(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.[7]
G06F 3/00

(11) 공개번호 특2000-0057082
(43) 공개일자 2000년09월15일

(21) 출원번호 10-1999-0060114
(22) 출원일자 1999년12월22일
(30) 우선권주장 9/219,970 1998년12월23일 미국(US)
(71) 출원인 휴렛트-팩카드 캄파니 디. 크레이그 노룬드
미합중국 캘리포니아주 (우편번호 94304) 팔로 알토 하노버 스트리트 3000
(72) 발명자 올리버토마스씨
미국콜로라도주80525반우드드라이브포트콜린스1742
(74) 대리인 김창세, 장성구

심사청구 : 없음

(54) 음성 제어 입력 방법

요약

동작을 제어하는 음성 명령을 수신하는 휴대용 핸드헬드 문서 스캐너 또는 디지탈 카메라와 같은 휴대용 캡쳐 장치(100)가 개시된다. "스캔", "저장", "삭제", "확대" 및 "송신"과 같은 명령은 예시를 위한 것이며, 여기서, 캡쳐 장치(100)는 명령 이름과 관련된 사전결정된 기능을 수행할 것이다. 일 실시예에서, 휴대용 캡쳐 장치(100)는 음성 분석 소프트웨어(124)를 통해 사용자의 구두 명령을 인식하기 위해 훈련된다. 음성 분석 소프트웨어(124)는 캡쳐 장치(100)내에 위치하거나 혹은, 호스트 컴퓨터 시스템(200)에 위치하여, 호스트 컴퓨터 시스템(200)에 연결된 동안 캡쳐 장치(100)에 의해 액세스될 수 있다. 캡쳐 장치(100)는 제어기(106)의 제어하에 오디오 입력/출력 시스템을 가진다. 음성 제어 입력 명령을 수신시에, 제어기(106)는 이 디지탈화된 음성 입력을 동적 메모리(118)에 저장한다. 그후, 제어기(106)는 이 수신된 명령을, 정적 메모리(116)에 유지되는 명령 인식 테이블(126)에 저장된 명령과 비교한다.

대표도

도1

명세서

도면의 간단한 설명

도 1은 본 발명의 캡쳐 장치의 블록도.

도 2는 본 발명의 캡쳐 장치와 통신하는 호스트 컴퓨터 시스템의 블록도.

도 3은 본 발명의 캡쳐 장치의 동작에 대한 음성 제어 입력의 전반적인 흐름을 도시한 흐름도.

도 4는 본 발명의 캡쳐 장치가 음성 제어 입력 명령을 처리하는 흐름도.

도 5는 본 발명의 캡쳐 장치가 명령을 실행하는 일반화된 흐름도.

도 6은 음성 제어 입력 명령을 인식하기 위해 본 발명의 캡쳐 장치를 훈련시키는 흐름도.

도면의 주요 부분에 대한 부호의 설명

100: 캡쳐 장치 102: 음성픽업 구성요소

110: 스피커 106: 제어기

116: 정적 메모리 118:동적 메모리

200: 호스트컴퓨터 시스템

발명의 상세한 설명

발명의 목적

발명이 속하는 기술분야 및 그 분야의 종래기술

본 발명은 핸드헬드 문서 스캐너(handheld document scanners) 또는 디지탈 카메라와 같은 휴대용 캡쳐 장치(portable capture devices)에 관한 것으로, 특히 휴대용 핸드헬드 문서 스캐너 또는 디지탈 카메라

에 대한 음성 제어 입력에 관한 것이다.

핸드헬드 문서 스캐너 또는 디지탈 카메라와 같은 휴대용 캡쳐 장치는 소정의 상황에서 상당히 유용한 수단인 것으로 알려져 있다. 이러한 캡쳐 장치의 주된 이점은 사용자의 사무실 또는 작업실로부터 멀리 떨어진 다양한 위치로부터의 정보를 포착 및 저장시에 그들의 휴대성 및 용이성에 있다.

이러한 휴대용 캡쳐 장치는 작지만, 보다 작고, 보다 휴대하기 편하도록 그들의 크기를 더욱 줄이는 것이 바람직하다. 그러나, 현재의 물리적 사용자 인터페이스 요건상 더 이상의 크기를 줄이는 일은 한계가 있다. 예를 들면, 대부분의 휴대용 핸드헬드 문서 스캐너는 어디서든지 사용자가 10 내지 15개의 사용자의 입력 버튼으로 다수의 상이한 동작을 제어할 수 있게 한다. 이러한 동작은 스캐닝 개시 및 중지; 스캔된 정보를 저장 및 삭제; 스캔된 정보를 송신; 스캔된 데이터를 스캐너 디스플레이상에 보여주고, 확대하고 작동시키는 것을 포함한다. 버튼은 사용자가 쉽게 제어 및 누룰 수 있도록 충분히 크고 적당한 간격을 두어야 한다. 또한, 문서를 스캔하기 위해 휴대용 스캐너를 사용하므로써 캡쳐 장치의 정상적인 전송 및 처리시에 구동되지 않는 각종 기능을 구동시키도록 버튼을 누룰 필요없이 휴대용 스캐너 장치를 다룰 수 있는 방식으로 버튼을 배치해야 한다. 버튼은 때때로 결합하여 사용되는 데, 이는 사용자가 스캐너 장치를 사용하기 불편하게 만든다. 사용자 입력 버튼이 차지하는 물리적 공간으로 인하여, 부득이 이러한 캡쳐 장치상의 출력 디스플레이는 종종 상당히 작을 수 있으므로, 원하는 것보다 적은 기능의 디스플레이를 사용하게 된다. 휴대용 디지탈 카메라의 경우에도 이와 동일 할 수 있다.

따라서, 휴대용 캡쳐 장치를 동작시키는 데 필요한 사용자 입력 버튼의 수를 감소키기고, 동시에 사용자 인터페이스의 복잡성을 감소시키는 개선된 방법 및 장치가 필요하다는 것은 명맥하다. 또한, 휴대용 캡쳐 장치의 휴대성을 크기하고 사용이 더욱 용이하도록 하기 위하여 크기를 감소시킬 필요가 있다. 본 기술분야에서는 휴대용 캡쳐 장치에 보다 크고 보다 판독가능한 디스플레이를 사용하면서 휴대용 캡쳐 장치의 감소된 전체 크기를 유지할 필요가 있다. 본 발명은 본 기술분야의 상기 및 다른 요구사항들을 만족시킨다.

발명이 이루고자 하는 기술적 과제

본 발명의 목적은 사용자 음성 입력을 사용하여 핸드헬드 문서 스캐너 또는 디지탈 카메라와 같은 휴대용 캡쳐 장치의 동작을 제어하는 데 있다.

본 발명의 다른 목적은 휴대용 캡쳐 장치상의 사용자 입력 버튼의 수를 감소시키려는 데 있다.

본 발명의 또다른 목적은 다수의 사용자 입력 버튼을 제거하여 휴대용 캡쳐 장치의 전체 크기를 감소시키려는 데 있다.

본 발명의 또다른 목적은 휴대용 캡쳐 장치의 출력 디스플레이 영역을 증가시키면서 제거된 다수의 사용자 입력 버튼이 이전에 차지한 소정의 물리적 영역을 활용하므로써 휴대용 캡쳐 장치의 전체 크기를 감소시키려는 데 있다.

본 발명의 또다른 목적은 휴대용 캡쳐 장치의 동작을 사용자가 이야기한 가청 패스워드(audible password)로 키잉화하려는 데 있다.

본 발명의 또다른 목적은 사용자의 음성 제어 입력 명령을 인식하도록 휴대용 캡쳐 장치를 훈련시키기 위하여 휴대용 캡쳐 장치를 호스트 컴퓨터로 연결(tether)시키려는 데 있다.

본 발명의 또다른 목적은 호스트 컴퓨터의 훈련이 필요없는 휴대용 캡쳐 장치에 제한된 음성 제어 입력 명령을 사용하려는 데 있다.

본 발명의 상기 및 다른 양상들은 음성 제어 입력 명령을 수신하여 그의 동작을 제어하는 휴대용 캡쳐 장치에서 이루어진다. 휴대용 핸드헬드 스캐너로써 스캔하는 것과 같이 휴대용 캡쳐 장치로써 동작을 개시하기 위하여, 사용자는 캡쳐 장치를 파워온하고 음성 제어 입력 명령 "스캔"을 입력하는 데, 그다음, 캡쳐 장치는 이 캡쳐 장치에 위치한 음성 픽업 구성요소를 통해 이 명령을 픽업한다. 캡처 장치는 명령 "스캔"을 인식시에 사용자가 문서상에 캡쳐 장치를 위치시키는 사전결정된 시간량, 주로 수 초를 기다린다. 시간 지연후에, 캡쳐 장치는 스캔할 준비가 되는 데, 이는 사용자에게 단어 "스캔"을 가청 반복하거나 혹은 발신음으로써 표시한다. 사용자는 문서의 표면을 가로질러 휴대용 핸드헬드 문서 스캐너를 이동시킨다. 사전결정된 시간주기동안 이동이 없다고 검출시에, 휴대용 핸드헬드 문서 스캐너는 "완료(done)" 또는 "중지(stop)"와 같은 또다른 가청 단어를 출력하거나 혹은 다시한번 발신음으로써 캡쳐 장치가 더 이상 스캔 모드가 아님을 사용자에게 알린다. 캡쳐 장치가 발신음 또는 가청 단어 출력으로부터 사전결정된 시간량내, 주로 수 초내에 더 이상의 이동이 없음을 검출한 경우, 휴대용 핸드헬드 문서 스캐너는 스캔 모드를 벗어나서, 휴대용 핸드헬드 문서 스캐너 디스플레이상으로 사용자에게 출력할 스캔 데이터의 처리를 시작한다. 본 발명의 다른 실시예에서, 사용자는 스캔 모드를 중지하기 위해 휴대용 핸드헬드 문서 스캐너상의 버튼은 누른다. 그후, 휴대용 핸드헬드 문서 스캐너는 사용자에게 출력하기 위한 스캔 데이터를 처리한다.

일단 이미지가 디스플레이로 출력되면, 사용자는 스캔된 이미지를 "저장" 또는 "삭제"하기 위한 음성 제어 입력 명령을 발생할 수 있다. 또한, 사용자는 "확대(zoom in)", "축소(zoom out)", "좌로", "우로", "위로" 또는 "아래로"와 같은 음성 제어 입력 명령을 발생하므로써 이미지의 상이한 부분들을 볼 수 있다. 사용자는 또한 "송신" 또는 "전체 송신"과 같은 음성 제어 입력 명령을 발생하므로써 설정된 연결부를 통해 호스트 컴퓨터로 스캔된 이미지 또는 몇몇 이미지를 전송할 수 있다. 일단 캡쳐 장치가 명령을 인식하게 되면 원하는 동작을 수행한다. 캡쳐 장치가 음성 제어 입력 명령을 처리하여 매칭되지 않음을 발견하는 경우, 사용자에게 매칭되지 않음을 가청 단어 또는 발신음 패턴과 같이 표시한다. 그후, 캡쳐 장치는 다음 음성 제어 입력 명령을 수신하기를 기다린다.

음성 제어 입력은 구두의 패스워드(a spoken password)를 사용하여 캡션장치를 위한 수단이 특정 사용자에게 "키(key)"가 되게 한다. 허용한다. 일단 캡쳐 장치가 파워온되면, 적절한 패스워드를 수신 및 처리

할 때 까지 동작하지 않을 것이다. 이것은 사용자의 패스워드가 도청당하지 않는 한 사용자외에는 누구도 이 캡쳐 장치를 사용하지 못하게 방지한다.

본 발명의 일 실시예에서, 사용자가 이야기한 음성 제어 입력 명령을 인식하도록 휴대용 핸드헬드 문서 스캐너와 같은 휴대용 캡쳐 장치를 음성 분석 소프트웨어를 통해 훈련시킨다. 음성 분석 소프트웨어는 캡쳐 장치내에 위치하거나 혹은, 호스트 컴퓨터 시스템에 위치하여 호스트 컴퓨터 시스템과 연결되어 있는 동안 캡쳐 장치가 이 소프트웨어를 액세스할 수 있다. 본 발명의 바람직한 실시예에서, 호스트 컴퓨터 시스템에 사용가능한 보다 큰 계산 능력을 이용하는 데, 그리고, 캡쳐 장치의 복잡도를 감소시키는 데 이 연결 모드(the tethered mode)를 사용할 수 있다.

예를 들면, 훈련 모드에서 음성 분석 소프트웨어를 사용시에, 사용자에게 음성 제어 입력 명령으로써 캡쳐 장치에 의해 실행될 수 있는 사전결정된 기능 리스트가 주어질 수 있다. 예를 들면, 명령 1은 문서 또는 이미지의 스캔 기능을 수행하기 위한 인스트럭션 집합을 나타낼 수 있다. 훈련 및 분석동안 명령 1을 선택하면, 음성 분석 소프트웨어는 스캔 기능을 위한 인스트럭션 집합을 야기하는 데 사용하기를 원하는 단어를 선택하도록 사용자에게 촉구할 수 있다. 그다음, 사용자는 선택한 단어를 다수번 반복하도록 촉구받는 다. 논리적 선택은 단어 "스캔"이지만, 사용자가 선택한 임의 단어가 사용될 수도 있다. 캡쳐 장치가 단어 "스캔"의 각 반복을 픽업하고, 음성 분석 소프트웨어가 이를 분석하여, "스캔" 명령을 발생시에 사용자 음성의 변동 및 억양을 포함하도록 인식 패턴을 전개(develope)한다. 사용자가 각종 기능을 야기하기 위해 선택한 모든 단어에 대한 인식 패턴은 캡쳐 장치의 정적 메모리에 명령 인식 테이블로 저장된다. 또한, 명령 인식 테이블에서 인식 패턴은 정적 메모리에 저장된 각종 기능을 위한 사전결정된 인스트럭션 집합과 각각 링크된다. 따라서, 캡쳐 장치가 구두 음성 제어 입력 명령어를 수신 및 인식할 때, 이 명령어와 관련된 인스트럭션 집합이 실행된다. 기능에 대한 인스트럭션 집합이 사용자의 단어 선택 및 후속되는 훈련과 그 단어 선택의 음성 분석과 결합(tie)되므로, 본 실시예는 음성 제어 입력 명령어를 위해 외국어가 사용될 수도 있도록 언어에 독립적이다.

본 발명의 또다른 실시예에서, 훈련 및 음성 분석을 위해 사용자에게 어떠한 단어 선택도 주어지지 않는다. 이 명령 인식 테이블에서 인식 패턴은 사전결정되며, 사용자가 사용해야 하는 특정 단어와 결합된다. 예를 들면, 캡쳐 장치가 사용자가 말한 명령을 인식할 때 까지, 사용자는 명령어 "스캔"의 그 또는 그녀의 발음을 변경시켜야 한다. 따라서, 본 실시예에서, 장치는 결과적인 동작을 표시하는 명령어가 있는 특정 언어를 주로 향하게 된다. 결과적인 동작을 표시하는 외국어를 사용하는 사용자를 위해 외국어 버전의 장치를 만들 수 있다.

휴대용 캡쳐 장치는 제어기의 제어하에 음성 오디오 입력/출력 시스템을 가진다. 제어기는 음성 제어 입력 명령을 수신시에 동적 메모리에 디지탈화된 음성 입력을 저장한다. 그후, 제어기는 이 명령을 처리하여, 이 명령에 대한 인식 패턴을 정적 메모리에 유지되는 명령 인식 테이블에 저장된 인식 패턴과 비교한다. 매칭이 발견되면, 인식 패턴과 결합된 인스트럭션 집합의 실행을 시작한다. 특정 명령에 대한 인스트럭션 집합은 가청 발신음, 명령 이름의 가청 재생 또는, 발광 다이오드(LED)을 출력하므로써 사용자에게 명령을 다시 확인응답하는 것을 포함한다. 또한, 사용자가 캡쳐 장치를 물리적으로 조작하거나 혹은 명령을 취소하는 데 시간을 허용하기 위해, 특정 명령은 인스트럭션 집합으로 구축된 하나 또는 그이상의 시간 지연을 가질 수 있다. 사용자가 방금 발생한 명령에 대해 그의 마음을 변경한 경우 혹은, 캡쳐 장치가 명령을 틀리게 해석한 경우, 사용자는 이 명령이 실행되기 전에 캡쳐 장치상의 취소 또는 클리어 버튼을 통해 혹은, 수신한 이전 명령을 취소하는 음성 제어 입력 명령을 통해 명령을 취소할 수 있다. 반면에, 명령을 취소하기 위한 입력이 없는 경우, 명령에 대한 인스트럭션 집합이 실행된다.

이를 테면, 디지탈 카메라 또는 문서 데이터 파일로써, 휴대용 스캐너로써 포착된 이미지 데이터 파일의 음성 주석을 가능하게 하는 휴대용 캡쳐 장치에 있어서, 캡쳐 장치는 음성 제어 입력 명령과 음성 주석을 구별한다. 본 발명의 일 실시예에서, 현 이미지 데이터 파일 또는 문서 데이터 파일에 바로 후속하여 음성 주석과 같은 음성 입력을 수신할 수 있도록 캡쳐 장치를 준비시키는 데 음성 제어 입력 주석 명령을 사용한다. 음성 주석이 종료했음을 표시하는 데 음성 입력없이 사전결정된 시간 길이의 침묵을 사용한다. 본 발명의 다른 실시예에서, 테이프 레코더상의 레코드 버튼과 유사한 사용 예를 활용한다. 후속된 음성 입력이 주석용임을 의미하기 위해 캡쳐 장치상의 버튼을 누르고, 누른채로 유지하는 데, 이것은 명령이 아니다. 일단 음성 주석이 완료되면, 사용자는 버튼을 해제하고, 캡쳐 장치는 포착한 음성 주석을 처리하여 이를 현 이미지 데이터 파일 또는 문서 데이터 파일로 연결시킨다.

발명의 구성 및 작용

본 발명의 상기 및 다른 양상, 특징 및 장점들은 첨부도면을 참조한 본 발명의 후속된 상세한 설명으로부터 보다 명백해질 것이다.

후속되는 설명은 본 발명을 수행할시에 고려되는 최선의 형태이다. 본 설명은 본 발명의 일반적 원리를 기술하기 위한 것일 뿐 이로 제한하려는 것은 아니다. 본 발명의 범주는 첨부된 특허청구의 범위를 기준으로 결정되어야 한다.

도 1은 본 발명의 캡쳐 장치의 블록도이다. 이제 도 1을 참조하면, 캡쳐 장치(100)는 이 캡쳐 장치상의 몇몇 제어 버튼(120)중의 하나인 파워온 버튼을 누르므로써 파워온된다. 캡쳐 장치(100)는 (도 1에는 도시되지 않은) 내부 배터리로부터 전력을 수신하거나 혹은, 캡쳐 장치(100)에 연결되고 (또한, 도 1에는 도시되지 않은) 전원에 플러그인된 전력 캐이블을 통하여 전력을 수신한다. 캡쳐 장치(100)를 제어하기 위한 음성 제어 입력 명령은 사용자가 음성 픽업 구성요소(102)가 픽업하기에 충분히 근접하여 이야기하므로써 주어진다. 음성 픽업 구성요소(102)는 사용자의 음성을 아날로그 신호로 변환한다. 음성 픽업 구성요소(102)에 아날로그-디지탈 변환기(104)가 연결되고, 이 변환기는 음성 픽업 구성요소(102)가 발생한 아날로그 신호를 디지탈 신호로 변환한다. 이 아날로그-디지탈 변환기(104)는 이 디지탈 신호를 제어기(106)로 송신하고, 제어기(106)는 이 신호를 이 제어기에 연결된 동적 메모리(118)에 저장한다. 그후, 본 발명의 바람직한 실시예에서, 제어기(106)는 정적 메모리(116)에 저장된 음성 분석 소프트웨어(124)를 호출하여, 동적 메모리(118)에 저장된 디지탈 신호상에 일련의 주파수 영역 변환을 수행한다. 음성 분석

소프트웨어(124)는 스페트랄 변환인 인식 패턴을 발생하고, 정적 메모리(116)의 명령 인식 테이블(126)에 저장된 명령에 대한 인식 패턴(또한, 스펙트랄 변환)과 비교한다. 당업자라면 본 발명에 스펙트랄 변환 대신에 음성 패턴을 인식하는 데에 다른 적당한 방법이 사용될 수 있음을 알 것이다.

매칭이 발견되면, 제어기(106)는 명령에 대한 인식 패턴과 링크된 명령 인식 테이블(126)의 인스트럭션 집합을 액세스한다. 예를 들면, 사용자가 문서를 스캔하기 위해 음성 제어 입력 명령을 말한 후에 캡쳐 장치(100)를 이동시켜, 이미지 픽업 구성요소(112)가 문서 표면의 일부 또는 전부를 접촉하도록 한다. 이미지 픽업 구성요소(112)는 문서의 표면으로부터 샘플 지점을 광학적으로 판독하고, 샘플링된 각 지점에 대한 명암 단계를 발생한다. 제어기(106)는 샘플 지점에 대한 명암 단계 값을 수신하고, 이들을 이미지 배열로 집합시킨다. 그 결과는 제어기(106)에 연결된 디스플레이(114)로 출력될 수 있는 데, 스캔된 문서의 표면을 시각적으로 표현한 것을 보여준다. 또한, 제어기(106)는 명암 단계 값을 디스플레이를 위한 또는 저장을 위한 이진 형태로 변환할 수 있다. 명암 단계 또는 이진 형태인 이미지 배열은 제어기(106)로부터 정적 메모리(116)에 문서 데이터 파일로서 저장된다.

문서를 스캐닝한 후에, 사용자는 음성 픽업 소자(102)에게 말하므로써, 설명적인 이야기 또는 사용자가 사용가능할 것으로 간주되는 다른 정보를 문서 데이터 파일에 음성 주석을 달 수 있다. 본 발명의 일 실시예에서 일반적으로 하나 또는 두 단어인 음성 제어 입력 명령으로부터 연장된 시간 주기에 걸쳐 완전히 연속된 음성 입력 열인 음성 주석을 구별하기 위하여, 사용자는 말하기 전에 몇몇 제어 버튼(120)중의 하나를 눌러 유지시키는 데, 이는 후속되는 음성 입력 열이 주석이며 명령이 아님을 가리킨다. 사용자가 음성 주석을 종료한 후에, 사용자는 제어 버튼(120)을 해제하고, 버튼 업 입력을 제어기(106)로 송신하여 음성 입력 열의 종료를 표시한다. 포착된 음성 입력 열은 음성 주석 파일로서 정적 메모리(116)에 저장되고, 스캔되어 정적 메모리(116)에 저장되어 있는 문서 데이터 파일에 연결된다.

본 발명의 다른 실시예에 있어서, 음성 제어 입력 명령중의 하나는 음성 주석 명령이다. 음성 제어 입력 주석 명령을 발생한 후에, 후속되는 음성 입력 열은 주석용으로 포착되어, 음성 주석 파일로서 저장되고, 캡쳐 장치에 포착되어 저장되어 있는 이미지 데이터 파일 또는 문서 데이터 파일에 연결된다. 사용자가 5 내지 10 초 사이와 같이 사전결정된 시간 주기 이상 동안 말하기를 중단할 때, 장치는 사전결정된 주기의 음성 무 입력을 음성 입력 열의 종료를 특징짓는 것으로 해석한다.

음성 제어 입력 주석 명령을 수신 및 인식시에, 혹은 후속되는 음성 입력이 주석용이라는 것을 표시하도록 몇몇 제어 버튼(120)중의 하나를 눌러서 유지할 시에, 음성 픽업 구성요소(102)는 사용자로부터의 음성 입력을 포착하고, 이 입력은 아날로그 신호로 변환된다. 아날로그-디지탈 변환기(104)는 음성 픽업 소자(102)에 의해 발생된 아날로그 신호를 디지탈 신호로 변환시킨다. 이 디지탈 신호는 제어기(106)로 송신된다. 제어기(106)는 음성 주석 디지탈 신호를 개별 음성 주석으로서 정적 메모리(116)에 저장하고, 이 음성 주석 파일을 이미지 데이터 파일 또는 문서 데이터 파일과 연결시킨다.

사용자는 문서 스캐닝 장치(100)가 음성 주석 파일을 재생하도록 요청할 수 있다. 제어기(106)는 요청받은 음성 주석 파일을 정적 메모리(116)로부터 검색하고, 이를 디지탈-아날로그 변환기(108)로 제공하고, 이 변환기는 디지탈 신호를 아날로그 신호로 변환한 뒤에 이 아날로그 신호를 스피커(110)로 공급하므로써, 이 스피커(110)는 오디오 출력을 발생한다. 또한, 명령에 대한 특정한 인스트럭션 집합은 디지탈-아날로그 변환기(108) 및 스피커(110)를 활용하여 혹은 (도 1에 도시되지 않은) LED를 비추므로써 명령의 수신을 확인응답하기 위해 사용자에게 가청 출력을 송신할 수 있다.

이미지 데이터 파일 또는 문서 데이터 파일 및 연결된 음성 주석 파일은 제어기(106)에 연결된 호스트 연결부(122)를 통해 호스트 컴퓨터 시스템(200)(도 2)와 같은 또다른 장치에 복사될 수 있다.

도 2는 본 발명과 관련된 호스트 컴퓨터 시스템의 블록도이다. 이제 도 2를 참조하면, 호스트 컴퓨터 시스템(200)은 처리 소자(202)를 포함한다. 처리 소자(202)는 시스템 버스(204)를 통해 호스트 컴퓨터 시스템(200)의 다른 소자와 통신한다. 키보드(206)는 사용자가 호스트 컴퓨터 시스템(200)으로 정보를 입력할 수 있게 하고, 그래픽 디스플레이(210)는 호스트 컴퓨터 시스템(200)이 사용자에게 정보를 출력할 수 있게 한다. 또한, 마우스(208)는 정보를 입력시에 사용되고, 저장 장치(212)는 호스트 컴퓨터 시스템(200)내에 데이터 및 프로그램을 저장하는 데 사용된다. 시스템 버스(204)에 연결된 통신 인터페이스(214)는 캡쳐 장치(100)(도 1)로부터 정보를 수신한다. 시스템 버스(204)에 연결된 스피커/사운드 카드(216)는 소정 오디오 정보를 사용자에게 출력한다. 소정 호스트 컴퓨터 시스템은 사운드 카드를 가질 수도 있으며, 이 경우에 스피커는 소프트웨어에 의해서만 구동된다. 또한, 시스템 버스(204)에 부착된 메모리(218)는 운영체제(220), 파일 전송 소프트웨어(222), 음성 분석 소프트웨어(224), 사용자 인터페이스 프로그램(226) 및 오디오 파일 변환 소프트웨어(228)를 포함한다.

파일 전송 소프트웨어(222)는 통신 인터페이스(214) 및 시스템 버스(204)를 통해 캡쳐 장치(100)의 호스트 연결부(122)(도 1)로부터 전송되는 이미지 데이터 파일 또는 문서 데이터 파일 및 연결된 음성 주석 파일을 수신하여, 이들을 저장 장치(212)에 저장한다. 사용자가 사용자 인터페이스 프로그램(226)을 액세스하고 음성 주석 파일을 가진 문서 데이터 파일 또는 이미지 데이터 파일을 선택할 때, 오디오 파일 변환 소프트웨어(228)는 압축해제하여 음성 주석 파일은 스피커/사운드 카드(216)가 인식할 만한 오디오 파일 포맷으로 변환시킨다. 스피커/사운드 카드(216)는 사용자에게 오디오 정보를 출력한다. 사용자는 오디오 정보를 들은 후에 이미지 데이터 파일 또는 문서 데이터 파일을 검토하기 위해 선택할 수 있다. 만약 그런 경우, 사용자 인터페이스 프로그램(226)은 중지되고, 이미지 데이터 파일 또는 문서 데이터 파일과 관련된 응용 프로그램이 호출되고, 그리고, 이 파일은 그래픽 디스플레이(210)에 디스플레이된다.

본 발명의 바람직한 언어 독립적 실시예에 있어서, 사용자는 임의 언어로 된 사용자의 음성 제어 입력 명령을 인식하도록 캡쳐 장치(100)를 훈련시키는 데 캡쳐 장치(100)에 위치한 음성 분석 소프트웨어(224)를 사용한다. 캡쳐 장치(100)는 먼저 보다 큰 계산 능력을 이용하기 위해 호스트 컴퓨터 시스템(200)에 연결된다. 그러면, 사용자는 음성 분석 소프트웨어(224)를 액세스하고, 스캔 기능과 같은 훈련시킬 기능을 야기하기 위해 사용자가 선택한 단어로 표현되는 특정 기능을 선택한다. 그후, 사용자는 스캔 기능을 나타내기 위해 선택한 단어를 다수번 반복한다. 사용자에 의해 가장 선택될 것으로 보이는 단어는 사용자가 말하는 언어에 상관없는 단어, 즉, 스캔 기능과 동격이거나 혹은 가장 가까운 단어이다. 영어권 사용

자의 경우, 선택되는 가장 흔한 단어는 "스캔"일 것이다. 픽업 소자(102)(도 1)가 포착한 사용자가 반복하는 단어"스캔"은 캡쳐 장치(100)에서 신호로 처리되어, 호스트 연결부(122)를 통해 통신 인터페이스(214)로 전송된다. 통신 인터페이스(214)는 각 신호를 시스템 버스(204)를 통해 메모리(218)로 전송하고, 여기서, 음성 분석 소프트웨어(224)는 "스캔" 명령을 발생시에 사용자의 음성의 변동 및 억양을 포함하도록 각 샘플 신호를 근거로 인식 패턴을 전개한다. 이 처리는 캡쳐 장치(100)에 대한 음성 입력 제어 명령으로써 야기될 수 있는 각 기능에 대해 반복된다. 그다음, 훈련을 위해 선택된 모든 단어에 대해 설정된 인식 패턴은 호스트 컴퓨터 시스템(200)으로부터 캡쳐 장치(100)로 다운로드되고, 후속된 제어 동작시에 사용하기 위해 정적 메모리(116)(도 1)의 명령 인식 테이블(126)에 저장된다.

도 3은 캡쳐 장치의 음성 제어 입력에 대한 동작의 전체 흐름을 도시한 흐름도이다. 이제 도 3을 참조하면, 단계(300)에서, 캡쳐 장치(100)(도 1)는 파워온된다. 단계(302)에서, 음성 픽업 구성요소(102)(도 1)가 본 발명의 바람직한 실시예에서 패스워드인 제1 음성 제어 입력 명령을 수신한다. 단계(304)에서, 도 4를 호출하여 음성 제어 입력 명령을 처리한다. 도 4로부터 복귀한 후에, 단계(306)에서, 도 4의 단계(408)에서 수행된 비교시에 단계(302)에서 수신된 음성 제어 입력 명령의 인식 패턴과 명령 인식 테이블(126)에 저장된 소정의 인식 패턴사이에 매칭이 발견되었는 지의 여부를 판정한다. 매칭이 발견되지 않은 경우, 단계(310)는 매칭되지 않음을 사용자에게 가청 단어 또는 특정한 발신음 패턴으로써 표시한다. 그후, 제어는 단계(302)로 복귀하고, 여기서, 캡쳐 장치(100)는 다음 음성 제어 입력 명령을 기다린다.

단계(306)에서, 도 4의 단계(408)에서 매칭이 발견되었다고 결정되면, 제어는 단계(308)로 넘어가고, 도 5를 호출하여 명령과 관련된 인스트럭션 집합을 실행한다. 도 5로부터 복귀시에, 단계(312)는 다음 음성 제어 입력 명령이 수신되었는 지 혹은 전원이 오프되었는 지의 여부를 판정한다. 다음 명령이 수신된 경우, 제어는 단계(302)로 복귀한다. 전원이 오프되면, 캡쳐 장치(100)의 동작은 종료된다.

도 4는 본 발명의 캡쳐 장치가 음성 제어 입력 명령을 처리하는 흐름도이다. 이제 도 4를 참조하면, 단계(400)에서, 음성 픽업 구성요소(102)(도 1)는 포착한 음성 제어 입력 명령을 아날로그 신호로서 출력한다. 단계(402)에서, 아날로그-디지탈 변환기(104)(도 1)는 아날로그 신호를 입력으로 수신하고, 이 아날로그 신호를 디지탈 신호로 변환하고, 디지탈 신호를 제어기(106)(도 1)로 출력한다. 단계(404)에서, 제어기(106)는 디지탈 신호를 입력으로 수신하고, 이 디지탈 신호를 동적 메모리(118)(도 1)에 저장한다. 단계(406)에서, 제어기(106)는 음성 분석 소프트웨어(120)를 호출하여 단계(404)에서 저장된 디지탈 신호상에 주파수 영역 변환을 수행하므로써 인식 패턴을 생성한다. 단계(408)에서, 제어기(106)는 단계(406)로부터의 인식 패턴을, 정적 메모리(116)(도 1)에 유지되는 명령 인식 테이블(126)에 저장된 음성 제어 입력에 대한 인식 패턴과 비교한다. 제어는 도 3으로 반환된다.

도 5는 본 발명의 캡쳐 장치가 명령을 실행하는 일반화된 흐름도이다. 당업자라면 단계의 순서는 특정 명령과 관련된 원하는 동작에 따라 상당히 변할 수 있다는 것을 알 수 있을 것이다. 이제 도 5를 참조하면, 단계(500)는 단계(302)에서 수신한 음성 제어 입력 명령과 매칭되는 인식 패턴으로 링크된 인스트럭션 집합을 액세스한다. 단계(502)는 인스트럭션 집합이 시간 지연 인스트럭션으로 시작하는 지의 여부를 판정한다. 만약 그러하다면, 단계(504)에서 시간 지연 인스트럭션을 실행하고, 시간 지연에서 명시된 시간량만큼 경과할 때 까지 인스트럭션 집합에서 나머지 인스트럭션의 더 이상의 실행을 중지한다. 단계(504)의 시간 지연 후 또는 단계(502)가 시간 지연 인스트럭션이 없다고 판정시에, 제어는 단계(506)로 넘어간다.

단계(506)에서, 인스트럭션 집합에서 다음 인스트럭션이 명령의 확인응답의 출력을 요구하는 지의 여부를 판정한다. 그러하다면, 단계(508)에서, 확인응답 인스트럭션을 실행한다. 특정한 명령에 따라, 가청 발신음, 수신한 음성 제어 입력 명령의 음성 재생, LED를 비추거나 혹은 다른 적절한 수단의 형태로 확인응답을 할 수 있다.

단계(510)에서, 인스트럭션 집합에서 다음 인스트럭션이 나머지 인스트럭션의 더 이상의 실행 전에 사용자에게 확인(confirmation) 입력을 요구하는 지의 여부를 판정한다. 삭제 명령과 같은 소정 명령은 유용한 데이터의 부주의한 파괴를 방지하기 위한 안전 예방조치로서 확인을 요구할 수 있다. 단계(510)에서 확인이 필요한 경우, 단계(512)는 사용자로부터 적절한 확인 입력을 수신하였는 지의 여부를 판정한다. 사용자는 확인하기 위해서는 특정한 명령을 근거로 몇몇 제어 버튼(120)중의 하나를 눌러야 한다. 그렇지 않으면, 사용자는 확인으로 또다른 음성 제어 입력 명령을 발생해야만 할 수도 있다. 단계(512)에서 적절한 확인 입력을 수신하지 못한 경우, 또는 수신한 입력이 없는 경우, 제어는 도 3의 단계(302)로 반환되고, 다음 음성 제어 입력 명령을 기다린다. 단계(512)에서 적절한 확인 입력을 수신한 경우 혹은, 단계(510)에서 확인 인스트럭션이 없다고 판정한 경우, 제어는 단계(514)로 넘어간다.

단계(514)는 명령을 취소하기 위한 입력을 수신했는 지의 여부를 판정한다. 단계(514)에서 취소 입력을 수신한 경우, 제어는 도 3의 단계(302)로 반환되어 다음 음성 제어 입력 명령을 기다린다. 단계(514)에서 취소 입력을 수신하지 못한 경우, 제어는 단계(516)로 넘어가서, 명령에 대한 인스트럭션 집합에서 나머지 인스트럭션을 실행한다. 인스트럭션 집합에서 마지막 인스트럭션을 실행시에, 제어는 도 3의 단계(312)로 반환된다.

도 6은 사용자 음성 제어 입력 명령을 인식하기 위해 본 발명의 캡쳐 장치를 훈련시키는 흐름도를 도시한다. 이제 도 6을 참조하면, 단계(600)에서, 음성 분석 소프트웨어(224)를 호스트 컴퓨터 시스템(200)(도 2)의 메모리(218)로 로딩한다. 단계(602)에서 캡쳐 장치(100)를 파워온한다. 단계(604)에서, 캡쳐 장치(100)(도 1)를 호스트 컴퓨터 시스템(200)에 연결한다. 이것은 캐이블, 적외선빔 또는 소정의 다른 적당한 연결부를 통해 가능하다. 단계(606)에서, 음성 분석 소프트웨어(224)는 사용자로부터의 입력을 수신하여 기능을 야기시키기 위한 명령어의 음성 분석 및 훈련을 위한 제1 기능을 선택한다. 그다음 단계(608)에서, 음성 분석 소프트웨어(224)는 캡쳐 장치(100)의 음성 픽업 구성요소(102)(도 1)로 제1 기능을 야기시키기 위해 사용자가 선택한 명령어를 다수번 가청 반복하도록 사용자에게 촉구한다. 단계(610)에서, 캡쳐 장치(100)는 음성 픽업 구성요소(102)에 의해 포착된 명령어의 다수의 음성 입력을 디지탈 신호로 처리하여, 호스트 컴퓨터 시스템(200)의 음성 분석 소프트웨어(224)로 송신한다. 단계(612)에서, 음성 분석 소프트웨어(224)는 단계(610)에서 수신된 다수의 디지탈 신호를 분석하고, 명령어에 대한 인식

패턴을 전개한다. 단계(614)에서, 단계(612)의 인식 패턴을 메모리(218)에 저장한다.

단계(616)는 사용자가 다음 기능을 야기하기 위한 명령어의 음성 분석 및 훈련을 위해 다음 기능을 선택했는 지의 여부 혹은, 사용자가 훈련 및 음성 분석을 위한 선택 기능을 완료했다는 표시를 수신하였는 지의 여부를 판정한다. 사용자가 선택 기능을 완료했다는 표시를 수신한 경우, 단계(618)에서, 음성 분석 소프트웨어(224)는 단계(612)에서 결정되어 단계(614)에 저장된 모든 인식 패턴을 단계(602)에서 설정된 연결부를 통해 캡쳐 장치(100)로 전송한다. 단계(620)에서, 단계(618)에서 전송된 인식 패턴은 정적 메모리(116)의 명령 인식 테이블(126)에 저장되고, 각 기능에 대한 인식 패턴은 명령 인식 테이블(126)에 또한 저장된 인스트럭션 집합으로 링크되고, 이 인스트럭션 집합은 인식 패턴으로 처리시에 단계(612)에서 결정된 인식 패턴중의 하나와 매칭되는 음성 제어 입력 명령을 수신시에 실행될 것이다. 단계(620)후에, 캡쳐 장치(100)의 훈련 및 음성 분석은 종료된다.

발명의 효과

따라서, 본발명의 바람직한 실시예를 기술하였지만, 당업자라면 특허청구의 범위에 정의되어진 본 발명의 범주를 벗어나지 않으면서 구성에서의 다수의 변경, 상이한 실시예 및 본 발명의 응용이 있을 수 있음을 알 것이다. 본 명세서에 개시된 설명은 본 발명을 예시하기 위한 것일 뿐 본 발명을 제한하려는 것은 아니며, 후속되는 특허청구의 범위내에서 정의된다.

(57) 청구의 범위

청구항 1

캡쳐 장치(100)를 위한 음성 제어 입력 방법에 있어서:

(가) 상기 캡쳐 장치(100)에서 음성 픽업 구성요소(102)로써 제1 음성 제어 입력 명령을 포착하는 단계와;

(나) 상기 제1 음성 제어 입력 명령을 제1 아날로그 신호로 변환하는 단계와;

(다) 상기 제1 아날로그 신호를 제1 디지탈 신호로 변환하는 단계와;

(라) 상기 제1 디지탈 신호를 제1 인식 패턴으로 변환하는 단계와;

(마) 상기 제1 인식 패턴을, 상기 캡쳐 장치(100)에서 정적 메모리(116)의 명령 인식 테이블(126)에 저장된 적어도 하나의 인식 패턴과 비교하는 단계와;

(바) 상기 제1 인식 패턴이 상기 명령 인식 테이블(126)에 저장된 상기 적어도 하나의 인식 패턴과 매칭하는 경우(306), 상기 적어도 하나의 인식 패턴에 링크된 제1 인스트럭션 집합을 실행하는 단계

를 포함하는 음성 제어 입력 방법.

청구항 2

제 1 항에 있어서,

상기 캡쳐 장치(100)는 스캐너 장치인 음성 제어 입력 방법.

청구항 3

제 1 항에 있어서,

상기 캡쳐 장치(100)는 디지탈 카메라인 음성 제어 입력 방법.

청구항 4

제1 항에 있어서,

상기 단계(나)는,

(나1) 상기 제1 아날로그 신호를 상기 캡쳐 장치(100)의 아날로그-디지탈 변환기(104)로 입력하는 단계를 더 포함하고,

상기 단계(다)는,

(다1) 상기 아날로그-디지탈 변환기(104)에서 상기 제1 아날로그 신호를 상기 제1 디지탈 신호로 변환시키는 단계와;

(다2) 상기 아날로그-디지탈 변환기(104)로부터의 상기 제1 디지탈 신호를 상기 캡쳐 장치(100)의 제어기(106)로 전송하는 단계와;

(다3) 상기 제어기(106)가 상기 제1 디지탈 신호를 상기 캡쳐 장치(100)의 동적 메모리에 저장하는(404) 단계를 더 포함하는

음성 제어 입력 방법.

청구항 5

제 1 항에 있어서,

상기 단계(라)는,

(라1) 상기 캡쳐 장치(100)의 동적 메모리(118)에 저장된 상기 제1 디지탈 신호상에 다수의 주파수 영역 변환을 수행하고(406), 상기 제1 인식 패턴을 발생하는 단계를 더 포함하고, 상기 제1 인식 패턴은 상기 제1 디지탈 신호의 스펙트랄 변환인 음성 제어 입력 방법.

청구항 6

제1 항에 있어서,

상기 단계(가)는, 상기 단계(가) 이전에,

(가1) 음성 분석 소프트웨어(224)를 호스트 컴퓨터 시스템(200)의 메모리(218)로 로딩하는(600) 단계와;

(가2) 상기 캡쳐 장치(100)와 상기 호스트 컴퓨터 시스템(200)을 연결시키는(604) 단계와;

(가3) 상기 음성 분석 소프트웨어(224)로써, 사전결정된 기능을 야기하기 위한 적어도 하나의 단어를 훈련 및 음성 분석하기 위하여 상기 사전결정된 기능을 선택하는(606) 단계와;

(가4) 상기 캡쳐 장치(100)의 상기 음성 픽업 구성요소(102)에서 상기 적어도 하나의 단어인 다수의 음성 입력을 포착하는(610) 단계와;

(가5) 상기 캡쳐 장치(100)에서 상기 다수의 음성 입력을 다수의 디지탈 신호로 처리하는(610) 단계와;

(가6) 상기 캡쳐 장치(100)로부터의 상기 다수의 디지탈 신호를 상기 호스트 컴퓨터 시스템(200)으로 송신하는(610) 단계와;

(가7) 상기 음성 분석 소프트웨어(224)로써 상기 다수의 디지탈 신호를 분석하는(612) 단계와;

(가8) 상기 음성 분석 소프트웨어(224)로써 상기 다수의 디지탈 신호의 상기 분석으로부터 상기 적어도 하나의 인식 패턴을 전개하는(612) 단계와;

(가9) 상기 적어도 하나의 인식 패턴을 상기 호스트 컴퓨터 시스템(200)의 상기 메모리(218)에 저장하는(614) 단계와;

(가10) 상기 호스트 컴퓨터 시스템(200)에서 상기 메모리(218)의 상기 적어도 하나의 인식 패턴을, 상기 캡쳐 장치(100)에서 상기 정적 메모리(116)의 상기 명령 인식 테이블로 전송하는(618) 단계로서, 상기 적어도 하나의 인식 패턴은 상기 사전결정된 기능을 수행하기 위해 상기 명령 인식 테이블(126)에 저장된 상기 제1 인스트럭션 집합으로 링크되는 상기 단계와;

(가11) 다수의 사전결정된 기능에 대해 단계(가3) 내지 단계(가10)를 반복하는 단계로서, 상기 적어도 하나인 다수의 단어에 대한 상기 다수의 음성 입력으로부터 다수의 인식 패턴이 전개되고, 상기 다수의 인식 패턴은 상기 캡쳐 장치(100)에서 상기 정적 메모리(116)의 상기 명령 인식 테이블(126)에 저장되고, 상기 다수의 인식 패턴의 각각은 상기 다수의 사전결정된 기능중의 하나를 수행하기 위해 상기 명령 인식 테이블(126)에 저장된 다수의 사전결정된 인스트럭션 집합중의 하나로 링크되는 상기 단계

를 더 포함하는 음성 제어 입력 방법.

청구항 7

제 6 항에 있어서,

단계(가1)은,

(가1)' 상기 캡쳐 장치(100)에서 상기 정적 메모리(116)의 음성 분석 소프트웨어(124)를 액세스하는 단계

인 새로운 단계에 의해 대체되고,

단계(가2), 단계(가6), 단계(가9)는 제거되고,

단계(가10)는,

(가10)' 상기 적어도 하나의 인식 패턴을 상기 캡쳐 장치(100)에서 상기 정적 메모리(116)의 상기 명령 인식 테이블(126)에 저장하는 단계로서, 상기 적어도 하나의 인식 패턴은 상기 사전결정된 기능을 수행하기 위해 상기 명령 인식 테이블(126)에 저장된 상기 제1 인스트럭션 집합으로 링크되는 상기 단계

인 새로운 단계에 의해 대체되는 음성 제어 입력 방법.

청구항 8

제 6 항 또는 제 7 항에 있어서,

상기 적어도 하나의 단어는 언어에 독립적인 음성 제어 입력 방법.

청구항 9

제 1 항에 있어서,

음성 주석 명령을 나타내는 상기 제1 인식 패턴은 상기 명령 인식 테이블(126)에 저장된 상기 적어도 하나의 인식 패턴과 매칭되고,

상기 제1 인스트럭션 집합이 실행되는 단계(바)는,

(바1) 사전결정된 주기의 음성 무입력이 발생되면, 단계(바2) 내지 단계(바6)을 수행하는 단계와;

(바2) 상기 캡쳐 장치(100)에서 상기 음성 픽업 소자(102)로써 음성 입력열을 포착하는 단계와;

(바3) 상기 음성 입력열을 제2 아날로그 신호로 변환시키는 단계와;

(바4) 상기 제2 아날로그 신호를 상기 캡처 장치(100)의 아날로그-디지탈 변환기(104)로 입력하는 단계와;

(바5) 상기 아날로그-디지탈 변환기(104)에서 상기 제2 아날로그 신호를 제2 디지탈 신호로 변환하는 단계와;

(바6) 상기 아날로그-디지탈 변환기(104)로부터의 상기 제2 디지탈 신호를 상기 캡쳐 장치(100)의 제어기(106)로 전송하는 단계와;

(바7) 상기 제어기(106)가 상기 제2 디지탈 신호를 음성 주석 파일로서 상기 캡처 장치(100)의 상기 정적 메모리(116)에 저장하는 단계와;

(바8) 상기 음성 주석 파일을 상기 정적 메모리(116)에 저장된 데이터 파일과 연결시키는 단계

를 더 포함하는 음성 제어 입력 방법.

청구항 10

제 1 항에 있어서,

(사) 제어기(106)가 상기 캡처 장치(100)상의 주석 제어 버튼으로부터 버튼 다운 입력을 수신하는 단계와;

(아) 상기 제어기가 상기 주석 제어 버튼으로부터 버튼 업 입력을 수신할 때 까지, 단계(자) 내지 단계(파)를 수행하는 단계와;

(자) 상기 음성 픽업 구성요소(102)로써 음성 입력열을 포착하는 단계와;

(차) 상기 음성 입력열을 제2 아날로그 신호로 변환하는 단계와;

(카) 상기 제2 아날로그 신호를 상기 캡처 장치(100)의 아날로그-디지탈 변환기(104)로 입력하는 단계와;

(타) 상기 아날로그-디지탈 변환기(104)에서 상기 제2 아날로그 신호를 제2 디지탈 신호로 변환하는 단계와;

(파) 상기 아날로그-디지탈 변환기(104)로부터의 상기 제2 디지탈 신호를 상기 캡처 장치(100)의 제어기(106)로 전송하는 단계와;

(하) 상기 제어기(106)가 상기 캡처 장치(100)상의 상기 주석 제어 버튼으로부터 상기 버튼 업 입력을 수신하는 단계와;

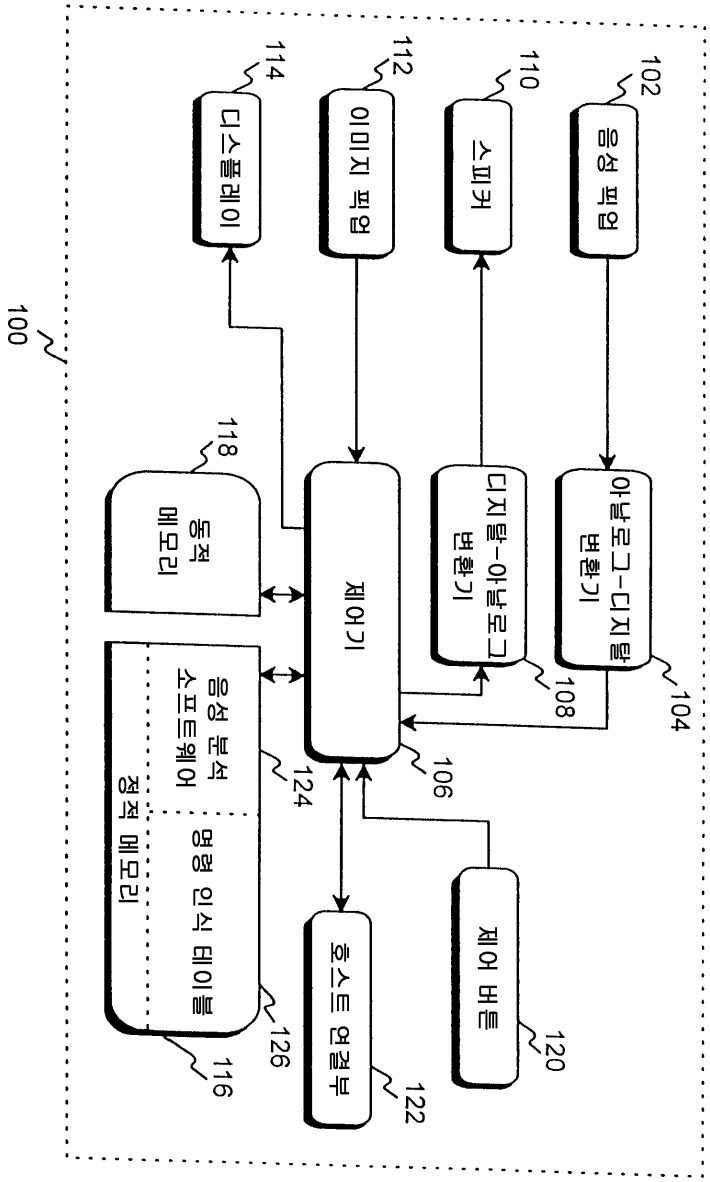
(우) 상기 제어기(106)가 상기 제2 디지탈 신호를 음성 주석 파일로서 상기 캡처 장치(100)의 상기 정적 메모리(116)에 저장하는 단계와;

(유) 상기 음성 주석 파일을 상기 정적 메모리(116)에 저장된 데이터 파일과 연결시키는 단계

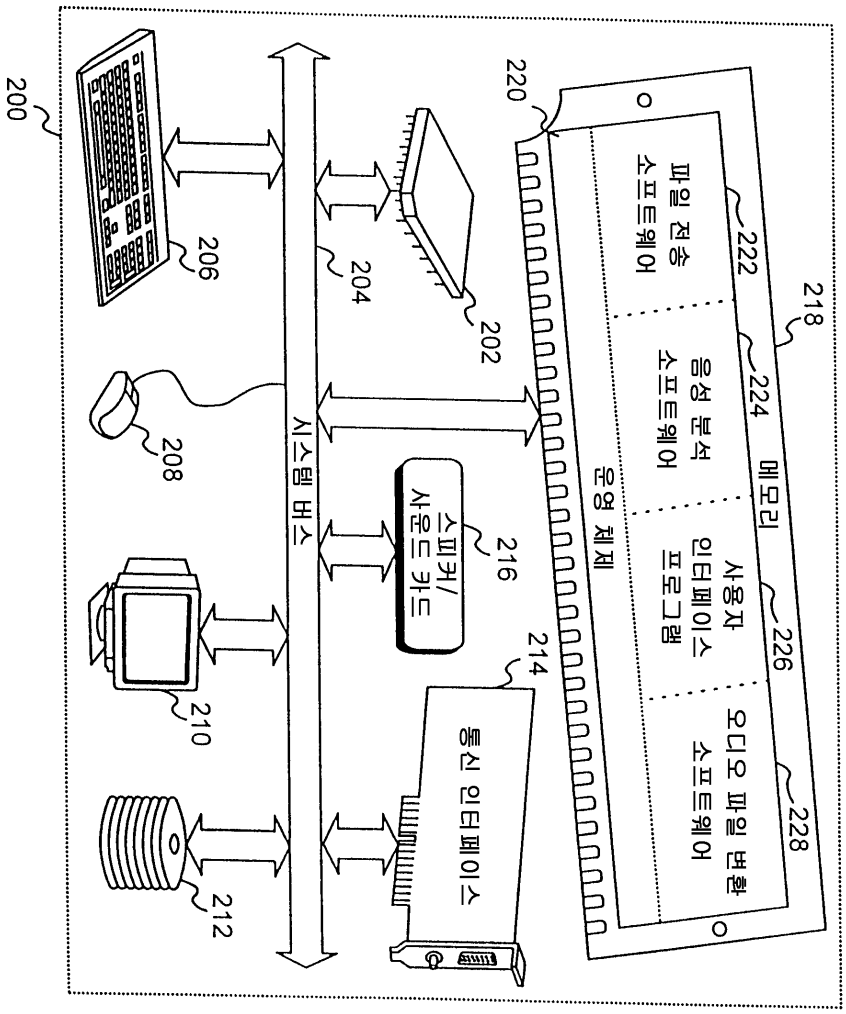
를 더 포함하는 음성 제어 입력 방법.

도면

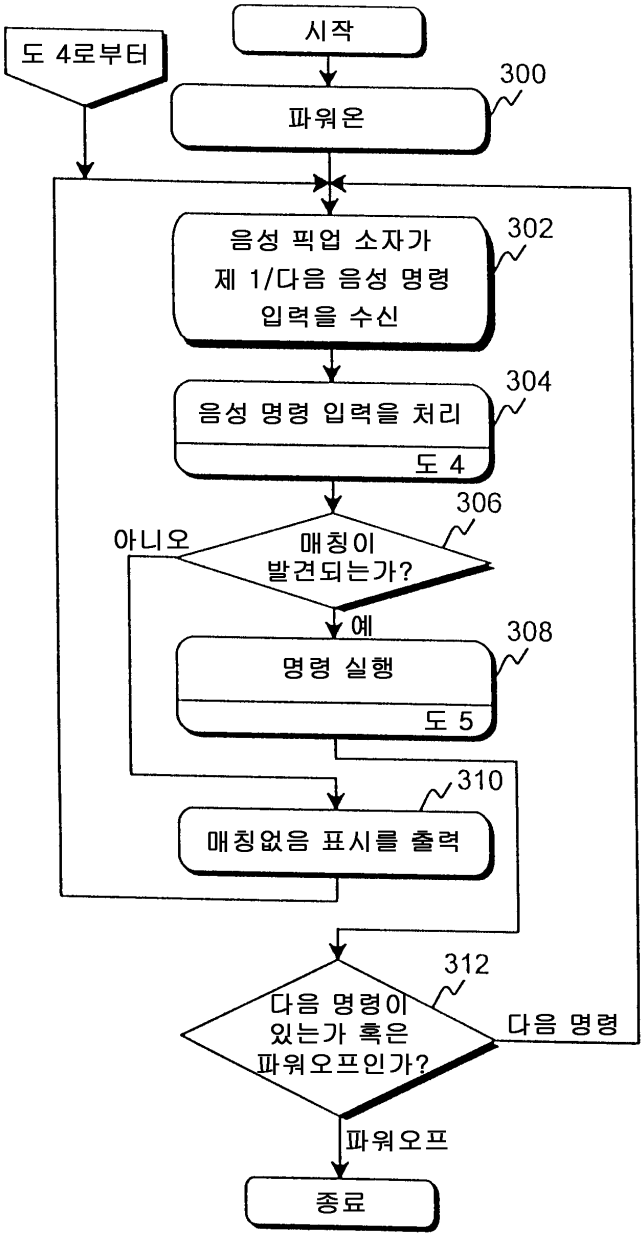
도면1


100
음성 픽업
102
아날로그-디지털 변환기
104
디지털-아날로그 변환기
108
스피커
110
이미지 픽업
112
디스플레이
114
제어기
106
동적 메모리
118
음성 분석 소프트웨어
124
명령 인식 테이블
126
정적 메모리
116
호스트 연결부
122
제어 버튼
120

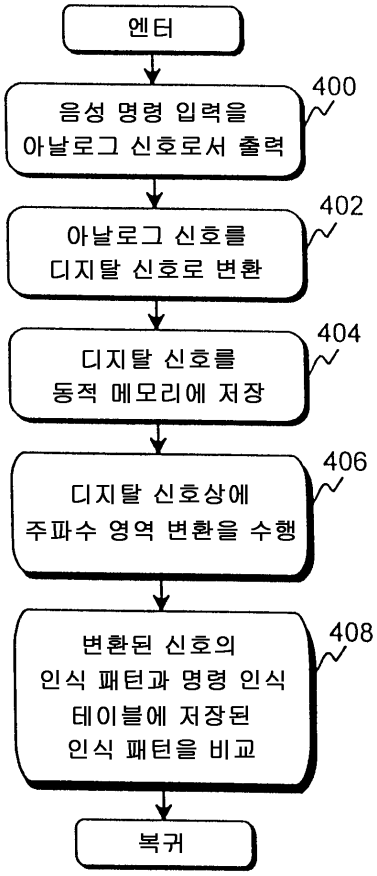
도면2


200
206
204
202
208
210
212
시스템 버스
스피커/
사운드 카드
216
214
통신 인터페이스
220
218
메모리
222
파일 전송
소프트웨어
224
음성 분석
소프트웨어
226
사용자
인터페이스
프로그램
228
오디오 파일 변환
소프트웨어
운영 체제

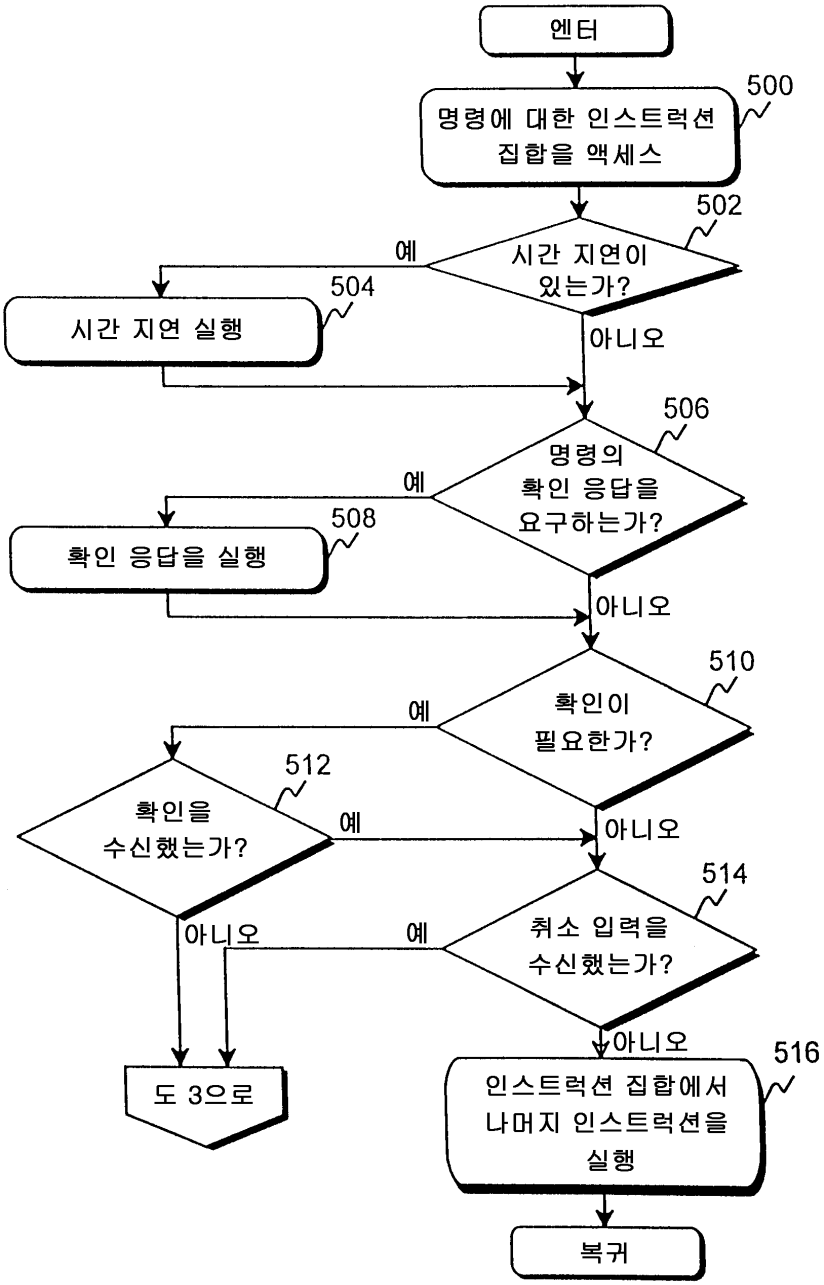
도면3


도 4로부터
시작
300
파워온
302
음성 픽업 소자가
제 1/다음 음성 명령
입력을 수신
304
음성 명령 입력을 처리
도 4
306
아니오
매칭이
발견되는가?
예
308
명령 실행
도 5
310
매칭없음 표시를 출력
312
다음 명령이
있는가 혹은
파워오프인가?
다음 명령
파워오프
종료

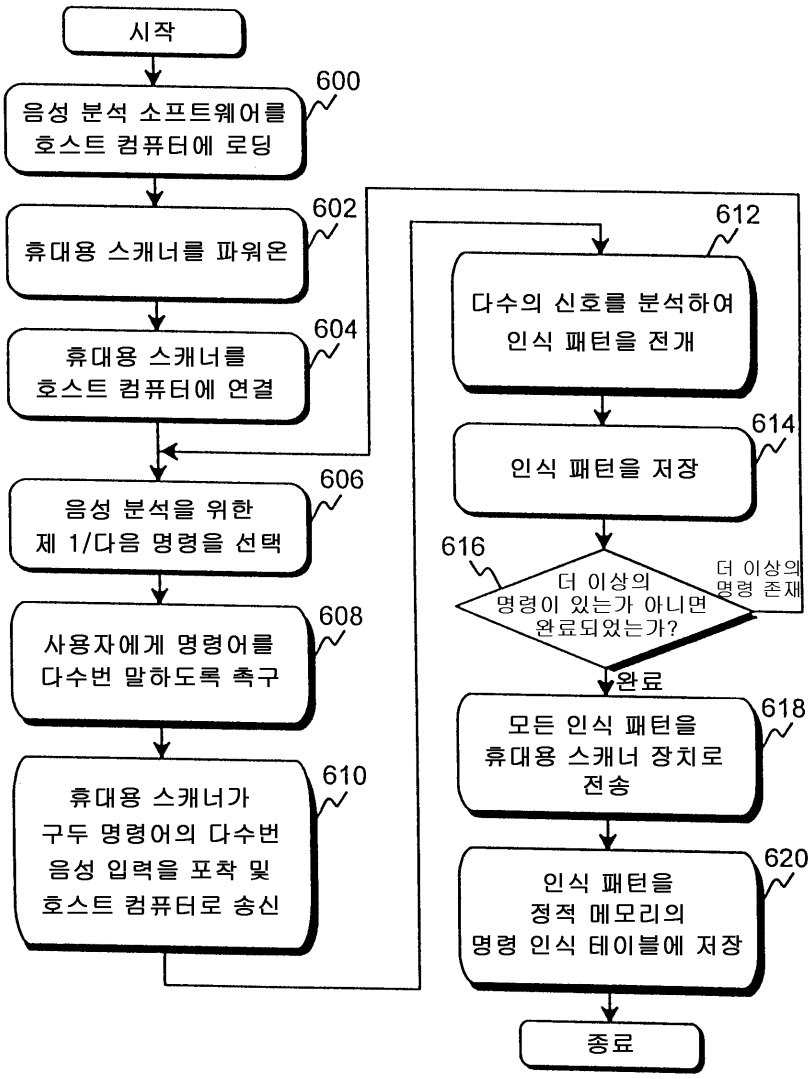
도면4


엔터
음성 명령 입력을 아날로그 신호로서 출력
400
아날로그 신호를 디지탈 신호로 변환
402
디지탈 신호를 동적 메모리에 저장
404
디지탈 신호상에 주파수 영역 변환을 수행
406
변환된 신호의 인식 패턴과 명령 인식 테이블에 저장된 인식 패턴을 비교
408
복귀

도면5


엔터
명령에 대한 인스트럭션 집합을 액세스
500
시간 지연이 있는가?
502
예
시간 지연 실행
504
아니오
명령의 확인 응답을 요구하는가?
506
예
확인 응답을 실행
508
아니오
확인이 필요한가?
510
예
확인을 수신했는가?
512
예
아니오
취소 입력을 수신했는가?
514
예
아니오
도 3으로
인스트럭션 집합에서 나머지 인스트럭션을 실행
516
복귀

도면6


시작
음성 분석 소프트웨어를 호스트 컴퓨터에 로딩
600
휴대용 스캐너를 파워온
602
휴대용 스캐너를 호스트 컴퓨터에 연결
604
음성 분석을 위한 제 1/다음 명령을 선택
606
사용자에게 명령어를 다수번 말하도록 촉구
608
휴대용 스캐너가 구두 명령어의 다수번 음성 입력을 포착 및 호스트 컴퓨터로 송신
610
다수의 신호를 분석하여 인식 패턴을 전개
612
인식 패턴을 저장
614
더 이상의 명령이 있는가 아니면 완료되었는가?
616
더 이상의 명령 존재
완료
모든 인식 패턴을 휴대용 스캐너 장치로 전송
618
인식 패턴을 정적 메모리의 명령 인식 테이블에 저장
620
종료