

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) Int. Cl.	(45) 공고일자	2006년10월13일
G10L 15/22 (2006.01)	(11) 등록번호	10-0632912
G10L 15/00 (2006.01)	(24) 등록일자	2006년09월29일
G10L 15/26 (2006.01)		

(21) 출원번호	10-2004-7010350	(65) 공개번호	10-2004-0072691
(22) 출원일자	2004년06월29일	(43) 공개일자	2004년08월18일
번역문 제출일자	2004년06월29일		
(86) 국제출원번호	PCT/US2002/040794	(87) 국제공개번호	WO 2003/058604
국제출원일자	2002년12월20일	국제공개일자	2003년07월17일

(30) 우선권주장 10/034,542 2001년12월29일 미국(US)

(73) 특허권자 모토로라 인코포레이티드
 미국, 일리노이 60196, 샤움버그, 이스트 엘공킨 로드 1303

(72) 발명자 발라스리야, 세나카
 미국 60559 일리노이주 웨스트몬트 넘버 209 웨스트 파운틴헤드 드라
 이브 63

(74) 대리인 주성민
 백만기
 이중희

심사관 : 경연정

(54) 멀티-레벨 분산 음성 인식을 위한 방법 및 장치

요약

멀티 레벨 분산 음성 인식용 시스템 및 방법은 마이크로폰(130)에 결합된 터미널 음성 인식기(136)를 구비하는 터미널 (122)을 포함한다. 터미널 음성 인식기(136)는 오디오 커맨드(37)를 수신하고, 터미널 신뢰값을 갖는 적어도 하나의 터미널 인식 오디오 커맨드를 생성한다. 적어도 하나의 네트워크 음성 인식기(150)를 갖는 네트워크 엘리먼트(124)도 또한 오디오 커맨드(149)를 수신하고, 네트워크 신뢰값을 갖는 적어도 하나의 네트워크 인식 오디오 커맨드를 생성한다. 비교기 (152)는 인식된 오디오 커맨드를 수신하고, 음성 인식 신뢰값과 비교한다. 비교기(152)는 출력(162)을 적어도 하나의 인식된 오디오 커맨드의 대화 관리자(dialog manager)에 제공하고, 다음에 대화 관리자는 적어도 하나의 인식된 오디오 커맨드에 기초하여 오퍼레이션을 실행하며, 적어도 하나의 인식된 오디오 커맨드를 사용자에게 제시하여 콘텐츠 서버를 검증 및 액세스하도록 한다.

대표도

도 2

색인어

콘텐츠 서버, 음성 인식 장치, 오디오 커맨드, 마이크로폰, 네트워크 신뢰값

명세서

기술분야

본 발명은 일반적으로 통신 디바이스 및 방법에 관한 것이고, 특히, 음성 인식용 통신 디바이스 및 방법에 관한 것이다.

배경기술

핸드헬드 디바이스, 이동 전화, 랩탑, PDA, 인터넷 설비, 데스크탑 컴퓨터, 또는 적합한 디바이스와 같은 터미널 디바이스를 포함하는 최근의 기술 영역은 복수의 입출력 포맷에서 정보 전달의 어플리케이션에 관한 것이다. 전형적으로 터미널 디바이스에 상주하는 것은 특정 정보 요청과 같은 정보를 사용자가 입력하도록 하는 입력 시스템이다. 예컨대, 사용자는 특정 도시에 대한 날씨 정보를 얻도록 날씨 데이터베이스에 액세스하기 위해 터미널 디바이스를 이용할 수 있다. 전형적으로, 사용자는 "시카고의 날씨"와 같은 특정 위치에 대한 날씨 정보를 문의하기 위해 음성 커맨드를 입력할 수 있다. 터미널 디바이스에 관련된 처리 한계로 인해, 음성 커맨드는 통신 링크를 통해 네트워크 엘리먼트로 전송될 수 있고, 네트워크 엘리먼트는 네트워크내의 복수의 네트워크 엘리먼트 중의 하나이다. 네트워크 엘리먼트는 음성 커맨드를 인식하고, 사용자 요청 정보를 실행 및 검색하는 음성 인식 엔진을 포함한다. 또한, 음성 인식 엔진은 네트워크내에 배치되어 네트워크 엘리먼트내에 상주하는 대신에 네트워크 엘리먼트에 동작가능하게 결합되어, 음성 인식 엔진은 복수의 네트워크 엘리먼트에 의해 액세스될 수 있다.

무선 기술의 진보에 따라, 무선 디바이스에 대한 사용자 어플리케이션은 증가되어 왔다. 많은 이러한 디바이스들은 더욱 상호작용하게 되고, 사용자가 커맨드 요청을 입력하고, 정보에 액세스할 수 있게 한다. 따라서, 이러한 무선 기술의 진보에 따라, 사용자가 특정 정보 요청을 제출하는 형태도 증가되어 왔다. 전형적으로, 사용자는 키패드를 통해 커맨드 요청을 입력할 수 있고, 터미널 디바이스는 이러한 입력을 부호화하여 네트워크 엘리먼트에 제공한다. 이러한 시스템의 통상적인 예는, 사용자가 계좌 번호 및 계좌 정보에 액세스하는 개인 식별 번호(PIN)를 입력하는 텔리폰 뱅킹 시스템이다. 터미널 디바이스 또는 네트워크 엘리먼트는, 키패드를 통해 입력을 수신하자마자, 입력을 듀얼 톤 멀티-프리퀀시 신호(DTMF)로 변환하여, 이 DTMF 신호를 뱅킹 서버에 제공한다.

게다가, 사용자는 음성 입력을 이용하여, 정보 요청과 같은 커맨드를 입력할 수 있다. 음성 인식 기술의 진보에도 불구하고, 프로세싱 및 메모리 저장장치의 요구사항들이 터미널 디바이스내에서 음성 인식 능력을 제약한다. 전형적으로, 음성 인식 엔진은 입력 음성 커맨드에 매칭하는 음성 모델의 라이브러리를 포함한다. 신뢰할만한 음성 인식을 위해, 방대한 라이브러리가 요구되고, 이에 따라 상당한 양의 메모리도 필요하게 된다. 또한, 음성 인식 용량의 증가에 따라, 전력 소비도 증가하여, 터미널 디바이스 배터리의 수명도 단축된다.

터미널 음성 인식 엔진은 적응적 시스템일 수 있다. 음성 인식 엔진은, 더 적은 인식된 커맨드의 라이브러리를 가지면서, 더 적응적이고, 톤, 억양, 액센트등과 같은 사용자의 차별적인 음성 패턴을 이해할 수 있게 된다. 따라서, 터미널내의 제한된 음성 인식 라이브러리는 더 높은 수준의 정확한 음성 인식 확률에 의해 오프셋된다. 사용자가 이전에 이름을 말하면 시스템이 자동적으로 관련 번호를 다이얼링하는 프로그램된 음성 활성화 다이얼링 기능과 같이, 이러한 시스템은 대부분 미리 터미널에 프로그램된 가장 통상적인 음성 커맨드들에만 속박된다.

음성 인식을 위한 다른 방법은 네트워크 엘리먼트에 완전한 음성 커맨드를 제공하는 것이다. 네트워크 음성 인식 엔진은 방대한 양의 가용 메모리에 의해 음성 인식 효율을 증가시키고, 관련된 전력 소비 요구에 대한 우려를 감소시킨다. 네트워크 엘리먼트상에서, 음성 인식 엔진이 복수의 네트워크 엘리먼트를 액세스하는 복수의 사용자에게 의해 액세스가능하여야 함에도 불구하고, 네트워크 음성 인식 엔진은 액센트등과 같은 특유의 음성 패턴을 인식할 수 있게 되는 것에 의해 제한되는 것은 아니다. 이와 같이, 네트워크 음성 인식 엔진은 음성-인식된 커맨드에 대한 더 많은 어휘를 제공할 수 있지만, 개별적인 사용자 음성 패턴내의 고유한 제한으로 인해, 적절하게 인식할 확률을 낮게 한다.

또한, 최근의 기술 발전은 터미널 디바이스가 음성 커맨드를 인식하기를 시도하는 멀티-레벨 분산 음성 인식을 제공하고, 터미널내에서 인식되지 않으면, 음성 커맨드는 부호화되고, 제2 음성 인식 시도를 위해 네트워크 음성 인식 엔진에 제공된다. Hedin등에게 허여된 미국 특허 제 6,185, 535 B1은 서비스 어플리케이션에 대한 사용자 인터페이스의 음성 제어를 위

한 시스템 및 방법을 개시하고 있다. 이러한 시스템은 스텝-와이즈(step-wise) 음성 인식을 제공하는데, 여기서는 적어도 하나의 네트워크 음성 인식 엔진만이 터미널 디바이스가 음성 커맨드를 인식하지 못하는 경우에 이용된다. 미국 특허 제 6,185,535는 오디오 커맨드가 터미널 음성 인식 엔진 또는 네트워크 음성 인식 엔진 중 하나로부터 정확하게 인식되는 것을 보장하는 단일 레벨만을 제공한다.

이와 같이, 음성 인식 엔진을 채용하는 개선된 통신 디바이스가 필요하게 된다.

도면의 간단한 설명

본 발명은 다음의 첨부도면에 의해 더욱 명확하게 이해될 것이다.

도 1은 종래의 무선 시스템을 도시하는 도면.

도 2는 본 발명의 일 실시예에 따른 멀티-레벨 분산 음성 인식용 장치의 블록도.

도 3은 본 발명의 일 실시예에 따른 멀티-레벨 분산 음성 인식을 위한 방법을 나타내는 플로우차트.

도 4는 본 발명의 일 실시예에 따른 멀티-레벨 분산 음성 인식을 위한 시스템의 블록도.

도 5는 본 발명의 일 실시예에 따른 멀티-레벨 분산 음성 인식을 위한 방법을 나타내는 플로우차트.

실시예

일반적으로, 터미널 디바이스의 오디오 서브시스템내의 마이크로폰에 동작가능하게 결합되고, 예컨대, "시카고의 날씨"와 같이 사용자로부터 제공된 음성 커맨드와 같은 오디오 커맨드를 수신하고, 적어도 하나의 터미널 인식된 오디오 커맨드를 생성하는 터미널 음성 인식 엔진을 통해 멀티-레벨 분산 음성 인식을 제공하는 시스템 및 방법이 제공되며, 여기서 적어도 하나의 터미널 인식된 오디오 커맨드는 대응하는 터미널 신뢰값(confidence value)을 가진다.

이러한 시스템 및 방법은 터미널내에서 마이크로폰에 동작가능하게 결합하고, 오디오 커맨드를 수신하고, 적어도 하나의 네트워크 인식된 오디오 커맨드를 생성하는 적어도 하나의 네트워크 음성 인식 엔진을 갖는 네트워크내의 네트워크 엘리먼트를 더 포함하고, 여기서 적어도 하나의 네트워크 인식된 오디오 커맨드는 대응하는 네트워크 신뢰값을 갖는다.

또한, 시스템 및 방법은 복수의 인식된 오디오 커맨드와 신뢰값을 비교하는 하드웨어 또는 소프트웨어로 구현된 모듈인 비교기를 포함한다. 비교기는 터미널 인식된 오디오 커맨드 및 터미널 음성 인식 신뢰값을 수신하기 위한 터미널 음성 인식 엔진에 동작가능하게 결합되고, 비교기는 네트워크 인식된 오디오 커맨드 및 네트워크 음성 인식된 신뢰값을 수신하기 위한 네트워크 음성 인식 엔진에 더 결합된다. 비교기는 터미널 음성 인식 신뢰값과 네트워크 음성 인식 신뢰값을 비교하고, 인식된 커맨드를 그 대응하는 신뢰값에 의해 컴파일링 및 분류한다. 일 실시예에서, 비교기는 특정 음성 인식 엔진에 기초하여 신뢰값에 가중 팩터를 부여하고, 특정 음성 인식 엔진으로부터의 신뢰값은 다른 신뢰값보다 더 큰 가중치가 부여된다.

대화 관리자(dialog manager)가 비교기에 동작가능하게 결합되는데, 대화 관리자는 당업자가 잘 아는, 음성 브라우저, 인터랙티브 음성 응답 유닛(IVR), 그래픽 브라우저, JAVA, 기반 어플리케이션, 소프트웨어 프로그램 어플리케이션, 또는 다른 소프트웨어/하드웨어 어플리케이션일 수 있다. 대화 관리자는 인식된 오디오 커맨드의 수신에 따라 커맨드를 수신, 해석 및 실행하는 하드웨어 또는 소프트웨어로 구현된 모듈이다. 대화 관리자는, 비교기에 N-최상 표시기(N-best indicator)를 제공할 수 있는데, N-최상 표시기는 대화 관리자로 제공될 최고의 신뢰값을 갖는 인식된 커맨드의 번호를 표시한다. 비교기는 인식된 오디오 커맨드 및 그 신뢰값, 즉, N-최상 인식된 오디오 커맨드 및 그 신뢰값의 관련 리스트를 대화 관리자에 제공한다. 또한, 비교기가 대화 관리자에게 임의의 인식된 오디오 커맨드를 제공하지 못하는 경우에는, 비교기는 에러 통지를 대화 관리자에게 제공한다.

대화 관리자가 하나 이상의 인식된 오디오 커맨드 및 대응하는 신뢰값을 수신할 때, 대화 관리자는 전술한 리스트를 더 한정하기 위한 추가적인 단계를 이용할 수 있다. 예컨대, 대화 관리자는 최고의 신뢰값으로 오디오 커맨드를 실행하거나 그 관련 리스트를 사용자에게 제시할 수 있어서, 사용자는 오디오 커맨드를 검증할 수 있다. 또한 대화 관리자가 에러 통지를 수신하거나, 인식된 오디오 커맨드의 어느것도 미리 정해진 최소 임계값 이상의 신뢰값을 갖지 않는 경우에는, 대화 관리자는 에러 메시지를 사용자에게 제공한다.

오디오 커맨드가 콘텐츠 서버로부터의 정보를 위해 요청되면, 대화 관리자는 콘텐츠 서버에 액세스하고, 부호화된 정보를 검색한다. 인터넷을 통해 결합된 상업적으로 이용가능한 서버, 인트라넷을 통한 로컬 상주 서버, बैं킹 시스템과 같은 상업적 어플리케이션 서버, 또는 임의의 다른 적절한 콘텐츠 서버와 같은 적어도 하나의 콘텐츠 서버가 대화 관리자에 동작가능하게 결합된다.

검색된 부호화된 정보는 대화 관리자에게 다시 제공되며, 하이퍼텍스트 마크-업 언어(HTML), 무선 마크-업 언어(WML), 확장 마크-업 언어(XML), 음성 확장 마크-업 언어(VoiceXML), 확장가능 하이퍼텍스트 마크업 언어(XHTML), 또는 다른 마크-업 언어와 같은 마크-업 언어로 부호화되어 복호화를 위해 대화 관리자에 제공된다. 따라서, 부호화된 정보는 대화 관리자에 의해 복호화되고, 사용자에게 제공된다.

이에 따라, 오디오 커맨드는, 제1 음성 인식 엔진은 터미널 디바이스상에 배치되고, 제2 음성 인식 엔진은 네트워크상에 배치되는 것과 같이, 멀티 레벨상에 배치될 수 있는 적어도 2개의 음성 인식 엔진간에 분산된다.

도 1은 사용자(102) 액세스를 터미널(108)과 네트워크 엘리먼트(110)사이에서 통신 링크(106)를 통해 적어도 하나의 콘텐츠 서버(104)에 제공하는 종래의 무선 통신 시스템(100)을 도시한다. 네트워크 엘리먼트(110)는 네트워크(112)내의 복수의 네트워크 엘리먼트(110)의 하나이다. 사용자(102)는 예컨대, "시카고의 날씨" 등의 음성 커맨드와 같은 입력 커맨드(114)를 터미널(108)에 제공한다. 터미널(108)은 커맨드를 해석하고, 표준 무선 접속과 같은 통신 링크(106)를 통해 네트워크 엘리먼트(110)에 커맨드를 제공한다.

네트워크 엘리먼트(110)는 커맨드를 수신하고, 커맨드를 처리, 즉, 음성 인식기(도시되지 않음)를 이용하여 입력 커맨드(114)를 인식 및 해석하고, 다음에 복수의 콘텐츠 서버(104)의 적어도 하나에 액세스하여 요청된 정보를 검색한다. 일단 정보가 검색되면, 네트워크 엘리먼트(110)에 다시 제공된다. 따라서, 요청된 정보는 통신 링크(106)를 통해 터미널(108)에 제공되며, 터미널(108)은 가청 메시지와 같은 출력(116)을 사용자에게 제공한다.

도 1의 종래의 시스템에서는, 입력 커맨드(114)는 터미널(108)에 제공되는 음성 커맨드일 수 있다. 터미널(108)은 음성 커맨드를 부호화하고, 부호화된 음성 커맨드를 통신 링크(106)를 통해 네트워크 엘리먼트(110)에 제공한다. 전형적으로, 네트워크 엘리먼트(110)내의 음성 인식 엔진은 음성 커맨드를 인식하기를 시도할 것이고, 이에 따라 요청된 정보를 검색한다. 전술한 바와 같이, 음성 커맨드(114)는 터미널(108)내에서 해석되고, 다음에 터미널은 네트워크 엘리먼트(110)에 요청된 정보에 대한 요청을 제공한다.

오디오 커맨드(114)를 터미널(108)에 제공하고, 터미널(108)이 이러한 커맨드를 해석하기를 시도하는 것이 본 기술분야에 잘 알려져 있다. 터미널(108)이 커맨드(114)를 해석할 수 없는 경우에, 오디오 커맨드(114)는 다음에, 적어도 하나의 네트워크 음성 인식 엔진(도시되지 않음)에 의해 인식되는 네트워크 엘리먼트(110)에 통신 링크(106)를 통해 제공된다. 이러한 종래의 시스템은 스텝-와이즈(step-wise) 음성 인식 시스템을 제공하는데, 여기서 적어도 하나의 네트워크 음성 인식 엔진만이, 터미널 음성 인식 엔진이 음성 인식 커맨드를 인식할 수 없는 경우에 액세스된다.

도 2는 본 발명의 실시예에 따른 멀티-레벨 분산 음성 인식용 장치를 도시한다. 오디오 서브시스템(120)은 미국 메사추세츠 02111, 보스턴, 아틀란틱 에비뉴 695에 주소를 둔 SpeechWorks International, Inc에 의해 제조된 OpenSpeech 인식 엔진 1.0과 같은 제1 음성 인식 엔진(122) 및 적어도 하나의 제2 음성 인식 엔진(124) 모두에 동작가능하게 결합된다. 당업자가 알 수 있는 바와 같이, 임의의 다른 적절한 음성 인식 엔진도 여기에 이용될 수 있다. 오디오 서브시스템(120)은 접속(126)을 통해 음성 인식 엔진(122 및 124)에 결합된다. 제1 음성 인식 엔진(122)은 접속(130)을 통해 비교기(128)에 동작가능하게 결합되고, 제2 음성 인식 엔진(124)은 접속(132)을 통해 비교기(128)에 동작가능하게 결합된다.

비교기(128)는 접속(136)을 통해 대화 관리자(134)에 결합된다. 대화 관리자는 접속(140)을 통해 콘텐츠 서버(138)에 결합되고, 접속(144)을 통해 음성 합성 엔진(142)에 결합된다. 게다가, 음성 합성 엔진은 접속(146)을 통해 오디오 서브 시스템(120)에 동작가능하게 더 결합된다.

도 2의 장치의 동작은 본 발명의 실시예에 따른, 멀티-레벨 분산 음성 인식을 위한 방법을 도시하는 도 3을 참조하여 설명된다. 이 방법은 단계(150)에서 개시되고, 단계(152)에서 장치는 오디오 커맨드를 수신한다. 전형적으로, 오디오 커맨드는 오디오 서브시스템(120)에 제공된다. 더 구체적으로, 오디오 커맨드는 오디오 서브시스템(120)내에 배치된 마이크로폰(도시되지 않음)을 통해 제공될 수 있다. 당업자가 알 수 있는 바와 같이, 오디오 커맨드는 메모리 로케이션으로부터 판독되거나, 어플리케이션으로부터 제공되는등의 임의의 다른 적절한 수단으로부터 제공될 수 있다.

오디오 커맨드를 수신하자마자, 오디오 서브시스템은 오디오 커맨드를 단계(154)에 지시된 바와 같이, 제1 음성 인식 엔진(122) 및 적어도 하나의 제2 음성 인식 엔진(124)에 제공한다. 오디오 커맨드는 접속(126)을 지나 제공된다. 다음에, 제1 음성 인식 엔진(122)은 오디오 커맨드를 인식하여, 적어도 하나의 인식된 오디오 커맨드를 생성하고, 여기서 적어도 하나의 제1의 인식된 오디오 커맨드는 단계(156)에 지시된 바와 같이 대응하는 제1 신뢰값을 갖는다. 또한, 단계(158)에 지시된 바와 같이, 적어도 하나의 제2의 음성 인식 엔진은 오디오 커맨드를 인식하고, 적어도 하나의 제2의 인식된 오디오 커맨드를 생성하며, 여기서 적어도 하나의 제2의 인식된 오디오 커맨드는 대응하는 제2의 신뢰값을 갖는다. 적어도 하나의 제2 음성 인식 엔진은 제1 음성 인식 엔진과 동일한 오디오 커맨드를 인식하지만, 제1 음성 인식 엔진과는 독립적으로 오디오 커맨드를 인식한다.

제1 음성 인식 엔진(122)은 다음에, 적어도 하나의 제1의 인식된 오디오 커맨드를 접속(130)을 통해 비교기(128)에 제공하며, 적어도 하나의 제2 음성 인식 엔진(124)은 적어도 하나의 제2 음성 인식된 오디오 커맨드를 접속(132)을 통해 비교기(128)에 제공한다. 본 발명의 일 실시예에서, 비교기는, 적어도 하나의 제1 신뢰값에 제1 가중 팩터 만큼의 가중치를 부여하고, 적어도 하나의 제2 신뢰값에 제2 가중 팩터 만큼의 가중치를 부여한다. 예컨대, 비교기는 제1 음성 인식 엔진의 인식에 추종하며, 이에 따라 단계(160)에 지시된 바와 같이, 제1 신뢰값은 .95의 스케일링 팩터에 의해 승산되고, 제2 신뢰값은 .90의 스케일링 팩터에 의해 승산된다.

다음에, 비교기는 단계(162)에 지시된 바와 같이, 적어도 하나의 제1 신뢰값 및 적어도 하나의 제2 신뢰값에 기초하여, 적어도 하나의 제1의 인식된 오디오 커맨드 및 적어도 하나의 제2의 인식된 오디오 커맨드로부터 인식된 오디오 커맨드 신뢰값을 갖는 적어도 하나의 인식된 오디오 커맨드를 선택한다. 본 발명의 일 실시예에서, 대화 관리자는 비교기에 요청된 인식 커맨드의 수를 표시하는 N-최상 표시기를 제공하며, 여기서 N-최상 표시기가 5이면, 5-최상 인식된 커맨드가 있는 것이다.

대화 관리자(134)는 접속(136)을 통해 비교기(128)로부터 N-최상 인식된 오디오 커맨드와 같은 인식된 오디오 커맨드를 수신한다. 대화 관리자는 다음에, 단계(164)에 지시된 바와 같이, 적어도 하나의 인식된 오디오 커맨드에 기초하여 적어도 하나의 오퍼레이션을 실행한다. 예컨대, 대화 관리자는 단계(166)에 지시된 바와 같이, 인식된 오디오 커맨드의 N-최상 리스트를 사용자 검증을 위해 사용자에게 제공함으로써, 적어도 하나의 인식된 오디오 커맨드를 검증하는 것을 시도할 수 있다. 본 발명의 일 실시예에서, 대화 관리자(134)는 인식된 오디오 커맨드의 N-최상 리스트를 접속(144)을 통해 음성 합성 엔진(142)에 제공한다. 음성 합성 엔진(142)은 N-최상 인식된 오디오 커맨드를 합성하고, 그것을 접속(146)을 통해 오디오 서브시스템(120)에 제공한다.

또한, 대화 관리자는, 0.65와 같은 최소 신뢰 레벨과 적어도 하나의 인식된 오디오 커맨드 신뢰값을 비교하는 것과 같이, N-최상 리스트상에서 필터링 오퍼레이션을 더 수행하여, 최고의 신뢰값을 갖는 인식된 오디오 커맨드를 적절한 인식된 오디오 커맨드로 간단하게 지정하게 된다. 여기서, 대화 관리자는, 접속(140)을 통해, 특정 도시에 대한 날씨 정보와 같은 요청된 정보를 검색하기 위해 콘텐츠 서버(138)에 액세스하는 것과 같은 커맨드를 실행한다.

게다가, 단계(168)에 지시된 바와 같이, 비교기는 적어도 하나의 제1 신뢰값 및 적어도 하나의 제2 신뢰값이 최소 신뢰 레벨 이하일 때, 에러 통지를 생성한다. 예컨대, 도 2를 참조하면, 비교기(128)는 제1 신뢰값 및 제2 신뢰값이 비교되는 0.55와 같은 내부 최소 신뢰 레벨을 가질 수 있다. 제1 신뢰값 또는 제2 신뢰값 어느 것도 최소 신뢰 레벨보다 높지 않으면, 비교기는 접속(176)을 통해 대화 관리자(134)에 에러 통지를 발행한다.

또한, 대화 관리자는 N-최상 인식된 오디오 커맨드내에서와 같이, 인식된 오디오 커맨드가 대화 관리자 최소 신뢰 레벨 이상의 인식된 신뢰값을 포함하지 못하는 경우에 에러 통지를 발행할 수 있다. 또한, 제1 음성 인식 엔진 및 적어도 하나의 제2 음성 인식 엔진이 임의의 오디오 커맨드를 인식할 수 없게 되거나, 인식된 오디오 커맨드가 제1 음성 인식 엔진, 제2 음성 인식 엔진 또는 비교기에 의해 지정되는 최소 신뢰 레벨 이하인 때에도 에러 통지가 비교기에 의해 생성된다.

비교기(128) 또는 대화 관리자(134)중 하나를 통해, 에러 통지가 발행될 때, 다음에 대화 관리자는 에러 커맨드를 실행하고, 여기서 에러 커맨드는 접속(144)을 통해 음성 합성 엔진(142)에 제공되고, 접속(146)을 통해 오디오 서브시스템(120)을 경유하여 말단 사용자에게도 더 제공된다. 당업자가 알수 있는 바와 같이, 에러 커맨드는 시각적 디스플레이를 이용하여 임의의 다른 적절한 수단을 통해 사용자에게 제공될 수 있다.

따라서, 도 2의 장치는 멀티-레벨 분산 음성 인식을 제공한다. 단계(170)에 지시된 바와 같이, 일단 대화 관리자가 적어도 하나의 인식된 커맨드에 응답하여 오퍼레이션을 실행하면, 방법은 종료된다.

도 4는 본 발명에 따른 멀티-레벨 분산 음성 인식 시스템을 도시한다. 시스템(200)은 터미널(202) 및 네트워크 엘리먼트(204)를 포함한다. 당업자가 알 수 있는 바와 같이, 네트워크 엘리먼트(204)는 네트워크(206)내의 복수의 네트워크 엘리먼트(204) 중의 하나이다.

터미널(202)은 여러 가지 중에서, 스피커(208) 및 마이크로폰(210)을 포함하는 오디오 서브시스템(206)을 구비한다. 오디오 서브시스템(206)은 터미널 음성 전달 인터페이스(212)에 동작가능하게 결합된다. 또한, 터미널 세션 제어(214)는 터미널(202)내에 배치된다.

터미널(202)는 또한 접속(218)을 통해 오디오 서브시스템(206)에 동작가능하게 결합되고, 미국 일리노이즈주, 60196, 슈버그, 이스트 알콘킨 로드, 1301에 주소를 둔 Motorola, Inc에 의해 제조되어, 음성 활성화 다이얼링을 제공하는, Motorola i90cTM에 나타나는 것과 같은 터미널 음성 인식 엔진(216)을 구비한다. 당업자가 알 수 있는 바와 같이, 다른 적절한 음성 인식 엔진이 여기에 이용될 수 있다. 터미널 음성 인식 엔진(216)은 오디오 서브시스템(206)내의 마이크로폰(210)을 통해 사용자(222)로부터 최초로 제공되는 오디오 커맨드(220)를 수신한다.

터미널 세션 제어(214)는 네트워크 엘리먼트(204)내에 배치된 네트워크 엘리먼트 세션 제어(222)에 동작가능하게 결합된다. 당업자가 알 수 있는 바와 같이, 터미널 세션 제어(214) 및 네트워크 엘리먼트 세션 제어(222)는 세션의 지속기간동안에는 통신 세션의 초기화시에 통신하고, 통신 세션의 종료시에도 통신한다. 예컨대, 초기화 스타트업동안 어드레스 지정이 터미널(202)내에 배치된 다양한 엘리먼트 및 네트워크 엘리먼트(204)에 제공된다.

터미널 음성 전달 인터페이스(212)는 네트워크 엘리먼트(204)내에 배치된 네트워크 엘리먼트 음성 전달 인터페이스(224)에 동작가능하게 결합된다. 네트워크 엘리먼트 음성 전달 인터페이스(224)는, 미국 메사추세츠 020111, 보스턴, 아틀란틱 애비뉴, 695에 주소를 둔, SpeechWorks International, Inc에 의해 제조된 OpenSpeech 인식 엔진 1.0과 같은 적어도 하나의 네트워크 음성 인식 엔진(226)에 동작가능하게 더 결합된다. 당업자가 알 수 있는 바와 같이, 임의의 다른 적절한 음성 인식 엔진도 여기에 이용될 수 있다. 적어도 하나의 네트워크 음성 인식 엔진(226)은 접속(230)을 통해 비교기(228)에 더 결합되고, 비교기는 여러 가지중에서, 하드웨어 또는 소프트웨어중 하나로 구현될 수 있고, 터미널 음성 인식 엔진(216) 및 네트워크 음성 인식 엔진(226)으로부터 수신되어 인식된 오디오 커맨드로부터 적어도 하나의 인식된 오디오 커맨드를 선택한다.

비교기(228)는 접속(232)을 통해, 터미널(202)내에 배치된 터미널 음성 인식 엔진(216)에 더 결합된다. 비교기(228)는 접속(236)을 통해 대화 관리자(234)에 결합된다. 대화 관리자(234)는 복수의 모듈에 동작가능하게 결합하고, 접속(240)을 통해 음성 합성 엔진(238)에 결합되며, 적어도 하나의 콘텐츠 서버(104)에 결합된다. 당업자가 알 수 있는 바와 같이, 대화 관리자는 도 4에서는 단순화를 위해 생략된 복수의 다른 컴포넌트에 결합될 수도 있다.

도 5는 본 발명의 일 실시예에 따른, 멀티-레벨 분산 음성 인식을 위한 방법을 도시한다. 도 4를 참조하여 알 수 있는 바와 같이, 도 5의 방법은 오디오 커맨드가 터미널(202)내에서 수신될 때, 단계(300)에서 개시한다. 전형적으로, 오디오 커맨드는 사용자(102)로부터 터미널(202)에 제공되어 오디오 입력을 오디오 서브시스템(206)의 마이크로폰(210)에 제공한다. 오디오 입력은 단계(304)에 지시된 바와 같이, 표준 부호화 포맷에 의해 부호화되고, 터미널 음성 인식 엔진(216)에 제공되며, 터미널 음성 전달 인터페이스(212) 및 적어도 하나의 네트워크 엘리먼트 음성 전달 인터페이스(224)를 통해 적어도 하나의 네트워크 음성 인식 엔진(226)에 더 제공된다.

도 2의 장치와 유사하게, 터미널 음성 인식 엔진은 오디오 커맨드를 인식하고, 적어도 하나의 터미널 인식된 오디오 커맨드를 생성하고, 여기서 적어도 하나의 터미널 인식된 오디오 커맨드는 단계(306)에 지시된 바와 같이 대응하는 터미널 신뢰값을 갖는다. 또한, 적어도 하나의 네트워크 음성 인식 엔진(226)은 오디오 커맨드를 인식하고, 적어도 하나의 네트워크 인식된 오디오 커맨드를 생성하며, 여기서 적어도 하나의 네트워크 인식된 오디오 커맨드는 단계(308)에 지시된 바와 같이, 대응하는 네트워크 신뢰값을 갖는다. 적어도 하나의 네트워크 음성 인식 엔진(226)은 동일한 오디오 커맨드를 터미널 음성 인식으로서 인식하고, 또한 터미널 음성 인식 엔진과는 독립적으로 오디오 커맨드를 인식한다.

일단 오디오 커맨드가 터미널 음성 인식 엔진(216)에 의해 인식되면, 적어도 하나의 터미널 인식된 오디오 커맨드는 접속(232)을 통해 비교기(228)에 제공된다. 또한, 일단 적어도 하나의 네트워크 음성 인식 엔진(226)이 오디오 커맨드를 인식하면, 적어도 하나의 네트워크 인식된 오디오 커맨드는 접속(230)을 통해 비교기(228)에 제공된다.

본 발명의 일 실시예에서, 비교기(228)는 단계(310)에 지시된 바와 같이, 적어도 하나의 터미널 신뢰값에 터미널 가중 팩터 만큼의 가중치를 부여하고, 적어도 하나의 네트워크 신뢰값에 네트워크 가중 팩터 만큼의 가중치를 부여한다. 예컨대, 비교기는 적어도 하나의 네트워크 음성 인식 엔진(226)의 인식 용량에 추종을 허용할 수 있고, 따라서, 조정, 즉, 네트워크 신뢰값을 스케일링 팩터에 의해 승산하여, 네트워크 신뢰값을 증가시키고, 또한, 조정, 즉, 터미널 신뢰값을 스케일링 팩터에 의해 승산하여 터미널 신뢰값을 감소시킨다.

게다가, 본 방법은 단계(312)에 지시된 바와 같이, 인식된 오디오 커맨드 신뢰값을 갖는 적어도 하나의 인식된 오디오 커맨드를 적어도 하나의 터미널 인식된 오디오 커맨드 및 적어도 하나의 네트워크 인식된 오디오 커맨드로부터 선택하는 단계를 제공한다. 구체적으로, 비교기(228)는 인식된 오디오 커맨드 신뢰값에 기초하여 복수의 인식된 오디오 커맨드를 선택한다. 본 발명의 일 실시예에서, 대화 관리자(234)는 비교기(228)에 N-최상 표시기를 제공하는데, 이 N-최상 표시기는 대화 관리자(234)에 제공하기 위한 인식된 오디오 커맨드의 수 N을 표시한다. 비교기(228)는 적어도 하나의 인식된 오디오 커맨드 및 적어도 하나의 네트워크 인식된 오디오 커맨드를 그 대응하는 신뢰값에 따라 분류하여, 그것으로부터 N-최상 커맨드를 추출한다.

본 발명의 일 실시예에서, 비교기(228)는 신뢰값에 대응하는 인식된 오디오 커맨드에 기초하여 적어도 하나의 터미널 인식된 오디오 커맨드 및 적어도 하나의 네트워크 인식된 오디오 커맨드를 필터링할 수 있다. 예컨대, 비교기는 인식된 오디오 커맨드 신뢰값이 비교되는 최소 신뢰값을 가질 수 있고, 최소 신뢰 레벨 이하의 신뢰값을 갖는 모든 인식된 오디오 커맨드는 제거된다. 이에 따라, 비교기는 대화 관리자에 N-최상 커맨드를 제공한다.

또한, 비교기는, 최소 신뢰 레벨 이상인 신뢰값을 갖는 커맨드가 N보다 적은 경우에, N보다 적은 커맨드를 대화 관리자에 제공할 수 있다. 비교기가 최소 신뢰 레벨 이상인 신뢰값을 갖는 임의의 인식된 커맨드를 수신하지 못하는 경우, 비교기는 에러 통지를 생성하고, 이 에러 통지는 접속(236)을 통해 대화 관리자에 제공된다. 그리고, 단계(314)에 지시된 바와 같이, 에러 통지는 적어도 하나의 터미널 신뢰값 및 적어도 하나의 네트워크 신뢰값이 최소 신뢰 레벨 이하인 때, 즉, 신뢰 레벨이 0.5이하인 때, 생성된다.

본 발명의 일 실시예에서, 단계(316)에 지시된 바와 같이, 대화 관리자는 적어도 하나의 인식된 오디오 커맨드를 검증하고, 검증되고 인식된 오디오 커맨드를 생성하고, 검증되고 인식된 오디오 커맨드에 기초하여 오퍼레이션을 실행한다. 예컨대, 대화 관리자는 N-최상 인식된 오디오 커맨드를, 음성 전달 인터페이스(212 및 214) 및 음성 합성 엔진(238)을 경유하여 스피커(208)를 통해 사용자에게 제공할 수 있다. 이에 따라, 사용자는 다음에 N-최상 커맨드 중 어느 것이 원래의 오디오 커맨드를 정확하게 반영하는지를 선택하고, 검증되고 인식된 오디오 커맨드를 생성한다.

검증되고 인식된 오디오 커맨드는 다음에, 원래의 오디오 커맨드가 제공되는 것과 동일한 방식으로 대화 관리자(234)에 다시 제공된다. 예컨대, N-최상 리스트의 4번째 인식된 오디오 커맨드가 적절한 커맨드라면, 사용자는 이 커맨드를 검증하고, 검증되고 인식된 오디오 커맨드를 생성하며, 사용자는 다음에, 마이크(206)으로 워드(4)를 말할 수 있고, 이것은 터미널 음성 인식 엔진(216) 및 적어도 하나의 네트워크 음성 인식 엔진(226) 모두에 제공되며, 또한 비교기(228)에도 제공되며, 비교기(228)에서는 대화 관리자(234)로 제공된다. 대화 관리자(234)는 검증되고 인식된 오디오 커맨드를 수신하자마자 이러한 검증되고 인식된 오디오 커맨드에 기초하여 오퍼레이션을 실행한다.

대화 관리자(234)는 적어도 하나의 인식된 오디오 커맨드, 또는 검증된 오디오 커맨드에 기초하여 복수의 오퍼레이션을 실행할 수 있다. 예컨대, 대화 관리자는 요청된 정보를 검색하기 위해, 상업용 데이터베이스와 같은 콘텐츠 서버(104)에 액세스할 수 있다. 또한, 대화 관리자는, 프로그램된 어플리케이션의 다음 단계로 진행하는 것과 같은, 프로그램내에서 오퍼레이션을 실행할 수 있다. 또한, 대화 관리자는 인식된 오디오 커맨드를 하나의 폼에 기입할 수 있고, 이에 따라 사용자로부터 폼에 대한 다음 엔트리 또는 입력을 요청한다. 당업자가 알 수 있는 바와 같이, 대화 관리자는 적어도 하나의 인식된 오디오 커맨드의 수신에 따라 임의의 적절한 오퍼레이션을 수행할 수 있다.

본 발명의 일 실시예에서, 단계(318)에 지시된 바와 같이, 대화 관리자는 적어도 하나의 인식된 오디오 커맨드를 수신하자마자, 적어도 하나의 인식된 오디오 커맨드 신뢰값에 기초하여 적어도 하나의 인식된 오디오 커맨드를 필터링하고, 최고로 인식된 오디오 커맨드 신뢰값을 갖는 인식된 오디오 커맨드에 기초하여 오퍼레이션을 실행한다. 예컨대, 대화 관리자는, 0.6이하와 같은, 소정의 설정값 이하인 신뢰값을 갖는 모든 인식된 오디오 커맨드를 제거할 수 있고, 나머지의 인식된 오디오 커맨드에 기초하여 오퍼레이션을 실행할 수 있다. 주지된 바와 같이, 대화 관리자는 적어도 하나의 인식된 오디오 커맨드에 응답하여 임의의 적절한 실행가능 오퍼레이션을 실행할 수 있다.

또한, 대화 관리자는 필터링에 기초하여, 비교기(236)에서 수행된 오퍼레이션과 유사하게, 소정의 신뢰 레벨 이하인 신뢰 값을 갖는 임의의 인식된 오디오 커맨드를 제거하길 시도할 수 있다. 예컨대, 대화 관리자는 비교기보다 더 높은 최소 신뢰 값을 설정할 수 있고, 이러한 최소 신뢰 레벨은 시스템(200)의 나머지와는 독립적으로 대화 관리자(234)에 의해 설정될 수 있다. 대화 관리자가 필터링후에, 대화 관리자 최소 신뢰 레벨이상인 임의의 인식된 오디오 커맨드를 포함하지 못하게 되는 경우에, 대화 관리자(234)는 비교기(228)와 유사하게 에러 통지를 생성하게 된다.

일단 에러 통지가 생성되면, 대화 관리자는 에러 커맨드(234)를 실행하고, 사용자(102)에게 오디오 커맨드가 적절하게 수신되지 않았다는 것을 알린다. 당업자가 알 수 있는 바와 같이, 대화 관리자는 비교기(228)에 의해 수행된 것과 같은 에러 통지를 생성하는 대신에 간단하게 에러 커맨드를 실행할 수 있다.

일단 대화 관리자가 오퍼레이션을 완전하게 실행하면, 멀티-레벨 분산 인식을 위한 방법은 단계(320)에 지시된 바와 같이 종료한다.

본 발명은 제1 음성 인식 엔진 및 적어도 하나의 제2 음성 인식 엔진을 통한 멀티-레벨 분산 음성 인식에 관한 것이다. 본 발명의 일 실시예에서, 제1 음성 인식 엔진은 터미널내에 배치되고, 적어도 하나의 제2 음성 인식 엔진은 네트워크내에 배치된다. 당업자가 알 수 있는 바와 같이, 음성 인식 엔진은 터미널, 네트워크 엘리먼트, 네트워크 엘리먼트에 동작가능하게 결합된 네트워크상의 별개의 서버내등에 배치될 수 있고, 음성 인식 엔진은 오디오 커맨드를 수신하고, 비교될 적어도 하나의 인식된 오디오 커맨드를 대화 관리자에 제공한다. 또한, 본 발명은, 제1 음성 인식 엔진에 제공되는 동일한 커맨드와는 독립적으로 제2 음성 인식 엔진에 오디오 커맨드를 제공하여 종래기술에 비해 개선된 점을 갖는다. 따라서, 제1 음성 인식 엔진의 인식 용량에 무관하게, 동일한 오디오 커맨드가 제2 음성 인식 엔진에 더 제공된다. 이에 따라, 본 발명은 비교기 및 수신하는 대화 관리자와 결합되는 복수의 음성 인식 엔진을 이용하여 음성 인식의 신뢰성을 향상시키고, 시스템 및 방법의 음성 인식의 정확도를 개량할 수 있게 한다.

당업자는 본 발명의 다른 변형 및 변경과 그 다양한 양태의 구현을 명확하게 알 수 있고, 본 발명은 여기 개시된 특정 실시예에 제한되는 것은 아니라는 것도 알 수 있다. 예컨대, 도 4의 비교기 및 대화 관리자는 네트워크 엘리먼트내에 상주하는 대신 네트워크 엘리먼트에 결합되는 서버상에 배치될 수 있다. 따라서, 임의의 모든 변형, 변경, 및 균등예가 여기서 클레임되고 개시된 본 발명의 근본 원리의 사상 및 범주를 벗어남이 없이 본 발명에 모두 포함될 수 있다는 것을 알 수 있다.

(57) 청구의 범위

청구항 1.

터미널 장치와 네트워크 장치 사이에서 멀티-레벨 분산 음성 인식(speech recognition)을 위한 방법에 있어서,

오디오 커맨드(audio command)를 상기 터미널 장치 내의 제1 음성 인식 엔진에 제공하는 단계;

상기 네트워크 장치 내의 적어도 하나의 제2 음성 인식 엔진에 상기 오디오 커맨드를 무선으로 제공하는 단계;

상기 제1 음성 인식 엔진 내에서 상기 오디오 커맨드를 인식하여, 적어도 하나의 제1의 인식된 오디오 커맨드를 생성하는 단계 - 상기 적어도 하나의 제1의 인식된 오디오 커맨드는 대응하는 제1 신뢰값(confidence value)을 가짐 -;

상기 제1 음성 인식 엔진에 의한 상기 오디오 커맨드의 인식과는 독립적으로, 상기 적어도 하나의 제2 음성 인식 엔진 내에서 상기 오디오 커맨드를 인식하여, 적어도 하나의 제2의 인식된 오디오 커맨드를 생성하는 단계 - 상기 적어도 하나의 제2의 인식된 오디오 커맨드는 대응하는 제2 신뢰값을 가짐 -;

상기 적어도 하나의 제1의 인식된 오디오 커맨드를 비교기에 무선으로 전송하는 단계;

상기 적어도 하나의 제2의 인식된 오디오 커맨드를 상기 비교기에 전송하는단계; 및

상기 적어도 하나의 제1 신뢰값 및 상기 적어도 하나의 제2 신뢰값에 기초하여, 상기 적어도 하나의 제1의 인식된 오디오 커맨드 및 상기 적어도 하나의 제2의 인식된 오디오 커맨드로부터, 인식된 오디오 커맨드 신뢰값을 갖는 적어도 하나의 인식된 오디오 커맨드를 선택하는 단계

를 포함하는 멀티-레벨 분산 음성 인식 방법.

청구항 2.

삭제

청구항 3.

제1항에 있어서,

상기 적어도 하나의 인식된 오디오 커맨드를 선택하기 전에,

상기 적어도 하나의 제1 신뢰값에 제1 가중 팩터 만큼의 가중치를 부여하고, 상기 적어도 하나의 제2 신뢰값에 제2 가중 팩터 만큼의 가중치를 부여하는 단계를 더 포함하는 멀티-레벨 분산 음성 인식 방법.

청구항 4.

제1항에 있어서,

상기 적어도 하나의 인식된 오디오 커맨드에 기초하여 적어도 하나의 오퍼레이션을 실행하는 단계를 더 포함하는 멀티-레벨 분산 음성 인식 방법.

청구항 5.

제1항에 있어서,

상기 적어도 하나의 인식된 오디오 커맨드를 검증하는 단계를 더 포함하는 멀티-레벨 분산 음성 인식 방법.

청구항 6.

제1항에 있어서,

상기 적어도 하나의 제1 신뢰값 및 상기 적어도 하나의 제2 신뢰값이 최소 신뢰 레벨이하일 때, 에러 통지를 생성하는 단계를 더 포함하는 멀티-레벨 분산 음성 인식 방법.

청구항 7.

터미널 장치와 네트워크 장치 사이에서 멀티-레벨 분산 음성 인식용 장치에 있어서,

오디오 커맨드를 수신하고, 적어도 하나의 터미널 인식된 오디오 커맨드를 생성하기 위해, 마이크로폰(microphone)에 동작가능하게 연결된 터미널 음성 인식 엔진 - 상기 적어도 하나의 터미널 인식된 오디오 커맨드는 대응하는 터미널 신뢰값을 가짐 -;

상기 오디오 커맨드를 상기 터미널 장치로부터 상기 네트워크 장치로의 무선 전송을 거쳐 수신하고, 상기 터미널 음성 인식 엔진과는 독립적으로, 적어도 하나의 네트워크 인식된 오디오 커맨드를 생성하기 위해, 상기 마이크로폰에 동작가능하게 연결된 적어도 하나의 네트워크 음성 인식 엔진 - 상기 적어도 하나의 네트워크 인식된 오디오 커맨드는 대응하는 네트워크 신뢰값을 가짐 -;

상기 적어도 하나의 터미널 인식된 오디오 커맨드를 무선 전송으로부터 수신하는 것이 동작가능할 수 있도록 상기 터미널 음성 인식 엔진에 동작가능하게 연결되고, 상기 적어도 하나의 네트워크 인식된 오디오 커맨드를 수신하도록 동작가능하게 결합된 비교기 - 상기 비교기는 상기 네트워크 장치 상에 배치됨(disposed) - 및

상기 비교기에 동작가능하게 연결된 대화 관리자(dialog manager)를 포함하고,

상기 비교기는 상기 적어도 하나의 터미널 신뢰값 및 상기 적어도 하나의 네트워크 신뢰값에 기초하여, 상기 적어도 하나의 터미널 인식된 오디오 커맨드 및 상기 적어도 하나의 네트워크 인식된 오디오 커맨드로부터, 인식된 신뢰값을 갖는 적어도 하나의 인식된 오디오 커맨드를 선택하고,

상기 선택된 적어도 하나의 인식된 오디오 커맨드는 상기 대화 관리자에 제공되는 멀티-레벨 분산 음성 인식용 장치.

청구항 8.

삭제

청구항 9.

제7항에 있어서,

상기 대화 관리자는 상기 적어도 하나의 인식된 오디오 커맨드 신뢰 레벨에 기초하여 상기 적어도 하나의 인식된 오디오 커맨드로부터 대화 관리자 오디오 커맨드를 결정하고,

상기 대화 관리자는 상기 대화 관리자 오디오 커맨드에 응답하여 오퍼레이션을 실행하는 멀티-레벨 분산 음성 인식용 장치.

청구항 10.

제9항에 있어서,

상기 대화 관리자는 콘텐츠 서버에 액세스하고, 상기 대화 관리자 오디오 커맨드에 응답하여 부호화된(encoded) 정보를 가져오는(retrieve) 멀티-레벨 분산 음성 인식용 장치.

청구항 11.

제10항에 있어서,

상기 대화 관리자에 동작가능하게 연결된 음성 합성 엔진을 더 포함하고,

상기 음성 합성 엔진은 상기 대화 관리자로부터 음성 부호화된 정보를 수신하고, 음성 포맷화된 정보를 생성하는 멀티-레벨 분산 음성 인식용 장치.

청구항 12.

제11항에 있어서,

상기 음성 합성 엔진에 오디오 서브시스템이 동작가능하게 연결되고,

상기 오디오 서브시스템은 상기 음성 포맷화된 정보를 수신하고, 출력 메시지를 제공하는 멀티-레벨 분산 음성 인식용 장치.

청구항 13.

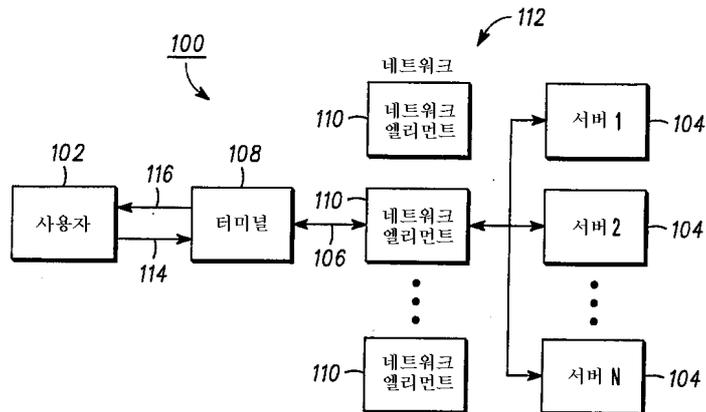
제12항에 있어서,

상기 비교기가 상기 대화 관리자에 예러 통지를 제공할 때, 상기 출력 메시지는 예러 서술문(statement)인 멀티-레벨 분산 음성 인식용 장치.

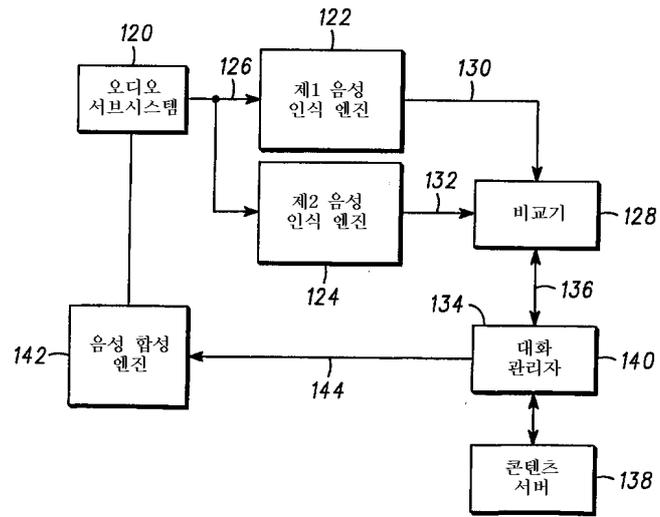
도면

도면1

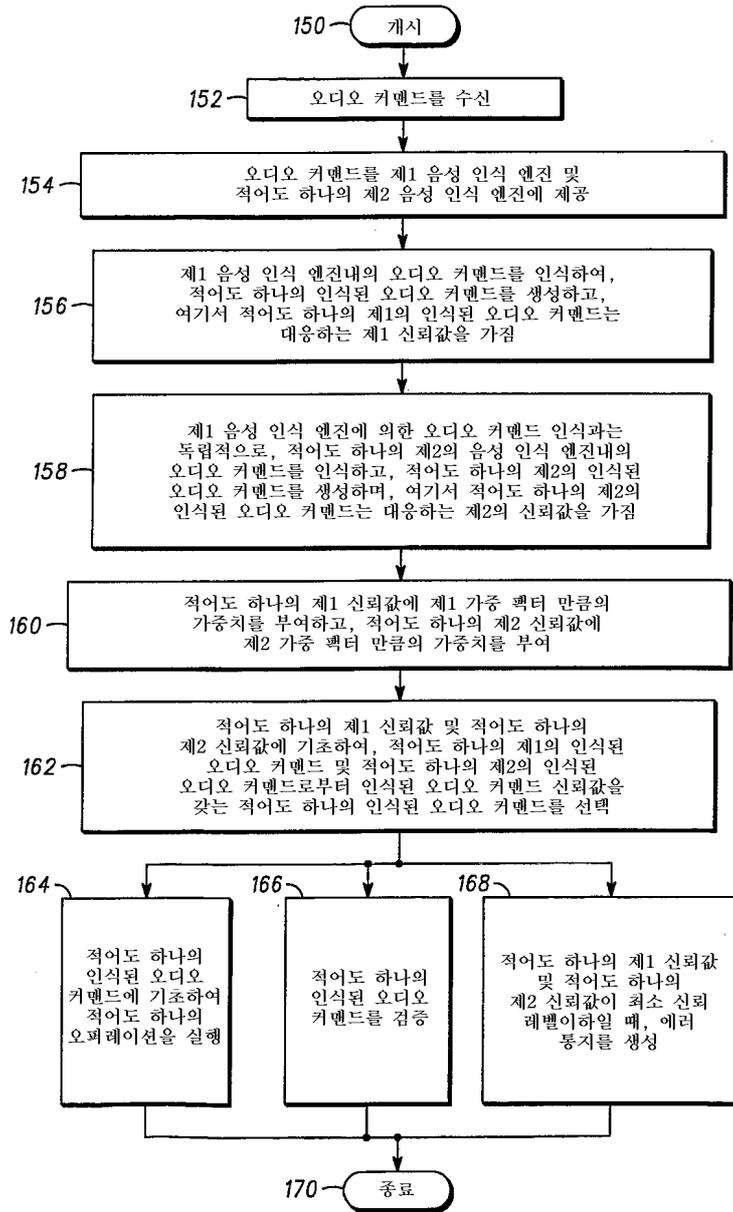
(종래 기술)



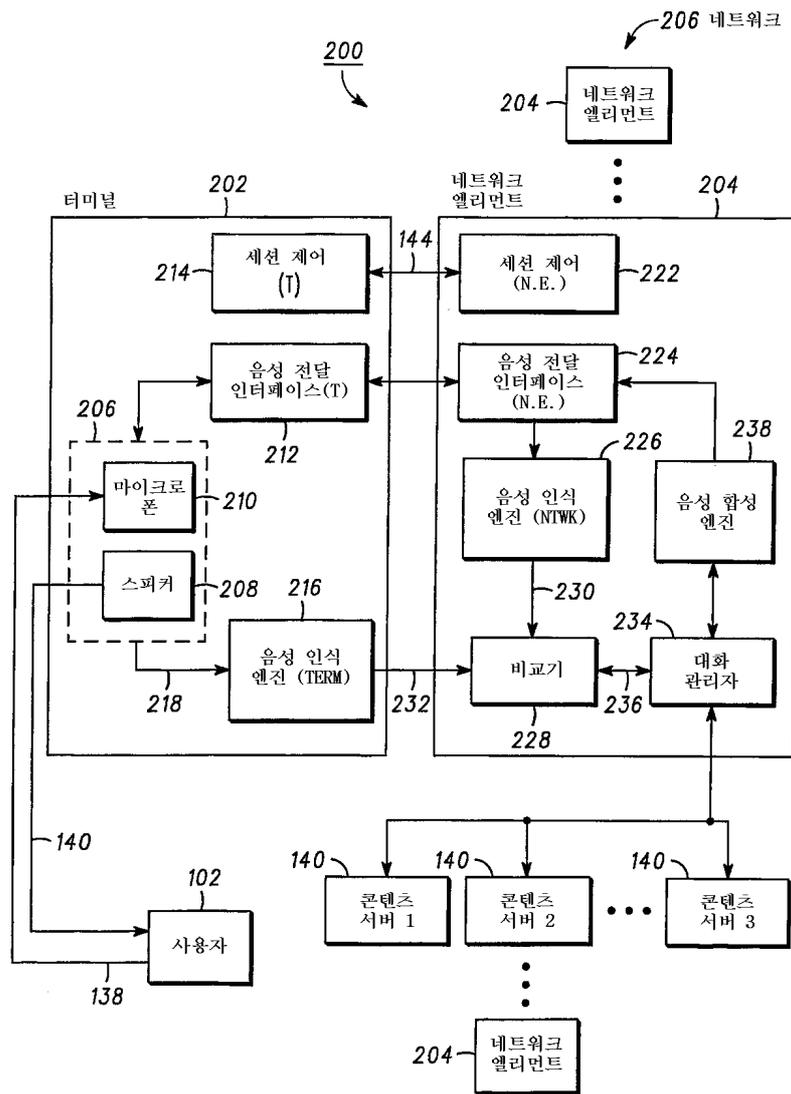
도면2



도면3



도면4



도면5

