



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년06월28일
(11) 등록번호 10-2269848
(24) 등록일자 2021년06월22일

(51) 국제특허분류(Int. Cl.)
H04N 5/60 (2006.01) H04N 5/44 (2011.01)
(21) 출원번호 10-2014-0146349
(22) 출원일자 2014년10월27일
심사청구일자 2019년10월15일
(65) 공개번호 10-2016-0049347
(43) 공개일자 2016년05월09일
(56) 선행기술조사문헌
KR1020110010233 A*
KR1020140095853 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
엘지전자 주식회사
서울특별시 영등포구 여의대로 128 (여의도동)
(72) 발명자
이승신
서울특별시 서초구 양재대로11길 19
정대권
서울특별시 서초구 양재대로11길 19
(74) 대리인
박장원

전체 청구항 수 : 총 9 항

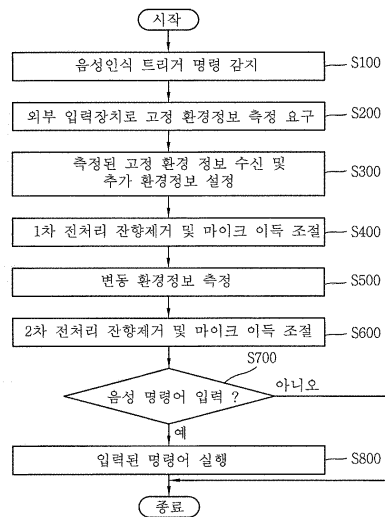
심사관 : 김혜린

(54) 발명의 명칭 영상표시기기 및 그의 원거리 음성 인식을 향상 방법

(57) 요약

본 발명은 사용 환경을 고려하여 영상 표시기기의 원거리 음성 인식 성능을 향상시킬 수 있는 영상표시기기 및 그의 제어 방법에 관한 것으로, 설치 공간의 고정 환경 정보를 수신하는 인터페이스부; 음성 명령어에 따른 동작을 실행하는 디스플레이부; 및 상기 수신된 설치 공간에 대한 고정 환경 정보 및 자체 측정된 영상표시기기의 시청 거리에 대한 변동 환경정보를 변수로 이용하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행하는 제어부;를 포함한다.

대표도 - 도4



명세서

청구범위

청구항 1

설치 공간의 고정 환경 정보를 수신하는 인터페이스부;

음성 명령어에 따른 동작을 실행하는 디스플레이부; 및

상기 수신된 설치 공간에 대한 고정 환경정보 및 자체 측정된 영상표시기기의 시청 거리에 대한 변동 환경정보를 변수로 이용하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행하는 제어부;를 포함하고,

상기 고정 환경정보는 거리 측정센서를 통해 직접 측정한 공간의 기본 환경정보를 포함하고,

상기 기본 환경정보는

영상표시기기가 설치된 공간의 크기, 영상표시기기의 설치 위치 및 외부 스피커까지의 거리를 포함하는 것을 특징으로 하는 영상표시기기.

청구항 2

제1항에 있어서, 상기 고정 환경정보는

외부 입력장치의 거리 측정센서에 의해 측정되는 것을 특징으로 하는 영상표시기기.

청구항 3

제1항에 있어서, 상기 고정 환경정보는

외부 입력장치의 사용자 입력부를 이용하여 영상표시기기의 디스플레이부에 표시된 메뉴상에서 설정한 부가 환경정보를 더 포함하는 것을 특징으로 하는 영상표시기기.

청구항 4

삭제

청구항 5

제3항에 있어서, 상기 부가 환경정보는

설치 공간의 종류와 벽 재질을 포함하는 것을 특징으로 하는 영상표시기기.

청구항 6

제1항에 있어서, 상기 제어부는

사용자로부터 음성의 트리거 키워드가 감지되면 외부 입력장치로 기본 환경정보의 측정을 요청하는 제어신호를 전송하고,

외부 입력장치로부터 기본 환경정보가 측정되어 수신되면 부가 환경정보를 설정하기 위한 메뉴를 디스플레이부에 표시하는 것을 특징으로 하는 영상표시기기.

청구항 7

제1항에 있어서, 상기 변동 환경정보는

발화자의 방향, 영상표시기기에서 발화자까지의 거리, 발화자의 평균 음량 및 잡음 음량을 포함하는 것을 특징으로 하는 영상표시기기.

청구항 8

제1항에 있어서, 상기 제어부는

설치 공간에 대한 고정 환경 정보를 이용하여 제1차 전처리 잔향제거 및 마이크 이득 조절을 수행하고, 영상표시기기의 시청거리에 따라 변동되는 변동 환경정보를 측정하여 제2차 전처리 잔향제거 및 마이크 이득 조절과 함께 인식기 언어 모델을 자동으로 설정하는 것을 특징으로 하는 영상표시기기.

청구항 9

제1항에 있어서, 상기 제어부는

이전의 음성인식에서 고정 환경 정보가 측정된 경우에는 변동 환경정보만을 측정하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행하는 것을 특징으로 하는 영상표시기기.

청구항 10

제1항에 있어서, 상기 제어부는

고정 환경 정보 및 변동 환경정보를 변수로 이용하여 영상표시기기의 음향, 화질, 3D 깊이감 및 커브드 TV의 곡률 반경을 자동으로 설정하는 것을 특징으로 하는 영상표시기기.

청구항 11

삭제

청구항 12

삭제

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

발명의 설명

기술 분야

[0001] 본 발명은 영상표시기기에 관한 것으로, 특히 사용 환경을 고려하여 영상 표시기기의 원거리 음성 인식 성능을 향상시킬 수 있는 영상표시기기 및 그의 제어 방법에 관한 것이다.

배경 기술

[0002] 영상표시기기는 동영상을 기록 및 재생하는 장치와 오디오를 기록 및 재생하는 장치를 모두 포함한다. 동영상을 기록 및 재생하는 장치는, 영상 표시 장치로서 TV, 컴퓨터 모니터, 프로젝터 등을 포함한다.

[0003] 이와 같은 영상표시기기는 기능이 다양화됨에 따라, 음악이나 동영상 파일의 재생 기능 외에도, 사진이나 동영상의 촬영, 게임, 방송의 수신 등의 복합적인 기능들을 갖춘 멀티영상표시기기(Multimedia player) 형태로 구현되고 있다. 나아가 영상표시기기의 기능 지지 및 증대를 위해, 영상표시기기의 구조적인 부분 및 소프트웨어적

인 부분을 개량하는 것이 고려될 수 있다.

- [0004] 일반적으로 음성 인식은 자동적 수단에 의하여 음성으로부터 언어적 의미 내용을 식별하는 것이다. 구체적으로 음성 인식은 음성 파형을 입력하여 단어나 단어열을 식별하고 의미를 추출하는 처리 과정으로, 넓은 의미로는 음성 분석, 음소 인식, 단어 인식, 문장 해석, 의미 추출로 분류될 수 있으며, 좁은 의미로는 음성 분석에서 단어 인식까지를 의미한다.
- [0005] 이와 같은 음성 인식은 사용자의 편의성 증대를 위하여 다양한 전자기기의 입력 수단으로 사용되고 있는데, 일 예로 TV나 에어컨과 같은 가전제품에서는 음성 인식장치를 구비하여 음성 인식 기능을 리모컨을 대신하여 사용하고 있다.
- [0006] 그런데, 현재 영상표시기기(e.g. TV)등에 구비된 음성 인식장치는 영상표시기가 설치된 위치에서 소정 거리 떨어져 있는(원거리) 사용자의 음성을 인식한다. 그 결과 발화자의 음성 세기가 정상보다 작을 경우에는 동일한 신호대 잡음비(SNR)에 대하여 음성 인식률이 감소하며, 주변 소음이 큰 경우 역시 음성 인식률이 현저하게 떨어지게 된다.
- [0007] 이러한 문제점을 방지하기 위하여 종래에는 영상표시기기에 다중 마이크(2개 또는 4개)를 구비하여 발화자의 방향을 탐지하여 음성 인식 성능을 개선하고, 마이크로 입력되는 발화자의 음성 증폭률을 주변 환경의 노이즈를 고려하여 조정하거나, 원거리 음성인식 성능에 중요한 환경인 잔향 시간(Reverberation time)의 영향을 줄이기 위해 빔포밍(Beamforming) 기술을 채용하거나 음성인식 엔진을 강화하는 등의 조치를 강구하고 있다.
- [0008] 그러나, 상기와 같은 종래의 음성 인식을 향상 방법은 발화자와 마이크간 거리에 따른 자동 이득조절은 고려되지 않고 발화자가 사용 가능한 예측거리 범위내에서 평균 마이크 이득을 사용하고 있을 뿐, 마이크 거리에 따른 자동 이득 조절과 상기 잔향 시간에 가장 큰 영향을 미치는 영상표시기기의 설치 환경, 예를들면 TV가 설치된 공간의 크기, 형상, 벽의 재질 및 TV설치 위치등은 전혀 고려하고 있지 않기 때문에 사용자 환경에 최적화된 원거리 음성 인식을 수행할 수 없는 단점이 있다. 그 결과, 주변 환경의 변화에 따라 음성인식 성능이 저하되고 음성 인식 성능이 보증 가능한 사용거리가 짧은 단점이 있다.

발명의 내용

해결하려는 과제

- [0009] 본 발명의 일 목적은 영상표시기기를 시청하는 사용자 환경에 최적화된 영상표시기기 및 그의 원거리 음성 인식을 향상 방법을 제공하는데 있다.
- [0010] 본 발명의 다른 목적은 사용자와 마이크간 거리에 따라 마이크 이득을 자동으로 조절할 수 있는 영상표시기기 및 그의 원거리 음성 인식을 향상 방법을 제공하는데 있다.
- [0011] 본 발명의 또 다른 목적은 사용자 환경을 고려하여 잔향시간을 감소시킬 수 있는 영상표시기기 및 그의 원거리 음성 인식을 향상 방법을 제공하는데 있다.

과제의 해결 수단

- [0012] 상기와 같은 목적을 달성하기 위하여 본 발명의 일 실시 예에 따른 영상표시기기는 설치 공간의 고정 환경 정보를 수신하는 인터페이스부; 음성 명령어에 따른 동작을 실행하는 디스플레이부; 및 상기 수신된 설치 공간에 대한 고정 환경 정보 및 자체 측정된 영상표시기기의 시청 거리에 대한 변동 환경정보를 변수로 이용하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행하는 제어부;를 포함한다.
- [0013] 상기 고정 환경정보는 외부 입력장치의 거리 측정센서를 통해 측정된다.
- [0014] 상기 고정 환경정보는 거리 측정센서를 통해 직접 측정된 공간의 기본 환경정보; 및 외부 입력장치의 사용자 입력부를 이용하여 영상표시기기의 디스플레이부에 표시된 메뉴상에서 설정한 부가 환경정보를 포함한다.
- [0015] 상기 기본 환경정보는 영상표시기기가 설치된 공간의 크기, 영상표시기기의 설치 위치 및 외부 스피커까지의 거리를 포함할 수 있다.
- [0016] 상기 부가 환경정보는 설치 공간의 종류와 벽 재질을 포함할 수 있다.
- [0017] 상기 제어부는 사용자로부터 음성의 트리거 키워드가 감지되면 외부 입력장치로 기본 환경정보의 측정을 요청하는 제어신호를 전송하고, 외부 입력장치로부터 기본 환경정보가 측정되어 수신되면 부가 환경정보를 설정하기

위한 메뉴를 디스플레이부에 표시할 수 있다.

- [0018] 상기 변동 환경정보는 발화자의 방향, 영상표시기기에서 발화자까지의 거리, 발화자의 평균 음량 및 잡음 음량을 포함할 수 있다.
- [0019] 상기 제어부는 설치 공간에 대한 고정 환경 정보를 이용하여 제1차 전처리 잔향제거 및 마이크 이득 조절을 수행하고, 영상표시기기의 시청거리에 따라 변동되는 변동 환경정보를 측정하여 제2차 전처리 잔향제거 및 마이크 이득 조절과 함께 인식기 언어 모델을 자동으로 설정할 수 있다.
- [0020] 상기 제어부는 이전의 음성인식에서 고정 환경 정보가 측정된 경우에는 변동 환경정보만을 측정하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행한다.
- [0021] 상기와 같은 목적을 달성하기 위하여 본 발명의 일 실시 예에 따른 영상표시기기의 원거리 음성 인식을 향상 방법은, 음성 인식이 트리거되면 영상표시기기의 설치 공간에 대한 고정 환경 정보를 수신하는 단계; 고정 환경정보를 이용하여 제1차 전처리 잔향 제거 및 마이크 이득 조절을 수행하는 단계; 영상표시기기의 시청 거리에 대한 변동 환경정보를 측정하는 단계; 및 상기 측정된 변동 환경정보를 이용하여 제2차 전처리 잔향제거, 마이크 이득 조절 및 인식기 언어 모델을 자동으로 설정하는 단계;를 포함한다.
- [0022] 상기 고정 환경 정보를 수신하는 단계는 사용자로부터 음성의 트리거 키워드가 입력되면 외부 입력장치로 기본 환경정보의 측정을 요청하는 단계; 외부 입력장치를 통해 기본 환경정보를 측정하는 단계; 상기 측정된 기본 환경정보가 수신되면 부가 환경정보를 설정하기 위한 메뉴를 디스플레이부에 표시하는 단계; 및 외부 입력장치를 통해 메뉴에서 부가 환경정보의 설정을 입력받는 단계;를 포함할 수 있다.
- [0023] 상기 기본 환경정보는 영상표시기기가 설치된 공간의 크기, 영상표시기기의 설치 위치 및 외부 스피커까지의 거리를 포함할 수 있다.
- [0024] 상기 부가 환경정보는 설치 공간의 종류와 벽 재질을 포함할 수 있다.
- [0025] 상기 변동 환경정보는 발화자의 방향, 영상표시기기에서 발화자까지의 거리, 발화자의 평균 음량 및 잡음 음량을 포함할 수 있다.
- [0026] 상기 음성 인식이 다시 트리거되면 변동 환경정보만을 측정하여 원거리 음성 인식을 위한 전처리 잔향제거 및 마이크 이득 조절을 수행하는 단계;를 더 포함할 수 있다.

발명의 효과

- [0027] 본 발명은 한번 설치되면 위치가 잘 변경되지 않는 영상표시기기(e.g., TV)의 특성을 감안하여 영상표시기기가 설치된 환경(설치 공간의 크기, 재질, 설치 위치, 외부 스피커의 위치등) 및 사용자의 위치에 따라 자동으로 전처리 잔향제거, 마이크 게인 조절, 음성인식기 언어 모델(Language model)을 자동으로 설정함으로써 원거리 인식 성능을 향상시킬 수 있는 효과가 있다.
- [0028] 또한, 본 발명은 사용자와 영상표시기기간 거리에 따라 전처리 잔향제거 및 마이크 게인 조절을 수행함으로써 먼 거리 또는 사용자의 목소리가 작은 경우에도 원거리 인식 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

- [0029] 도 1은 본 발명과 관련된 영상표시기기 및 외부 입력 장치를 설명하기 위한 블록도.
- 도 2는 도 1의 외부 입력 장치의 상세 블록도.
- 도 3은 본 발명과 관련된 영상표시기기 및 외부 입력 장치의 상호 동작을 설명하기 위한 블록도.
- 도 4는 본 발명의 실시예에 따른 영상표시기기의 원거리 음성 인식을 향상 방법을 나타낸 순서도.
- 도 5는 외부 입력 장치를 통해 주변환경을 측정하는 일 예를 나타낸 도면.
- 도 6은 추가 환경정보를 설정하기 위한 메뉴 구성의 일 예를 나타낸 도면.
- 도 7은 음성 인식 성능을 향상시키기 위하여 초기에 리모컨에 의해 수행되는 고정 환경 정보 측정 및 설정동작을 나타낸 순서도.
- 도 8은 음성 인식 성능을 향상시키기 위해 TV에 의해 수행되는 변동 사용자 정보를 측정 동작을 나타낸 순서도.

도 9는 본 발명의 다른 실시예에 따른 영상표시기기의 원거리 음성 인식을 향상 방법을 나타낸 순서도.

도 10은 본 발명의 실시예에 따른 영상표시기기의 근거리 음성 인식을 향상 방법을 나타낸 순서도.

발명을 실시하기 위한 구체적인 내용

- [0030] 이하, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 정도로 상세하게 설명하기 위하여, 본 발명의 실시 예가 첨부된 도면을 참조하여 설명한다. 하지만, 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 여기에서 설명하는 실시 예에 한정되지 않는다. 그리고, 도면에서 본 발명을 명확하게 설명하기 위해 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통해 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0031] 본 명세서에서, 영상표시기기는 방송을 수신하여 표시하거나, 동영상을 기록 및 재생하는 장치와 오디오를 기록 및 재생하는 장치를 모두 포함한다. 이하, 이러한 예로서, 텔레비전을 예를 들어 설명한다.
- [0032] 도 1은 본 발명과 관련된 영상표시기기(100) 및 외부 입력 장치(200)를 보여주는 블록도이다. 영상표시기기(100)는, 튜너(110), 복조부(120), 신호 입출력부(130), 인터페이스부(140), 제어부(150), 저장부(160), 디스플레이부(170) 및 오디오 출력부(180)를 포함한다. 다만, 외부 입력 장치(200)는 영상표시기기(100)와 별도의 장치이나, 영상표시기기(100)의 일 구성요소로 포함될 수도 있다.
- [0033] 도 1을 참조하면, 튜너(110)는 안테나를 통해 수신되는 RF(Radio Frequency) 방송 신호 중 사용자에게 의해 선택된 채널에 대응하는 RF 방송 신호를 선택하고, RF 방송 신호를 중간 주파수 신호 또는 베이스 밴드 영상/음성 신호로 변환한다. 예를 들어, RF 방송 신호가 디지털 방송 신호이면, 튜너(110)는 RF 방송 신호를 디지털 IF 신호(DIF)로 변환한다. 반면, RF 방송 신호가 아날로그 방송 신호이면, 튜너(110)는 RF 방송 신호를 아날로그 베이스 밴드 영상/음성신호(CVBS/SIF)로 변환된다. 이와 같이, 튜너(110)는 디지털 방송 신호와 아날로그 방송 신호를 처리할 수 있는 하이브리드 튜너일 수 있다.
- [0034] 튜너(110)에서 출력되는 디지털 IF 신호(DIF)는 복조부(120)로 입력되고, 튜너(110)에서 출력되는 아날로그 베이스 밴드 영상/음성신호(CVBS/SIF)는 제어부(160)로 입력될 수 있다. 튜너(120)는 ATSC(Advanced Television Systems Committee) 방식에 따른 단일 캐리어의 RF 방송 신호 또는 DVB(Digital Video Broadcasting) 방식에 따른 복수 캐리어의 RF 방송 신호를 수신할 수 있다.
- [0035] 비록 도면에는 하나의 튜너(110)가 도시되나, 이에 한정되지 않고, 영상표시기기(100)는 다수의 튜너, 예를 들어, 제 1 및 제 2 튜너를 구비할 수 있다. 이런 경우, 제 1 튜너는 사용자가 선택한 방송 채널에 대응하는 제 1 RF 방송 신호를 수신하고, 제 2 튜너는 기저장된 방송 채널에 대응하는 제 2 RF 방송 신호를 순차적으로 또는 주기적으로 수신할 수 있다. 제 2 튜너는 제 1 튜너와 마찬가지로 방식으로 RF 방송 신호를 디지털 IF 신호(DIF) 또는 아날로그 베이스 밴드 영상/음성신호(CVBS/SIF)로 변환할 수 있다.
- [0036] 복조부(120)는 튜너(110)에서 변환되는 디지털 IF 신호(DIF)를 수신하여 복조 동작을 수행한다. 예를 들어, 튜너(110)에서 출력되는 디지털 IF 신호(DIF)가 ATSC 방식이면, 복조부(120)는 8-VSB(8-Vestigal Side Band) 복조를 수행한다. 이때, 복조부(120)는 트렐리스 복호화, 디인터리빙(de-interleaving), 리드 솔로몬 복호화 등의 채널 복호화를 수행할 수도 있다. 이를 위해, 복조부(120)는 트렐리스 디코더(Trellis decoder), 디인터리버(de-interleaver) 및 리드 솔로몬 디코더(Reed Solomon decoder) 등을 구비할 수 있다.
- [0037] 다른 예를 들어, 튜너(110)에서 출력되는 디지털 IF 신호(DIF)가 DVB 방식이면, 복조부(120)는 COFDM(Coded Orthogonal Frequency Division Modulation) 복조를 수행한다. 이때, 복조부(120)는 컨벌루션 복호화, 디인터리빙, 리드 솔로몬 복호화 등의 채널 복호화를 수행할 수도 있다. 이를 위해, 복조부(120)는 컨벌루션 디코더(convolution decoder), 디인터리버 및 리드-솔로몬 디코더 등을 구비할 수 있다.
- [0038] 신호 입출력부(130)는 외부 기기와 연결되어 신호 입력 및 출력 동작을 수행하고, 이를 위해, A/V 입출력부(미도시됨) 및 무선 통신부(미도시됨)를 포함할 수 있다.
- [0039] A/V 입출력부는 이더넷(Ethernet) 단자, USB 단자, CVBS(Composite Video Banking Sync) 단자, 컴포넌트 단자, S-비디오 단자(아날로그), DVI(Digital Visual Interface) 단자, HDMI(High Definition Multimedia Interface) 단자, MHL (Mobile High-definition Link) 단자, RGB 단자, D-SUB 단자, IEEE 1394 단자, SPDIF 단자, 리퀴드(Liquid) HD 단자 등을 포함할 수 있다. 이러한 단자들을 통해 입력되는 디지털 신호는 제어부(150)에 전달될 수 있다. 이때, CVBS 단자 및 S-비디오 단자를 통해 입력되는 아날로그 신호는 아날로그-디지털 변환부(미도

시)를 통해 디지털 신호로 변환되어 제어부(150)로 전달될 수 있다.

- [0040] 무선 통신부는 무선 인터넷 접속을 수행할 수 있다. 예를 들어, 무선 통신부는 WLAN(Wireless LAN)(Wi-Fi), Wibro(Wireless broadband), Wimax(World Interoperability for Microwave Access), HSDPA(High Speed Downlink Packet Access) 등을 이용하여 무선 인터넷 접속을 수행할 수 있다. 또한, 무선 통신부는 다른 전자기기와 근거리 무선 통신을 수행할 수 있다. 예를 들어, 무선 통신부는 블루투스(Bluetooth), RFID(Radio Frequency Identification), 적외선 통신(IrDA, infrared Data Association), UWB(Ultra Wideband), 지그비(ZigBee) 등을 이용하여 근거리 무선 통신을 수행할 수 있다.
- [0041] 신호 입출력부(130)는 DVD(Digital Versatile Disk) 플레이어, 블루레이(Blu-ray) 플레이어, 게임기기, 캠코더, 컴퓨터(노트북), 휴대기기, 스마트 폰 등과 같은 외부 기기로부터 제공되는 영상 신호, 음성 신호 및 데이터 신호를 제어부(150)로 전달할 수 있다. 또한, 메모리장치, 하드디스크 등과 같은 외부 저장 장치에 저장된 다양한 미디어 파일의 영상 신호, 음성 신호 및 데이터 신호를 제어부(150)로 전달할 수 있다. 또한, 제어부(150)에 의해 처리된 영상 신호, 음성 신호 및 데이터 신호를 다른 외부 기기로 출력할 수 있다.
- [0042] 신호 입출력부(130)는 상술한 각종 단자 중 적어도 하나를 통해 셋톱 박스, 예를 들어, IPTV(Internet Protocol TV)용 셋톱 박스와 연결되어 신호 입력 및 출력 동작을 수행할 수 있다. 예를 들어, 신호 입출력부(130)는 양방향 통신이 가능하도록 IPTV용 셋톱 박스에 의해 처리된 영상 신호, 음성 신호 및 데이터 신호를 제어부(150)로 전달할 수 있고, 제어부(150)에 의해 처리된 신호들을 IPTV용 셋톱 박스로 전달할 수도 있다. 여기서, IPTV는 전송 네트워크에 따라 구분되는 ADSL-TV, VDSL-TV, FTTH-TV 등을 포함할 수 있다.
- [0043] 복조부(120) 및 신호 출력부(130)에서 출력되는 디지털 신호는 스트림 신호(TS)를 포함할 수 있다. 스트림 신호(TS)는 영상 신호, 음성 신호 및 데이터 신호가 다중화된 신호일 수 있다. 예를 들어, 스트림 신호(TS)는 MPEG-2 규격의 영상 신호, 돌비(Dolby) AC-3 규격의 음성 신호 등이 다중화된 MPEG-2 TS(Transport Stream)일 수 있다. 여기서, MPEG-2 TS는 4 바이트(byte)의 헤더와 184 바이트의 페이로드(payload)를 포함할 수 있다.
- [0044] 인터페이스부(140)는 외부 입력 장치(200)로부터 전원 제어, 채널 선택, 화면 설정 등을 위한 입력 신호를 수신하거나, 제어부(160)에 의해 처리된 신호를 외부 입력 장치(200)로 전송할 수 있다. 인터페이스부(140)와 외부 입력 장치(200)는 유선 또는 무선으로 연결될 수 있다.
- [0045] 상기 인터페이스부(140)의 일 예로서, 센서부가 구비될 수 있으며, 센서부는 원격조정기, 예를 들어 리모컨으로부터 상기 입력 신호를 감지하도록 이루어진다.
- [0046] 네트워크 인터페이스부(미도시)는, 영상표시기기(100)를 인터넷망을 포함하는 유/무선 네트워크와 연결하기 위한 인터페이스를 제공한다. 네트워크 인터페이스부(230)는, 유선 네트워크와의 접속을 위해, 이더넷(Ethernet) 단자 등을 구비할 수 있으며, 무선 네트워크와의 접속을 위해, WLAN(Wireless LAN)(Wi-Fi), Wibro(Wireless broadband), Wimax(World Interoperability for Microwave Access), HSDPA(High Speed Downlink Packet Access) 통신 규격 등이 이용될 수 있다.
- [0047] 네트워크 인터페이스부(미도시)는, 네트워크를 통해, 소정 웹 페이지에 접속할 수 있다. 즉, 네트워크를 통해 소정 웹 페이지에 접속하여, 해당 서버와 데이터를 송신 또는 수신할 수 있다. 그 외, 콘텐츠 제공자 또는 네트워크 운영자가 제공하는 콘텐츠 또는 데이터들을 수신할 수 있다. 즉, 네트워크를 통하여 콘텐츠 제공자 또는 네트워크 제공자로부터 제공되는 영화, 광고, 게임, VOD, 방송 신호 등의 콘텐츠 및 그와 관련된 정보를 수신할 수 있다. 또한, 네트워크 운영자가 제공하는 펌웨어의 업데이트 정보 및 업데이트 파일을 수신할 수 있다. 또한, 인터넷 또는 콘텐츠 제공자 또는 네트워크 운영자에게 데이터들을 송신할 수 있다.
- [0048] 또한, 네트워크 인터페이스부(미도시)는, 네트워크를 통해, 공중에 공개(open)된 애플리케이션들 중 원하는 애플리케이션을 선택하여 수신할 수 있다.
- [0049] 제어부(150)는 영상표시기기(100)의 전반적인 동작을 제어할 수 있다. 보다 구체적으로, 제어부(150)는 영상의 생성 및 출력을 제어하도록 형성된다. 예를 들어, 제어부(150)는 사용자가 선택한 채널 또는 기저장된 채널에 대응하는 RF 방송 신호를 튜닝(tuning)하도록 튜너(110)를 제어할 수 있다. 비록 도면에는 도시되지 않았으나, 제어부(150)는 역다중화부, 영상 처리부, 음성 처리부, 데이터 처리부, OSD(On Screen Display) 생성부 등을 포함할 수 있다. 또한, 제어부(150)는 하드웨어적으로 CPU 나 주변기기 등을 포함할 수 있다.
- [0050] 제어부(150)는 스트림 신호(TS), 예를 들어, MPEG-2 TS를 역다중화하여 영상 신호, 음성 신호 및 데이터 신호로 분리할 수 있다.

- [0051] 제어부(150)는 역다중화된 영상 신호에 대한 영상 처리, 예를 들어, 복호화를 수행할 수 있다. 좀더 상세하게, 제어부(150)는 MPEG-2 디코더를 이용하여 MPEG-2 규격의 부호화된 영상 신호를 복호화하고, H.264 디코더를 이용하여 DMB(Digital Multimedia Broadcasting) 방식 또는 DVB-H에 따른 H.264 규격의 부호화된 영상 신호를 복호화할 수 있다. 또한, 제어부(150)는 영상 신호의 밝기(brightness), 틴트(tint) 및 색조(color) 등이 조절되도록 영상 처리할 수 있다. 제어부(150)에 의해 영상 처리된 영상 신호는 디스플레이부(170)로 전달되거나, 외부 출력 단자를 통해 외부 출력 장치(미도시)로 전달될 수 있다.
- [0052] 제어부(150)는 역다중화된 음성 신호에 대한 음성 처리, 예를 들어, 복호화를 수행할 수 있다. 좀더 상세하게, 제어부(150)는 MPEG-2 디코더를 이용하여 MPEG-2 규격의 부호화된 음성 신호를 복호화하고, MPEG 4 디코더를 이용하여 DMB 방식에 따른 MPEG 4 BSAC(Bit Sliced Arithmetic Coding) 규격의 부호화된 음성 신호를 복호화하며, AAC 디코더를 이용하여 위성 DMB 방식 또는 DVB-H에 따른 MPEG 2의 AAC(Advanced Audio Codec) 규격의 부호화된 음성 신호를 복호화할 수 있다. 또한, 제어부(150)는 베이스(Base), 트레블(Treble), 음량 조절 등을 처리할 수 있다. 제어부(150)에서 처리된 음성 신호는 오디오 출력부(180), 예를 들어, 스피커로 전달되거나, 외부 출력 장치로 전달될 수 있다.
- [0053] 제어부(150)는 아날로그 베이스 밴드 영상/음성신호(CVBS/SIF)에 대한 신호 처리를 수행할 수 있다. 여기서, 제어부(150)에 입력되는 아날로그 베이스 밴드 영상/음성신호(CVBS/SIF)는 튜너(110) 또는 신호 입출력부(130)에서 출력된 아날로그 베이스 밴드 영상/음성신호일 수 있다. 신호 처리된 영상 신호는 디스플레이부(170)를 통해 표시되고, 신호 처리된 음성 신호는 오디오 출력부(180)를 통해 출력된다.
- [0054] 제어부(150)는 역다중화된 데이터 신호에 대한 데이터 처리, 예를 들어, 복호화를 수행할 수 있다. 여기서, 데이터 신호는 각각의 채널에서 방영되는 방송프로그램의 시작시간, 종료시간 등의 방송정보를 포함하는 EPG(Electronic Program Guide) 정보를 포함할 수 있다. EPG 정보는, 예를 들어, ATSC 방식에서는 TSC-PSIP(ATSC-Program and System Information Protocol) 정보를 포함하고, DVB 방식에서는 DVB-SI(DVB-Service Information) 정보를 포함할 수 있다. ATSC-PSIP 정보 또는 DVB-SI 정보는 MPEG-2 TS의 헤더(4 byte)에 포함될 수 있다.
- [0055] 제어부(150)는 OSD 처리를 위한 제어 동작을 수행할 수 있다. 좀더 상세하게, 제어부(150)는 영상 신호 및 데이터 신호 중 적어도 하나 또는 외부 입력 장치(200)로부터 수신되는 입력 신호에 근거하여 각종 정보를 그래픽(Graphic)이나 텍스트(Text) 형태로 표시하기 위한 OSD 신호를 생성할 수 있다. OSD 신호는 영상표시기기(100)의 사용자 인터페이스 화면, 메뉴 화면, 위젯, 아이콘 등의 다양한 데이터를 포함할 수 있다.
- [0056] 저장부(160)는 제어부(150)의 신호 처리 및 제어를 위한 프로그램이 저장될 수도 있고, 신호 처리된 영상 신호, 음성 신호 및 데이터 신호를 저장할 수도 있다. 저장부(160)는 플래시 메모리(flash memory), 하드디스크(hard disk), 멀티미디어 카드 마이크로 타입(multimedia card micro type), 카드 타입의 메모리(예를 들어 SD 또는 XD 메모리 등), 램(random access memory; RAM), SRAM(static random access memory), 롬(read-only memory; ROM), EEPROM(electrically erasable programmable read-only memory), PROM(programmable read-only memory), 자기 메모리, 자기 디스크, 광디스크 중 적어도 하나의 저장매체를 포함할 수 있다.
- [0057] 디스플레이부(170)는 제어부(150)에 의해 처리된 영상 신호, 데이터 신호, OSD 신호 등을 RGB 신호로 변환하여 구동 신호를 생성할 수 있다. 이를 통하여, 디스플레이부(170)는 영상을 출력하게 된다. 디스플레이부(170)는 플라즈마 디스플레이 패널(Plasma Display Panel: PDP), 액정 디스플레이(Liquid Crystal Display: LCD), 박막 트랜지스터 액정 디스플레이(Thin Film Transistor-Liquid Crystal Display: TFT-LCD), 유기 발광 다이오드(Organic Light Emitting Diode: OLED), 플렉시블 디스플레이(flexible display), 3차원 디스플레이(3D display), 전자잉크 디스플레이(e-ink display) 등의 다양한 형태로 구현될 수 있다. 또한, 디스플레이(180)는 터치 스크린으로 구현되어 입력 장치의 기능도 수행할 수 있다.
- [0058] 오디오 출력부(180)는 제어부(150)에 의해 처리된 음성 신호, 예를 들어, 스테레오 신호 또는 5.1 채널 신호를 출력한다. 오디오 출력부(180)는 다양한 형태의 스피커로 구현될 수 있다.
- [0059] 한편, 영상표시기기(100)는 피사체를 촬영하는 촬영부(미도시)와 오디오 입력부(미도시)를 더 구비할 수 있다. 촬영부(미도시)는 1 개의 카메라로 구현되는 것이 가능하나, 이에 한정되지 않으며, 복수 개의 카메라로 구현되는 것도 가능하다. 촬영부(미도시)에서 촬영된 영상 정보는 제어부(270)에 입력된다. 상기 오디오 입력부는 2개 또는 4의 마이크로 구성될 수 있다.
- [0060] 한편, 사용자의 제스처를 감지하기 위해, 상술한 바와 같이, 터치 센서, 음성 센서, 위치 센서, 동작 센서, 거

리측정 센서 중 적어도 하나를 구비하는 센싱부(미도시)가 영상표시기기(100)에 더 구비될 수 있다. 센싱부(미도시)에서 감지된 신호는 사용자입력 인터페이스부(140)를 통해 제어부(150)로 전달될 수 있다.

- [0061] 제어부(150)는, 촬영부(미도시)로부터 촬영된 영상, 또는 센싱부(미도시)로부터의 감지된 신호를 각각 또는 조합하여 사용자의 제스처를 감지할 수도 있다.
- [0062] 전원 공급부(미도시)는, 영상표시기기(100) 전반에 걸쳐 해당 전원을 공급한다. 특히, 시스템 온 칩(System On Chip, SOC)의 형태로 구현될 수 있는 제어부(150)와, 영상 표시를 위한 디스플레이부(170), 및 오디오 출력을 위한 오디오 출력부(180)에 전원을 공급할 수 있다.
- [0063] 이를 위해, 전원 공급부(미도시)는, 교류 전원을 직류 전원으로 변환하는 컨버터(미도시)를 구비할 수 있다. 한편, 예를 들어, 디스플레이부(170)가 다수의 백라이트 램프를 구비하는 액정패널로서 구현되는 경우, 휘도 가변 또는 디밍(dimming) 구동을 위해, PWM 동작이 가능한 인버터(미도시)를 더 구비할 수도 있다.
- [0064] 외부 입력 장치(200)는 무선으로 인터페이스부(140)와 연결되며, 사용자 입력에 따라 생성되는 입력 신호 및 측정된 사용자 주변의 환경 정보를 인터페이스부(140)로 전송한다. 외부 입력 장치(200)는 원격조정기, 마우스, 키보드 등을 포함할 수 있다. 원격조정기는 블루투스(Bluetooth), RF 통신, 적외선 통신, UWB(Ultra Wideband), 지그비(ZigBee) 방식 등을 통해 입력 신호를 인터페이스부(140)로 전송할 수 있다. 원격조정기는 공간 원격 제어 장치로서 구현될 수 있다. 공간 원격 제어 장치는 공간에서 본체의 동작을 감지하여 입력 신호를 생성할 수 있다.
- [0065] 영상표시기기(100)는 ATSC 방식(8-VSB 방식)의 디지털 방송, DVB-T 방식(COFDM 방식)의 디지털 방송, ISDB-T 방식(BST-OFDM 방식)의 디지털 방송 등 중 적어도 하나를 수신 가능한 고정형 디지털 방송 수신기로 구현될 수 있다. 또한, 영상표시기기(100)는 지상파 DMB 방식의 디지털 방송, 위성 DMB 방식의 디지털 방송, ATSC-M/H 방식의 디지털 방송, DVB-H 방식(COFDM 방식)의 디지털 방송, 미디어플로(Media Forward Link Only) 방식의 디지털 방송 등 중 적어도 하나를 수신 가능한 이동형 디지털 방송 수신기로 구현될 수 있다. 또한, 영상표시기기(100)는 케이블, 위성통신, IPTV용 디지털 방송 수신기로 구현될 수 있다.
- [0066] 도 2는 도 1의 외부 입력 장치(200)를 상세하게 보여주는 블록도이다. 외부 입력 장치(200)는 무선 통신부(210), 사용자 입력부(220), 감지부(230), 출력부(240), 전원 공급부(250), 저장부(260) 및 제어부(270)를 포함한다.
- [0067] 도 2를 참조하면, 무선 통신부(210)는 영상표시기기(100)로 신호를 송신하거나 영상표시기기(100)로부터 신호를 수신한다. 이를 위해, 무선 통신부(210)는 RF 모듈(211) 및 IR 모듈(212)을 구비할 수 있다. RF 모듈(211)은 RF 통신 규격에 따라 영상표시기기(100)의 인터페이스부(140)와 연결되어 신호를 송수신하고, IR 모듈(212)은 IR 통신 규격에 따라 영상표시기기(100)의 인터페이스부(140)와 신호를 송수신한다.
- [0068] 사용자 입력부(220)는 입력 수단으로 키 패드, 키 버튼, 스크롤 키, 조그 키 등을 구비할 수 있다. 사용자는 사용자 입력부(220)를 조작하여 영상표시기기(100)와 관련된 명령을 입력할 수 있다. 이러한 명령은, 예를 들어, 사용자가 사용자 입력부(200)의 하드 키 버튼의 푸쉬 동작을 통해 입력될 수 있다.
- [0069] 감지부(230)는 자이로 센서(231), 가속도 센서(232) 및 거리 측정 센서를 구비할 수 있다. 자이로 센서(231)는 외부 입력 장치(200)의 공간 움직임을 x축, y축 및 z축을 기준으로 감지할 수 있다. 가속도 센서(232)는 외부 입력 장치(200)의 이동 속도 등을 감지할 수 있다. 거리 측정 센서는 영상표시기기(100)가 설치된 공간의 크기, 설치위치 및 외부 스피커와의 거리를 측정한다.
- [0070] 출력부(240)는 사용자 입력부(230)의 조작에 따른 정보 및 영상표시기기(100)의 전송 신호에 대응하는 정보를 출력한다. 따라서, 사용자는 출력부(240)를 통해 사용자 입력부(230)의 조작 상태 또는 영상표시기기(100)의 제어 상태를 인지할 수 있다. 예를 들어, 출력부(240)는, 사용자 입력부(230)의 조작 또는 무선 통신부(210)을 통한 신호 송수신에 응답하여, 점등되는 LED 모듈(241), 진동을 발생하는 진동 모듈(242), 음향을 출력하는 음향 출력 모듈(243) 및 영상을 출력하는 디스플레이 모듈(244)을 구비할 수 있다.
- [0071] 전원 공급부(250)는 외부 입력 장치(200)의 각종 전자 소자에 전원을 공급한다. 전원 공급부(250)는 외부 입력 장치(200)가 소정 시간 동안 움직이지 않은 경우, 전원 공급을 중단함으로써, 전원 낭비를 줄일 수 있다. 전원 공급부(250)는 외부 입력 장치(200)의 소정 키가 조작되는 경우, 전원 공급을 재개할 수 있다.
- [0072] 저장부(260)는 외부 입력 장치(200)의 제어 또는 동작과 관련된 각종 프로그램, 지문 인식 프로그램, 애플리케이션, 주파수 대역 정보 및 하나 이상의 사용자 지문을 저장할 수 있다. 제어부(270)는 외부 입력 장치(200)의

전반적인 제어 동작을 수행한다.

- [0073] 본 발명은 음성 인식시 영상표시기기가 설치된 주변 환경 및 사용자 거리에 따라 최적화된 전처리 잔향 제거 및 마이크 이득을 조절함으로써 사용자의 음성인식 성능을 향상시킬 수 있는 다양한 방안을 제안한다.
- [0074] 상기 주변 환경은 영상표시기기가 설치된 공간의 크기(넓이), 형상, 설치 위치, 벽면 재질 및 외부 스피커 위치 등을 포함할 수 있다.
- [0075] 상기 영상표시기기는 대화면기기로 TV를 포함할 수 있다.
- [0076] 상기 영상표시기기는 2개 또는 4개의 다중 마이크를 구비한다.
- [0077] 도 3은 본 발명과 관련된 영상표시기기(100) 및 외부 입력 장치(200)의 상호 동작을 보여주는 개념도이다. 여기에서, 영상 표시기기(100)의 예로서 TV 수상기가 도시되고, 외부 입력 장치(200)의 예로서 원격 조정기가 도시된다.
- [0078] 도 3을 참조하면, 본 발명에서 외부 입력 장치(200)는 원격조정기로 사용된다. 외부 입력 장치(200)는 내장된 거리 측정센서(초음파 또는 레이저)를 이용하여 영상표시기기(100)가 설치된 공간의 크기, 영상표시기기의 설치 위치 및 외부 스피커와의 거리를 측정하고, 측정된 환경 정보를 RF(또는 IR) 통신 규격에 따라 영상 표시기기(100)로 전송할 수 있다.
- [0079] 영상표시기기(100)의 화면에는 음성 인식의 수행 여부 또는 환경 정보의 측정 여부를 묻는 팝-업 메시지가 표시될 수 있으며, 상기 측정된 환경 정보에 대한 추가 정보를 설정할 수 있는 설정 메뉴를 표시할 수 있다. 상기 설정 메뉴는 공간의 종류(거실, 방, 공공장소 및 기타) 및 공간의 재질(벽지, 대리석, 벽돌 및 기타)을 설정할 수 있는 항목을 포함할 수 있다.
- [0080] 외부 입력 장치(200)는 근거리 음성인식시 거리측정 센서를 이용하여 사용자와의 거리를 측정하여, 측정된 거리에 따라 자체 내장된 마이크의 이득을 조절할 수 있다. 또한, 상기 외부 입력 장치(200)는 사용자 인증을 위한 지문 스캐너(50)를 포함할 수 있다.
- [0081] 위에서 살펴본 본 발명의 일 실시 예에 따른 영상표시기기에서는 외부 입력 장치에서 측정된 환경 정보와 영상표시기기와 사용자간 거리에 근거하여 전처리 잔향제거, 마이크 이득조절 및 음성인식기 언어 모델을 자동으로 조절하여 음성인식 성능을 향상시킨다.
- [0082] 이하에서는, 영상표시기기가 설치된 주변환경을 측정하여 음성 인식률을 향상시키는 방법에 대하여 첨부된 도면과 함께 보다 구체적으로 살펴본다.
- [0083] 본 발명은 한번 설치되면 위치가 잘 변경되지 않는 영상표시기기(e.g., TV)의 특성을 감안하여 영상표시기기가 설치된 환경(설치 공간의 크기, 재질, 설치 위치, 외부 스피커의 위치등)을 최초 원거리 음성인식 사용시에 측정한 후 변동 가능한 변수(사용자와 영상표시기기간 거리, 사용자의 목소리 크기 및 기타)를 추가로 고려하여 사용자 개인에 맞는 전처리 잔향제거, 마이크 게인 조절, 음성인식기 언어 모델(Language model)을 자동으로 설정함으로써 원거리 인식 성능을 향상시킨다.
- [0084] 도 4는 본 발명의 실시예에 따른 영상표시기기의 원거리 음성 인식을 향상 방법을 나타낸 순서도이고, 도 5는 외부 입력 장치를 이용하여 주변환경을 측정하는 일 예이다. 또한, 도 6은 추가 환경정보를 설정하기 위한 메뉴 구성의 일 예를나타낸다. 본 발명은 설명의 편의를 위하여 영상표시기기는 TV, 외부 입력장치는 리모컨을 예로 들어 설명한다.
- [0085] 도 4에 도시된 바와같이, 사용자는 사전에 약속된 트리거 키워드를 발화하여 TV(100)에게 원거리 음성인식 동작이 트리거되었음을 알릴 수 있다. 상기 트리거 키워드는 예를들면 소정 단어 "음성" 또는 단어의 조합 "음성 인식"일 수 있으며 단순한 음이어도 무방하다. 상기 트리거 키워드는 메뉴에서 설정하거나 직접 녹음하여 설정할 수 있다.
- [0086] TV(100)의 전면에는 사용자와 대향되도록 일정 거리를 두고 적어도 하나 이상의 마이크가 구비되어 있다. 일 예로 TV(100)의 상측 하측의 좌우에 구비될 수 있다.
- [0087] 사용자로부터 사전에 약속된 음성의 트리거 키워드가 감지되면(S100), 제어부(150)는 디스플레이부(170)에 사용자에게 음성 인식이 시작됨을 나타내는 정보를 표시할 수 있다. 상기 정보는 팝-업 또는 기타 식별 가능한 아이콘일 수 있다.

- [0088] 원거리 음성인식이 트리거되면 제어부(180)는 리모컨(200)으로 기본 환경정보의 측정을 요구하는 제어신호를 전송한다(S200). 리모컨(200)은 제어신호에 따라 거리 측정센서(초음파 또는 레이저)를 동작시켜 도 5에 도시된 바와같이, 공간(300)을 향해 여러 방향으로 초음파 또는 레이저를 송출하고, 반사되는 신호를 분석하여 현재 TV(100)가 설치된 공간(300)의 크기, TV 설치 위치 및 외부 스피커까지의 거리를 측정한다. 상기 리모컨(200)이 측정하는 기본 환경 정보는 변동이 적은 정보로 초기 음성인식률 보정에 이용된다.
- [0089] 기본 환경정보가 측정되면 리모컨(200)은 기본 환경정보를 RF(또는 IR) 통신 규격에 따라 TV(100)로 전송한다. 리모컨(200)에서 측정된 기본 환경 정보가 도 1의 인터페이스부(140)를 통해 수신되면 제어부(150)는 저장부(160)에 저장함과 함께 도 6에 도시된 바와같이, 설치 공간(300)에 대한 추가 환경정보를 설정할 수 있는 메뉴(60)를 디스플레이부(1760)에 표시한다(S300).
- [0090] 상기 환경정보 설정 메뉴는 설치 공간의 종류와 재질을 선택할 수 있는 항목으로 구성되며, 각 항목이 선택하면 상세 선택 항목이 표시된다. 상기 설치 공간(300)의 종류는 거실, 방, 공공장소 등이 있으며, 상기 설치 공간(300)의 재질은 벽지, 대리석, 벽돌등을 포함한다. 상기 설치 공간(300)의 종류는 잔향 시간에 큰 영향을 미치는 방의 넓이를 사용자가 선택하기 위한 것이며, 상기 방의 재질은 잔향시간에 큰 영향을 미치는 흡음률을 선택하기 위함이다. 상기 메뉴(60)에서의 추가 환경정보는 도 2의 리모컨(200)의 사용자 입력부(220)을 통해 설정될 수 있다. 상기 측정된 기본 환경정보와 추가 환경정보는 고정 환경정보로 차후 잔향제거 및 마이크 이득조절의 변수로 이용된다.
- [0091] 고정 환경정보의 설정이 완료되면 제어부(150)는 기본 환경정보와 추가 환경정보를 이용하여 전처리 잔향 제거를 수행함과 함께 마이크들의 이득을 조정한다(1차 잔향제거 및 마이크 이득조절)(S400).
- [0092] 그런데, 사용자가 TV(100)를 시청할 경우 TV위치나 공간의 크기 및 재질은 크게 변하지 않지만, 사용자의 움직임에 의해 사용자의 발화 방향, 사용자와 TV간 거리(시청거리), 사용자의 음량, 주변 노이즈 등은 실시간으로 변경될 수 있다.
- [0093] 따라서, 1차 잔향제거 및 마이크 이득조절이 완료되면, 제어부(150)는 변동 환경 정보를 측정한다(S500). 상기 변동 환경정보는 사용자의 TV(100) 시청양태(위치, 방향, 시청거리)에 의해 변경되는 정보로서, 앞에서 설명한 사용자의 발화 방향, 사용자와 TV간 거리, 사용자의 음량, 잡음 음량(주변 노이즈)와 같은 변동 환경변수를 나타낸다.
- [0094] 제어부(180)는 2개의 마이크로 입력되는 음성 시간차를 이용하여 사용자가 발화한 방향을 감지하고, 감지된 방향으로 거리측정센서를 사용하여 TV(100)에서 발화자까지의 거리를 측정할 수 있다. 제어부(150)는 감지된 방향에 따라 2개의 마이크중에서 감지된 방향에 위치하는 마이크는 주 마이크로 사용하고 다른 방향에 있는 마이크는 보조 마이크로 사용할 수 있다. 이때, 주 마이크에는 해당 방향의 음을 집중하여 취득할 수 있는 빔 포밍 기법이 함께 사용될 수 있다. 또한, 제어부(180)는 사용자의 평균 음량을 측정하고 주변 노이즈를 측정하여 잡음 음량을 계산한다.
- [0095] 이와같이 변동 환경변수의 측정이 완료되면 제어부(180)는 TV(100)에서 발화자까지의 거리, 사용자의 평균음량 및 잡음 음량등의 변동 환경변수를 이용하여 최적화된 전처리 잔향 제어, 마이크 이득 조절 및 음성 인식기 언어 모델을 자동으로 조절하여 2차 잔향제거 및 마이크 이득조절을 수행한다(S600).
- [0096] 상기 음성 인식기 언어 모델은 사용자마다 특정 언어에서 발음 및 억양이 상이하기 때문에 각 사용자에게 맞는 발음 및 억양을 학습과정을 통해 축적한 모델로, 예를들면, 경상도 사투리를 쓰는 사용자 A에 대한 음성 인식을 수행할 경우에는 해당 사용자에게 따른 음성 인식기 언어 모델을 자동으로 적용한다. 이때, 사용자 식별은 음성 식별 및 카메라를 통한 얼굴인식을 통해 수행될 수 있다.
- [0097] 상기 2차 잔향제거 및 마이크 이득조절이 완료되면, 제어부(180)는 TV(100)의 동작을 제어하기 위한 사용자의 음성 명령어에 따라 해당 동작을 수행한다(S700, S800).
- [0098] 도 7은 음성 인식 성능을 향상시키기 위하여 초기에 리모컨에 의해 수행되는 고정 환경 정보 측정 및 설정동작을 나타낸 순서도이다.
- [0099] 도 7을 참조하면, 리모컨(200)은 TV(100)로부터 환경정보 측정을 요구하는 제어신호가 수신되면, 제어신호에 따라 거리 측정센서(초음파 또는 레이저)를 동작시켜 TV(100)가 설치된 공간(300)의 크기, TV 설치 위치 및 외부 스피커까지의 거리를 측정한다(S301). 상기 제어신호에는 측정할 항목이 포함되어 있으며, 사용자가 측정항목을 선택하지 않은 경우 모든 항목(디폴트 항목)이 측정된다.

- [0100] 환경 정보가 측정되면 리모컨(200)은 측정된 환경 정보를 환경 정보를 RF(또는 IR) 통신 규격에 따라 TV(100)로 전송하고, 사용자는 리모컨(200)을 이용하여 TV(100)의 제어부(150)가 디스플레이부(1760)에 표시하는 추가 환경정보 설정 메뉴(60)에서 설치 공간(300)의 종류(거실, 방, 공공장소)와 재질(벽지, 대리석, 벽돌)을 선택할 수 있다(S302, S303)(도 6).
- [0101] 도 7에 도시된 고정 환경 정보 측정 및 설정 동작은 음성 인식 초기에 수행되는 동작으로, 일단 음성인식이 수행되어 환경 정보가 저장부(160)에 저장되어 있다면, 수행되지 않는 동작이다.
- [0102] 도 8은 음성 인식 성능을 향상시키기 위해 TV에 의해 수행되는 변동 사용자 정보를 측정 동작을 나타낸 순서도이다.
- [0103] 도 8에 도시된 바와같이, 1차 잔향제거 및 마이크 이득조절이 완료되면(S400), 제어부(150)는 변동 가능한 사용자 정보를 측정한다.
- [0104] 제어부(180)는 2개의 마이크로 입력되는 음성 시간차를 이용하여 사용자가 발화한 방향을 감지하고(S501), 감지된 방향으로 거리측정센서를 사용하여 TV에서 발화자까지의 거리를 측정할 수 있다(S502). 또한, 제어부(180)는 사용자의 평균 음량을 측정하고 주변 노이즈를 측정하여 잡음 음량을 계산한다(S503, S504).
- [0105] 따라서, 상기와 같이 측정된 변동 사용자 정보는 음성 인식 성능을 향상시키기 위한 2차 잔향제거 및 마이크 이득조절에 이용된다.
- [0106] 도 9는 본 발명의 다른 실시예에 따른 영상표시기기의 원거리 음성 인식을 향상 방법을 나타낸 순서도이다.
- [0107] 도 9는 도 1에 의해 음성인식이 한번 수행된 후에 다시 사용자에게 의해 음성인식이 트리거되었을 때 수행되는 동작을 나타낸다. 즉, 사용자에게 의해 음성인식이 트리거되면(S100), 제어부(180)는 저장부(160)에 기측정된 고정 환경 정보가 존재하는지 체크한다(S900). 체크결과, 저장부(160)에 기측정된 고정 환경 정보가 존재하지 않으면, 전술한 단계(S200~S400)를 수행하여 1차잔향제거 및 마이크 이득조절을 수행한다.
- [0108] 반면에, 저장부(160)에 기측정된 고정 환경 정보가 존재하면, 제어부(150)는 바로 전술한 단계(S500)를 수행하여 변동 환경정보 설정 및 2차잔향제거 및 마이크 이득조절을 수행한다.
- [0109] 다른 실시예로 제어부(150)는 저장부(160)에 기측정된 고정 환경 정보가 존재하더라도 디스플레이부(170)에 고정 환경정보의 측정 여부를 묻는 팝-업을 표시할 수도 있다.
- [0110] 도 10은 본 발명의 실시예에 따른 영상표시기기의 근거리 음성 인식을 향상 방법을 나타낸 순서도이다.
- [0111] 도 10에 도시된 실시예는 리모컨(200)에 내장된 마이크를 통해 사용자의 음성을 인식하여 TV(100)로 전송하는 예에 적용된다.
- [0112] 사용자의 트리거 키워드가 인식되면 도 2에 도시된 리모컨(200)의 제어부(270)는 내장된 거리측정 센서(미도시)를 이용하여 사용자와 리모컨간의 거리를 측정한다.
- [0113] 제어부(270)는 측정된 거리에 따라 내장된 마이크의 이득(증폭 이득)을 조절함으로써 음성인식 성능을 향상시킬 수 있다.
- [0114] 본 발명은 환경정보를 음성인식 성능 향상에 적용하였으나 이에 한정되지 않고 최적의 음향, 화질, 3D의 깊이조절 및 커브드 TV의 곡률 반경등의 자동 설정에 활용될 수 있다.
- [0115] 상술한 바와같이 본 발명은 한번 설치되면 위치가 잘 변경되지 않는 영상표시기기(e.g., TV)의 특성을 감안하여 영상표시기기가 설치된 환경(설치 공간의 크기, 재질, 설치 위치, 외부 스피커의 위치등) 및 사용자의 위치에 따라 자동으로 전처리 잔향제거, 마이크 게인 조절, 음성인식기 언어 모델(Language model)을 자동으로 설정함으로써 원거리 인식 성능을 향상시킬 수 있다.
- [0116] 그리고, 본 발명은 사용자와 영상표시기기간 거리에 따라 전처리 잔향제거 및 마이크 게인 조절을 수행함으로써 먼 거리 또는 사용자의 목소리가 작은 경우에도 원거리 인식 성능을 향상시킬 수 있다.
- [0117] 또한, 본 발명은 음성인식을 위해 측정된 사용 환경을 이용하여 시청자에 맞는 최적의 음향, 화질, 3D 깊이감 및 커브드 TV의 곡률반경등의 자동 설정활용에 이용 가능하다.
- [0118] 본 발명의 일 실시예에 의하면, 전술한 방법은, 프로그램이 기록된 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장

되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 매체의 예로는, ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어, 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한, 상기 컴퓨터는, 단말기의 제어부를 포함할 수도 있다.

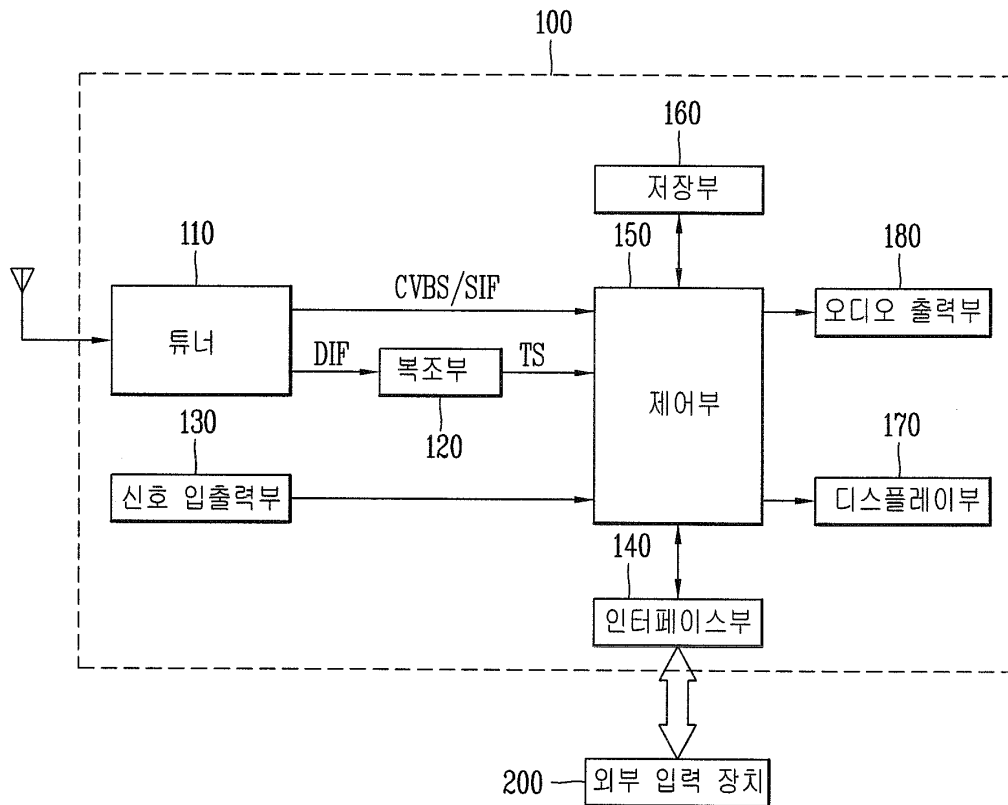
[0119] 상기와 같이 설명된 영상표시기기 및 그의 원거리 음성 인식을 향상 방법은 상기 설명된 실시예들의 구성과 방법이 한정되게 적용될 수 있는 것이 아니라, 상기 실시예들은 다양한 변형이 이루어질 수 있도록 각 실시예들의 전부 또는 일부가 선택적으로 조합되어 구성될 수도 있다.

부호의 설명

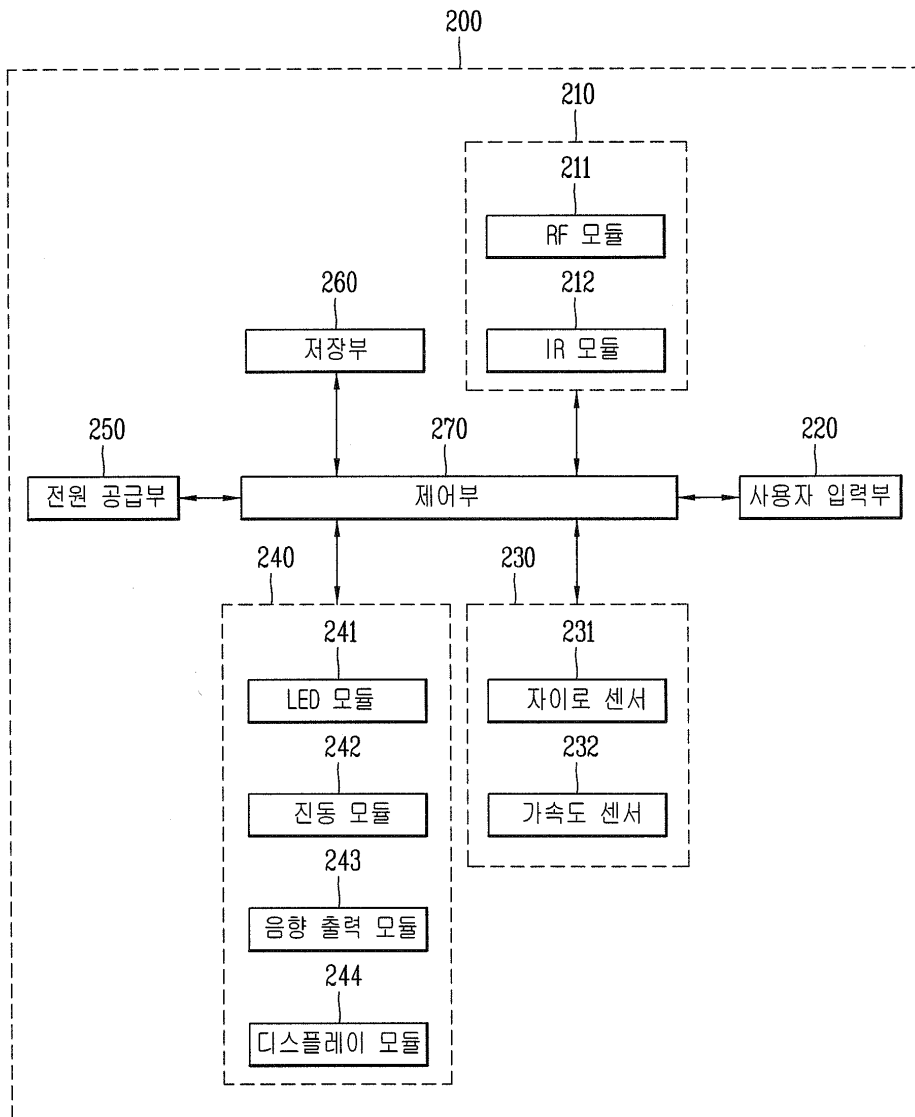
- [0120] 50 : 지문 스캐너
- 60 : 설정 메뉴
- 100 : 영상표시기기
- 140 : 인터페이스부
- 150 : 제어부
- 160 : 저장부
- 170 : 디스플레이부
- 200 : 외부 입력장치
- 300 : 설치 공간

도면

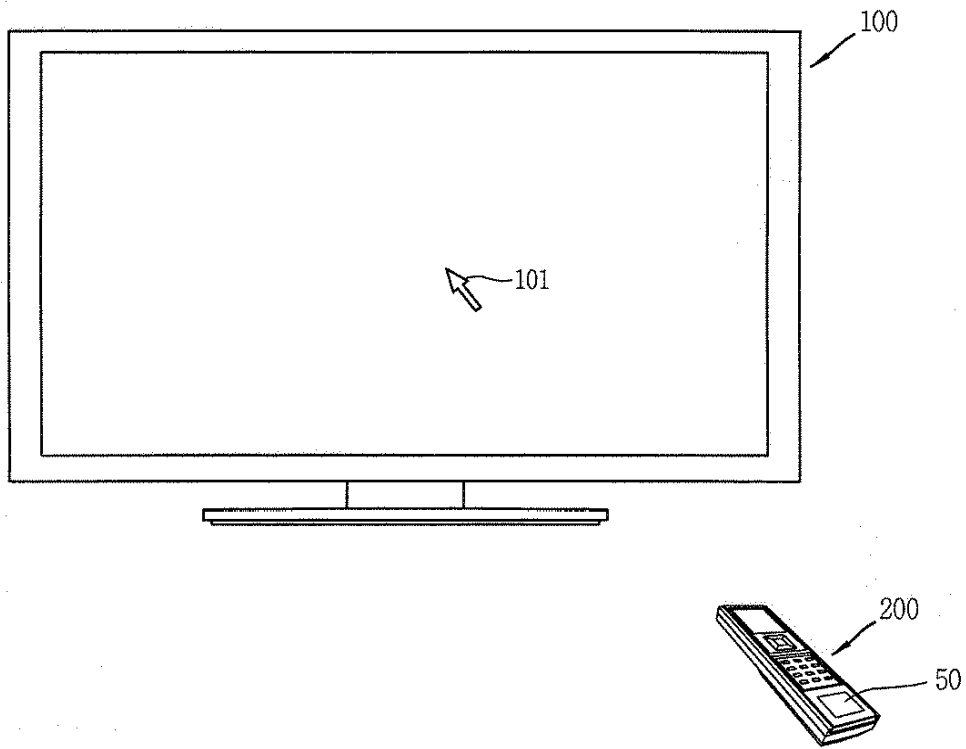
도면1



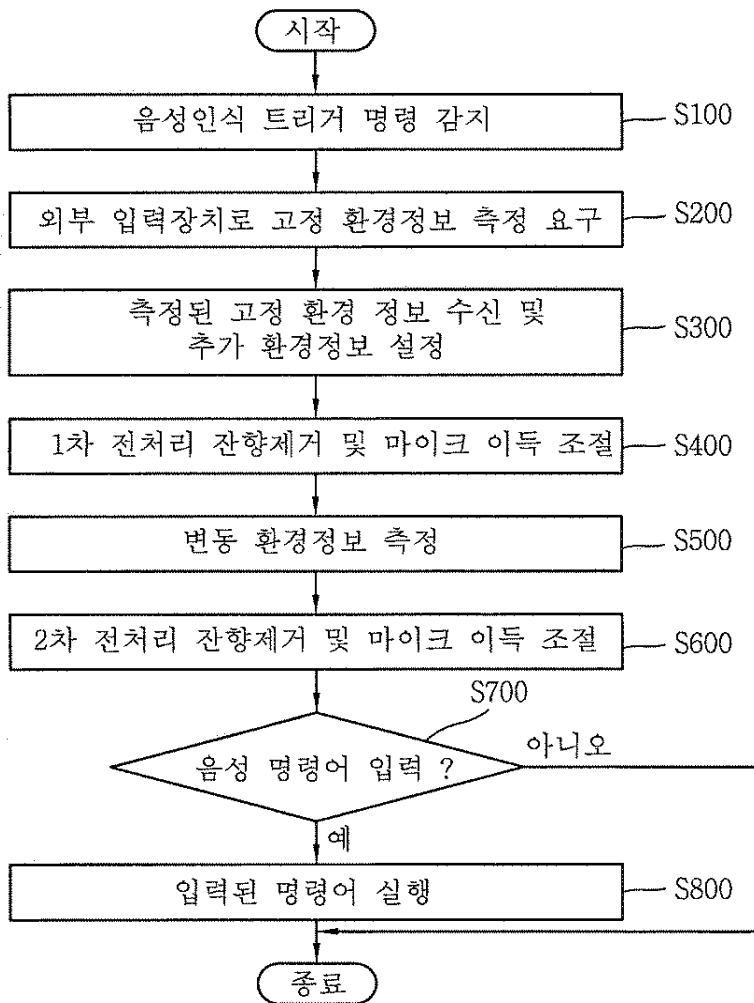
도면2



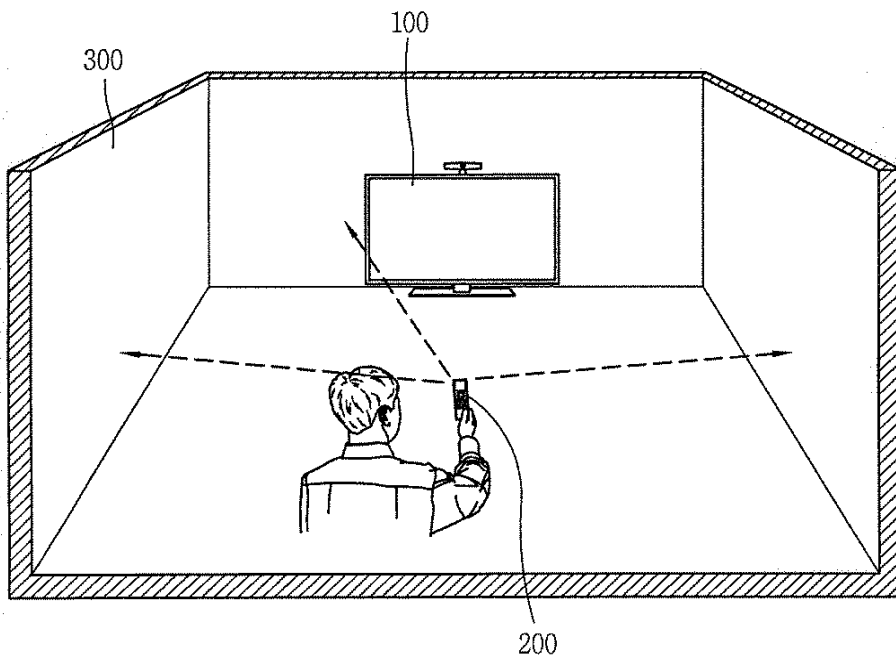
도면3



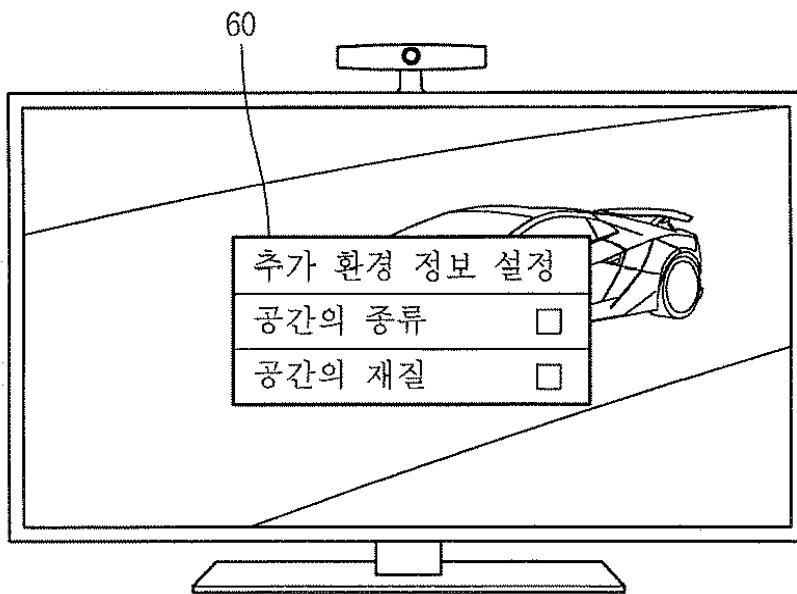
도면4



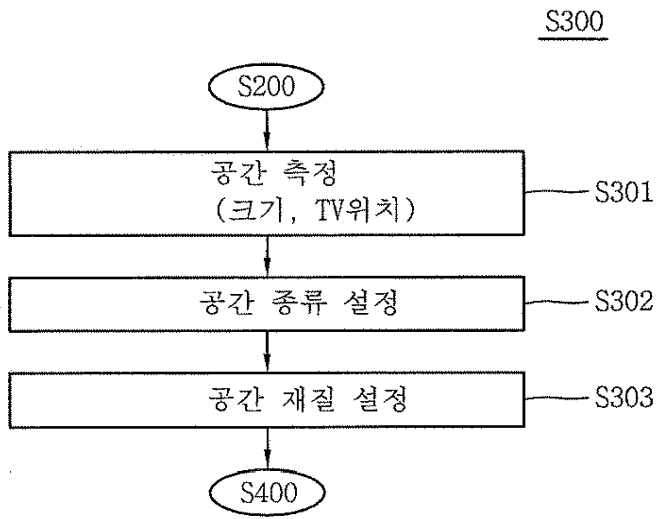
도면5



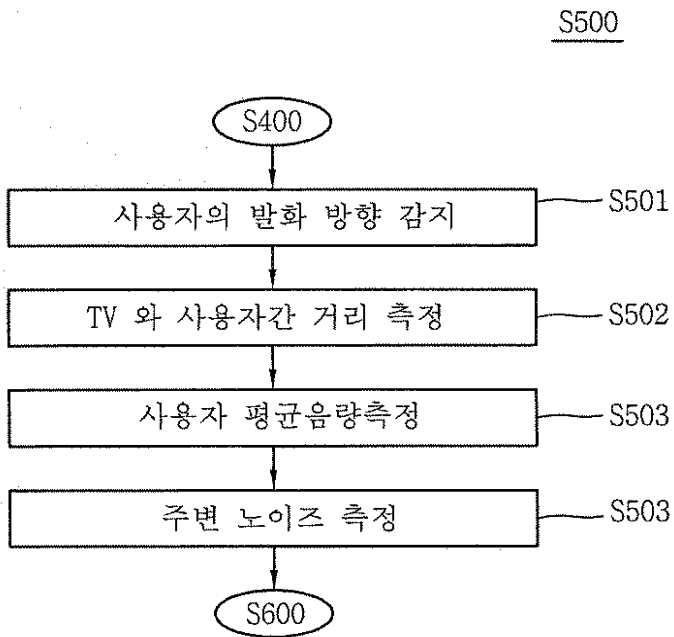
도면6



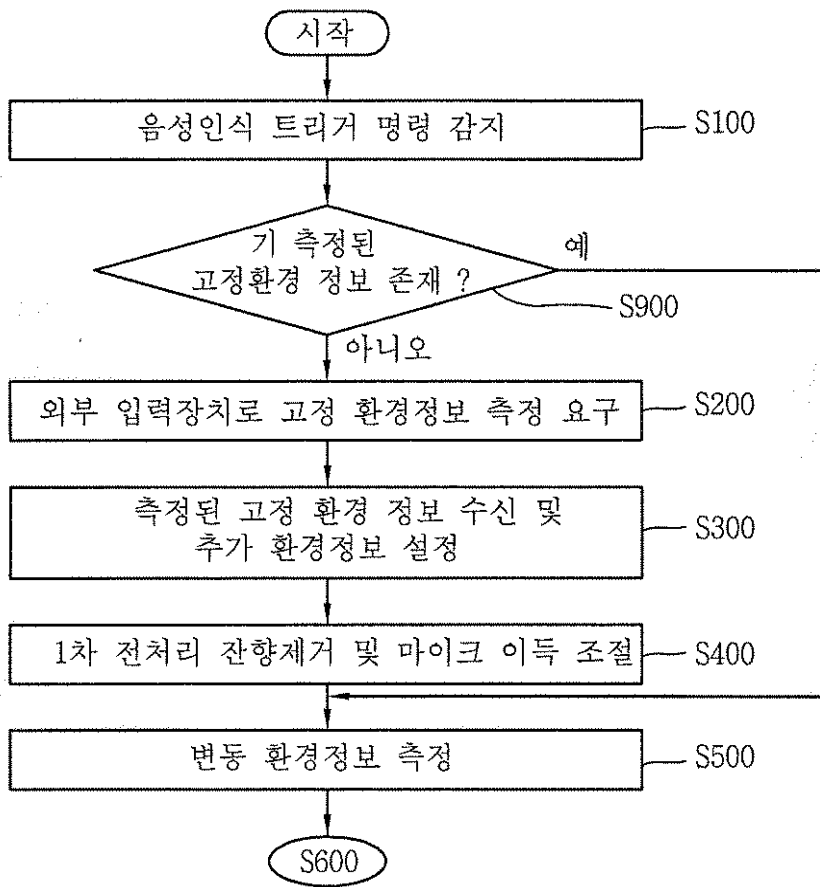
도면7



도면8



도면9



도면10

