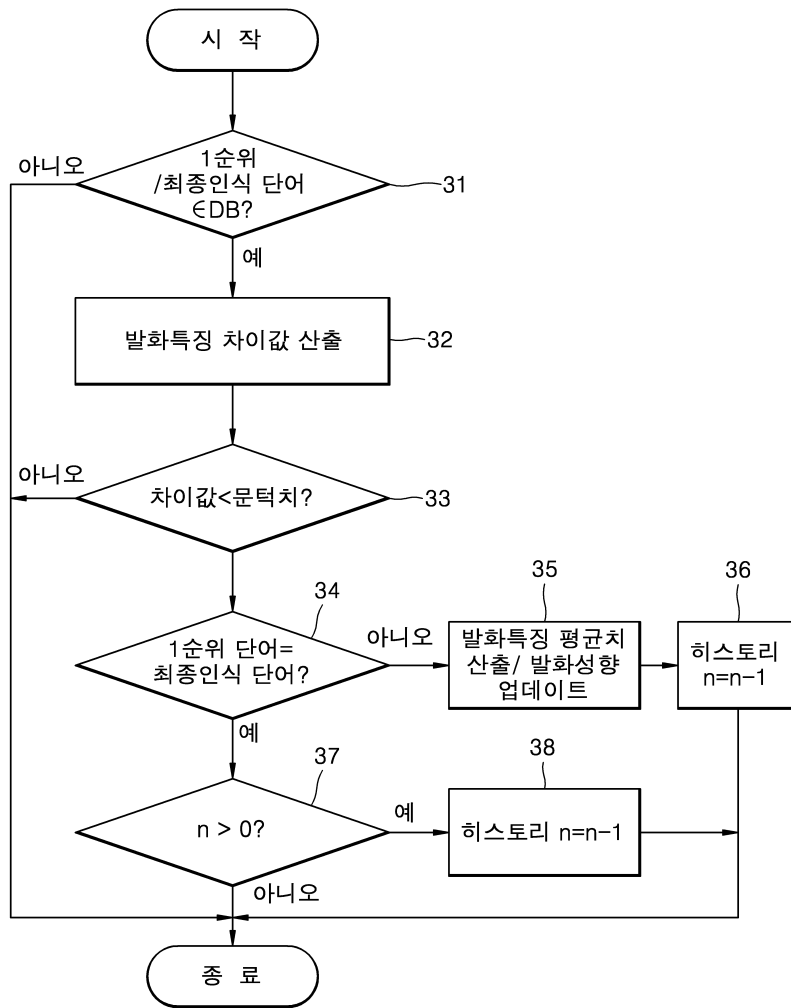
**도면1**



도면2



도면3

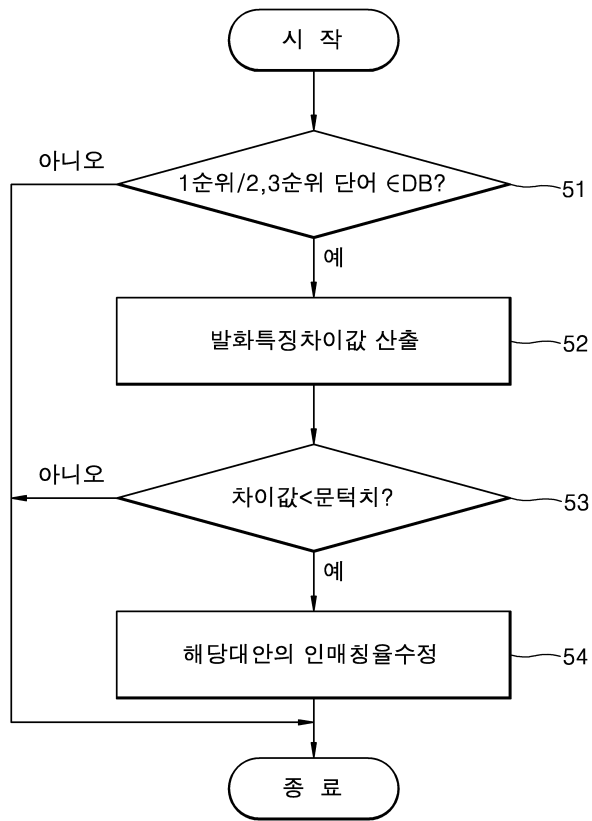




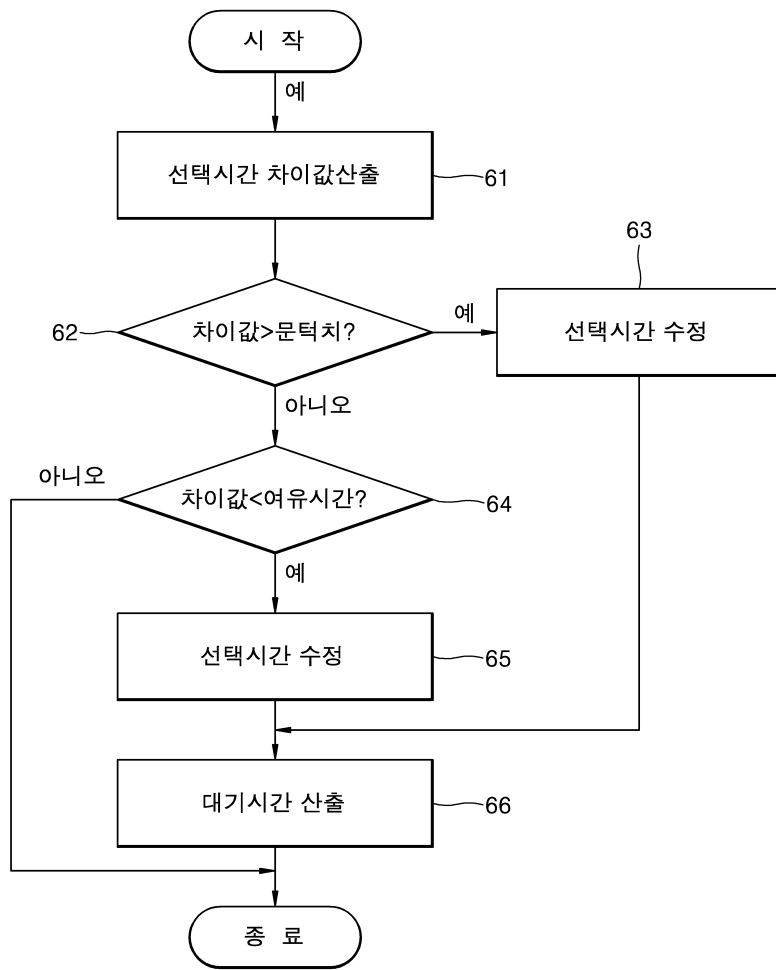
도면4

환경두	인식결과 1순위 단어	41			환경동	최종인식 단어	42		
	사용자 발화특정1	XX	...	사용자 발화특정n		43	발화성향	XX	44
	45	히스토리(n)			3				

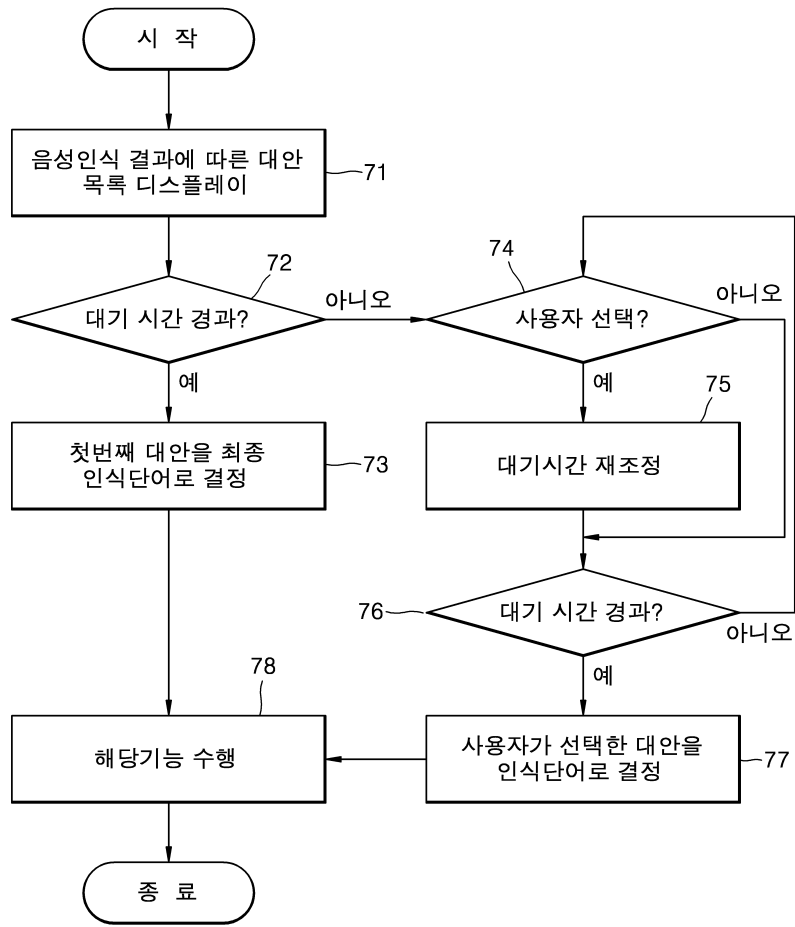
도면5



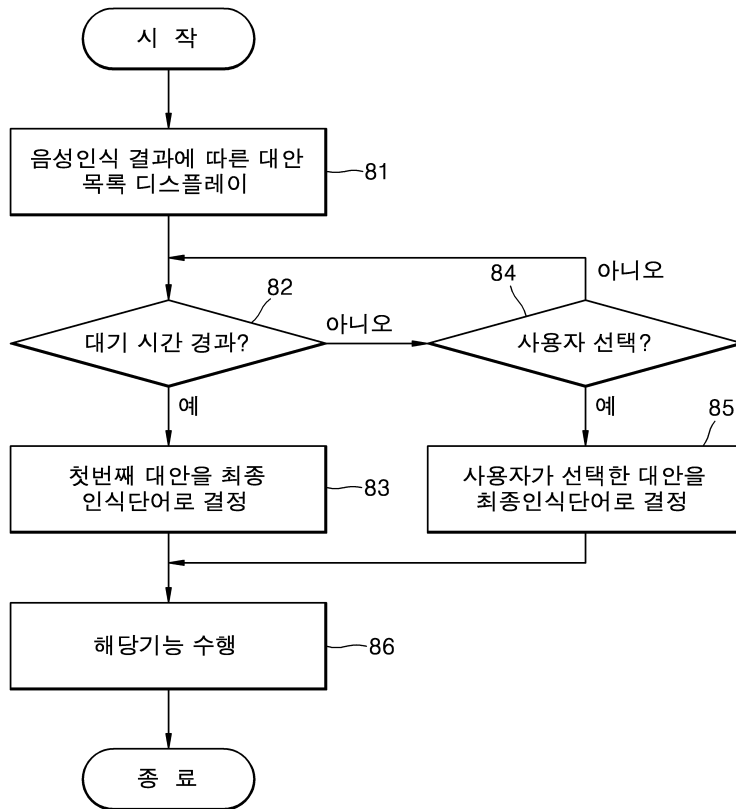
도면6



도면7



도면8



도면9

