



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2017년04월13일  
 (11) 등록번호 10-1726738  
 (24) 등록일자 2017년04월07일

(51) 국제특허분류(Int. Cl.)  
*G10L 15/10* (2006.01) *G10L 15/20* (2006.01)  
 (21) 출원번호 10-2010-0121309  
 (22) 출원일자 2010년12월01일  
 심사청구일자 2015년10월06일  
 (65) 공개번호 10-2012-0059837  
 (43) 공개일자 2012년06월11일  
 (56) 선행기술조사문헌  
 JP2010136080 A\*  
 US20020173864 A1\*  
 \*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
**삼성전자주식회사**  
 경기도 수원시 영통구 삼성로 129 (매탄동)  
 (72) 발명자  
**김기범**  
 경기도 성남시 분당구 미금일로86번길 20 (구미동)  
 (74) 대리인  
**특허법인세림**

전체 청구항 수 : 총 15 항

심사관 : 정성윤

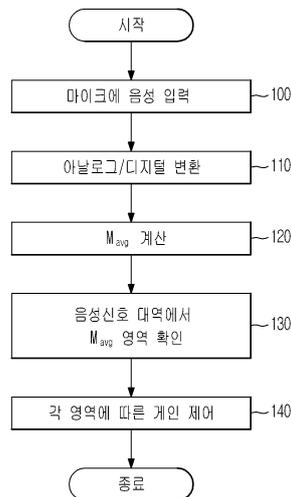
(54) 발명의 명칭 **음성처리장치 및 그 방법**

**(57) 요약**

마이크로 입력되는 신호의 특성에 기반해 계인을 조절하여 음성을 처리하는 음성처리장치 및 그 방법을 개시한다.

음성처리장치는 음성 신호를 입력받는 마이크와, 마이크에서 출력되는 아날로그 음성 신호를 디지털 음성 신호로 변환하는 아날로그/디지털 변환부와, 아날로그/디지털 변환부로부터 출력되는 음성 신호의 복수 개의 프레임의 평균값을 산출하고, 복수 개의 프레임의 평균값이 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역 상의 어느 영역에 위치하는지 확인하고, 음성 신호가 처리될 수 있는 최적의 영역의 크기를 가지도록 복수 개의 프레임의 평균값이 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 자동 계인 제어부를 포함하므로, 음원의 거리에 구속받지 않고 음성 인식을 극대화할 수 있다.

**대표도** - 도4



**명세서**

**청구범위**

**청구항 1**

음성 신호를 입력받는 마이크;

상기 마이크에서 출력되는 아날로그 음성 신호를 디지털 음성 신호로 변환하는 아날로그/디지털 변환부; 및

상기 아날로그/디지털 변환부로부터 출력되는 음성 신호의 복수 개의 프레임의 평균값을 산출하고, 상기 복수 개의 프레임의 평균값이 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역 상의 어느 영역에 위치하는지 확인하고, 상기 음성 신호가 처리될 수 있는 최적의 영역의 크기를 가지도록 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 자동 계인 제어부를 포함하되,

상기 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역은,

상기 음성의 세기에 따라 상기 마이크에서 감지할 수 있는 최저 한계 크기, 상기 마이크에서 감지할 수 있는 최고 한계 크기, 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기 및 하이 크기, 상기 음성 신호 대역의 중간 지점을 경계로 하여 복수 개의 영역으로 구분되는 것인 음성처리장치.

**청구항 2**

삭제

**청구항 3**

제 1 항에 있어서,

상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 작으면 상기 계인을 높여 입력되는 음성 신호를 증폭시키는 음성처리장치.

**청구항 4**

제 3 항에 있어서,

상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 작아 계인을 높여 입력되는 음성 신호를 증폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에는 조정된 계인을 적용하지 않는 음성처리장치.

**청구항 5**

제 1 항에 있어서,

상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 크지만 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기보다 작으면 상기 계인을 높여 입력되는 음성 신호를 증폭시키는 음성처리장치.

**청구항 6**

제 5 항에 있어서,

상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 크지만 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기보다 작아 상기 계인을 높여 입력되는 음성 신호를 증폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에 조정된 계인을 적용하는 음성처리장치.

**청구항 7**

제 1 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 존재하면 상기 게인을 조정하지 않는 음성처리장치.

**청구항 8**

제 1 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 존재하면 상기 음성 신호의 복수 개의 프레임의 평균값에 게인이 적용될 경우 상기 음성 신호 대역의 중간 지점에 세기가 맞춰지도록 상기 게인을 조정하는 음성처리장치.

**청구항 9**

제 1 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 하이크기보다 크고 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 작으면 상기 게인을 낮춰 입력되는 음성 신호를 감폭시키는 음성처리장치.

**청구항 10**

제 9 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 하이크기보다 크고 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 작아 상기 게인을 낮춰 입력되는 음성 신호를 감폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에 조정된 게인을 적용하는 음성처리장치.

**청구항 11**

제 1 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 크면 상기 게인을 낮춰 입력되는 음성 신호를 감폭시키는 음성처리장치.

**청구항 12**

제 11 항에 있어서,

상기 자동 게인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 커서 상기 게인을 낮춰 입력되는 음성 신호를 감폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에는 조정된 게인을 적용하지 않는 음성처리장치.

**청구항 13**

제 1 항에 있어서,

상기 자동 게인 제어부로부터 디지털 신호를 입력받아 음성신호로 처리하는 음성처리부를 더 포함하며,

상기 자동 게인 제어부는 상기 음성처리부의 음성 크기 변형 게인을 파라미터로 하여 상기 음성 신호 대역을 복수 개의 영역으로 구분하는 음성처리장치.

**청구항 14**

마이크로부터 아날로그 음성 신호를 입력받아 디지털 음성 신호로 변환하고;

상기 디지털 음성 신호의 복수 개의 프레임의 평균값을 산출하고;

상기 복수 개의 프레임의 평균값이 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역 상의 어느 영역에 위치하는지 확인하고;

상기 음성 신호가 처리될 수 있는 최적의 영역의 크기를 가지도록 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하되,

상기 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역은,

상기 음성의 세기에 따라 상기 마이크에서 감지할 수 있는 최저 한계 크기, 상기 마이크에서 감지할 수 있는 최고 한계 크기, 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기 및 하이 크기, 상기 음성 신호 대역의 중간 지점을 경계로 하여 복수 개의 영역으로 구분되는 것인 음성처리방법.

**청구항 15**

삭제

**청구항 16**

제 14 항에 있어서,

상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 것은,

상기 복수 개의 프레임의 평균값에 조정된 계인을 적용할 경우 그 세기가 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 위치하도록 상기 계인을 조정하는 것인 음성처리방법.

**청구항 17**

제 14 항에 있어서,

상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 것은,

상기 복수 개의 프레임의 평균값에 조정된 계인을 적용할 경우 그 세기가 상기 음성 신호 대역의 중간 지점에 위치하도록 상기 계인을 조정하는 것인 음성처리방법.

**발명의 설명**

**기술 분야**

[0001] 마이크의 계인을 자동으로 조정할 수 있는 음성처리장치 및 그 방법에 관한 것이다.

**배경 기술**

[0002] 최근에는 음성 인식에 대한 소비자의 요구 증대에 따라 음성인식을 위한 마이크가 부착된 로봇 제품이 개발되어 시판되고 있다. 음성 인식 수단을 구비한 로봇은 음성 센서에 의해 감지된 음성이 오디오 주파수 대역 또는 고 주파수 대역을 가질 경우에 그 음성에 의해 정해진 한가지 동작을 수행하게 된다.

[0003] 로봇의 음성인식수단은 고정이득이나 자동이득이 적용된 마이크와 그 마이크를 통해 입력되는 음원데이터를 저장할 버퍼와, 저장된 음원데이터를 신호처리할 음성처리부로 구성되며 음성인식에 따라 로봇이 동작한다.

[0004] 한편, 마이크로 입력되는 음성의 크기는 마이크와 화자와의 거리 및 화자의 발화의 크기에 따라 정해진다. 대부분의 마이크는 마이크와 화자와의 거리에 따라 계인을 조정하나, 화자는 마이크와의 거리가 가까우면 작게 발성하려고 하고, 마이크와의 거리가 멀면 크게 발성하려고 하기 때문에 거리만을 고려해 계인을 조정하는 것은 문제가 있다. 즉, 마이크와 화자와의 거리에 따라 계인을 조정하는 것은 화자의 발성 상태를 고려하지 않는다는 문제가 있다.

**발명의 내용**

[0005] 본 발명의 일측면은 마이크로 입력되는 신호의 특성에 기반해 계인을 조절하는 방식을 제공한다.

[0006] 이를 위한 본 발명의 일실시예에 의한 음성처리장치는 음성 신호를 입력받는 마이크;와 상기 마이크에서 출력되는 아날로그 음성 신호를 디지털 음성 신호로 변환하는 아날로그/디지털 변환부; 및 상기 아날로그/디지털 변환부로부터 출력되는 음성 신호의 복수 개의 프레임의 평균값을 산출하고, 상기 복수 개의 프레임의 평균값이 음

성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역 상의 어느 영역에 위치하는지 확인하고, 상기 음성 신호가 처리될 수 있는 최적의 영역의 크기를 가지도록 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 자동 계인 제어부를 포함할 수 있다.

- [0007] 상기 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역은, 상기 음성의 세기에 따라 상기 마이크에서 감지할 수 있는 최저 한계 크기, 상기 마이크에서 감지할 수 있는 최고 한계 크기, 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기 및 하이 크기, 상기 음성 신호 대역의 중간 지점을 경계로 하여 복수 개의 영역으로 구분될 수 있다.
- [0008] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 작으면 상기 계인을 높여 입력되는 음성 신호를 증폭시킬 수 있다.
- [0009] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 작아 계인을 높여 입력되는 음성 신호를 증폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에는 조정된 계인을 적용하지 않을 수 있다.
- [0010] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 크지만 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기보다 작으면 상기 계인을 높여 입력되는 음성 신호를 증폭시킬 수 있다.
- [0011] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호 대역에 설정된 상기 최저 한계 크기보다 크지만 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기보다 작아 상기 계인을 높여 입력되는 음성 신호를 증폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에 조정된 계인을 적용할 수 있다.
- [0012] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 존재하면 상기 계인을 조정하지 않을 수 있다.
- [0013] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 존재하면 상기 음성 신호의 복수 개의 프레임의 평균값에 계인이 적용될 경우 상기 음성 신호 대역의 중간 지점에 세기가 맞춰지도록 상기 계인을 조정할 수 있다.
- [0014] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 하이 크기보다 크고 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 작으면 상기 계인을 낮춰 입력되는 음성 신호를 감폭시킬 수 있다.
- [0015] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 음성 신호가 처리될 수 있는 최적의 영역의 하이 크기보다 크고 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 작아 상기 계인을 낮춰 입력되는 음성 신호를 감폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에 조정된 계인을 적용할 수 있다.
- [0016] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 크면 상기 계인을 낮춰 입력되는 음성 신호를 감폭시킬 수 있다.
- [0017] 상기 자동 계인 제어부는 상기 음성 신호의 복수 개의 프레임의 평균값이 상기 마이크에서 감지할 수 있는 최고 한계 크기보다 커서 상기 계인을 낮춰 입력되는 음성 신호를 감폭시키는 경우 상기 음성 신호의 복수 개의 프레임의 평균값을 계산하는데 사용된 프레임에는 조정된 계인을 적용하지 않을 수 있다.
- [0018] 상기 자동 계인 제어부로부터 디지털 신호를 입력받아 음성신호로 처리하는 음성처리부를 더 포함하며, 상기 자동 계인 제어부는 상기 음성처리부의 음성 크기 변형 계인을 파라미터로 하여 상기 음성 신호 대역을 복수 개의 영역으로 구분할 수 있다.
- [0019] 그리고, 본 발명의 일실시예에 의한 음성처리방법은 아날로그 음성 신호를 입력받아 디지털 음성 신호로 변환하고, 상기 디지털 음성 신호의 복수 개의 프레임의 평균값을 산출하고, 상기 복수 개의 프레임의 평균값이 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역 상의 어느 영역에 위치하는지 확인하고, 상기 음성 신호가 처리될 수 있는 최적의 영역의 크기를 가지도록 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정할 수 있다.
- [0020] 상기 음성의 세기에 따라 복수 개의 영역으로 구분된 음성 신호 대역은, 상기 음성의 세기에 따라 상기 마이크

에서 감지할 수 있는 최저 한계 크기, 상기 마이크에서 감지할 수 있는 최고 한계 크기, 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기 및 하이 크기, 상기 음성 신호 대역의 중간 지점을 경계로 하여 복수 개의 영역으로 구분될 수 있다.

[0021] 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 것은, 상기 복수 개의 프레임의 평균값에 조정된 계인을 적용할 경우 그 세기가 상기 음성 신호가 처리될 수 있는 최적의 영역의 로우 크기와 하이 크기 사이에 위치하도록 상기 계인을 조정할 수 있다.

[0022] 상기 복수 개의 프레임의 평균값이 상기 음성 신호 대역 상의 어느 위치에 존재하는지 여부에 따라 계인을 조정하는 것은, 상기 복수 개의 프레임의 평균값에 조정된 계인을 적용할 경우 그 세기가 상기 음성 신호 대역의 중간 지점에 위치하도록 상기 계인을 조정할 수 있다.

[0023] 한편, 상술한 본 발명의 일측면에 의하면 음원의 거리에 구속받지 않고 음성 인식율을 극대화할 수 있다.

**도면의 간단한 설명**

[0024] 도 1은 본 발명의 일실시예에 의한 음성처리장치의 기능블록도

도 2는 본 발명의 일실시예에 의한 음성처리장치의 마이크의 음성 신호의 크기 영역을 도시한 도면

도 3은 본 발명의 다른 실시예를 설명하기 위해 도 2의 마이크의 음성 신호의 크기 영역의 중간 지점을 도시한 도면

도 4는 본 발명의 일실시예에 의한 음성처리장치의 제어흐름도

도 5는 본 발명의 일실시예에 의한 음성처리장치의 계인 조정을 설명하기 위한 제어흐름도

**발명을 실시하기 위한 구체적인 내용**

[0025] 이하에서는 첨부도면을 참조하여 본 발명에 대해 상세히 설명한다.

[0026] 도 1은 본 발명의 일실시예에 의한 음성처리장치의 기능블록도이고, 도 2는 본 발명의 일실시예에 의한 음성처리장치의 마이크의 음성 신호의 크기 영역을 도시한 것이며, 도 3은 본 발명의 다른 실시예를 설명하기 위해 도 2의 마이크의 음성 신호의 크기 영역의 중간 지점을 도시한 도면이다.

[0027] 음성처리장치(1)는 마이크(10), 아날로그/디지털 변환부(20), 자동 계인 제어부(30), 음성처리부(40) 및 음성인식부(50)를 포함할 수 있다.

[0028] 마이크(10)는 화자로부터 전송되는 음성신호를 수신한다.

[0029] 아날로그/디지털 변환부(20)는 마이크(10)를 통해 입력되는 아날로그 음성 신호를 디지털 음성 신호로 전환한다.

[0030] 자동 계인 제어부(30)는 디지털 음성 신호의 계인을 조절한다. 자동 계인 제어부(30)는 잡음 제거기(미도시) 또는 음성 인식기(미도시)와 같은 특정 용도에 맞게 블록 처리된 음성처리부(40)에 계인이 조절된 디지털 음성 신호를 전송한다.

[0031] 자동 계인 제어부(30)는 마이크(10)로 직접 입력되는 음성 신호의 특성에 기반하여 계인을 조절한다. 자동 계인 제어부(30)는 음성처리부(40)에서 음성이 처리되기에 적합한 영역으로 음성의 크기를 맞추어 전송한다.

[0032] 도 2를 참조하면, 자동 계인 제어부(30)는 음성의 세기에 따라 음성 신호 대역을 복수 개의 영역으로 분리한다. 자동 계인 제어부(30)는 음성 신호의 크기를 5개의 영역으로 설정할 수 있다.

[0033] 제 1 영역 :  $M_{avg} \leq L_{lower}$

[0034] 제 2 영역 :  $L_{lower} < M_{avg} \leq M_{min}$

[0035] 제 3 영역 :  $M_{min} < M_{avg} \leq M_{max}$

[0036] 제 4 영역 :  $M_{max} < M_{avg} \leq L_{upper}$

[0037] 제 5 영역 :  $M_{avg} > L_{upper}$

- [0038]  $M_{avg}$ 는 마이크(10)로 입력되는 음성 신호의 프레임 단위의 크기를 평균한 값이며(단, 프레임의 개수는 미리 정해짐),  $L_{lower}$ 는 마이크(10)에서 감지할 수 있는 다이내믹 레인지(Dynamic Range)의 최저 한계 크기이며,  $L_{upper}$ 는 마이크(10)에서 감지할 수 있는 다이내믹 레인지(Dynamic Range)의 최고 한계 크기이며,  $M_{min}$ 과  $M_{max}$ 는 다이내믹 레인지(Dynamic Range)내에서 신호가 처리될 수 있는 최적의 영역의 로우값 및 하이값을 의미한다. 한편,  $L_{lower}$  및  $L_{upper}$ 는 마이크(10)의 종류 및 사양에 따라 결정되는 값이며,  $M_{min}$ 과  $M_{max}$ 는 후술하는 것처럼 자동 게인 제어부(30)에 연결되는 음성처리부(40)에 따라 유동적으로 변화될 수 있는 값이다.
- [0039] 자동 게인 제어부(30)는 입력되는 음성 신호의 크기를 프레임 단위로 계산하여 평균하고, 평균값( $M_{avg}$ )이 위치하는 영역을 확인하여 게인을 조정한다. 자동 게인 제어부(30)는 입력되는 음성 신호의 크기의 평균값에 조정된 게인을 적용하여 그 영역이 다이내믹 레인지의 제 3 영역에 위치하도록 제어한다.
- [0040] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 작아서 제 1 영역에 있다고 판단되면, 게인을 높이고 이후에 입력되는 신호에 조정된 게인을 적용하여 증폭시킨다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.
- [0041] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 크지만 다이내믹 레인지(Dynamic Range)내의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )보다 작아서 제 2 영역에 있다고 판단되면, 게인을 높이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 제 2 영역에 있는 음성 신호는 조정된 게인을 적용하여 최적의 영역 즉, 제 3 영역에 위치하는 신호로 변경가능하기 때문에, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.
- [0042] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )과 신호가 처리될 수 있는 최적의 영역의 하이값( $M_{max}$ ) 사이에 위치하게 되면 최적의 영역에 위치하는 것이므로 게인 조정없이 음성처리부(40)에 전송한다.
- [0043] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최적의 영역의 하이값( $M_{max}$ )보다 크고 최고 한계 크기( $L_{upper}$ )보다 작아서 제 4 영역에 있다고 판단되면, 게인을 줄이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 제 4 영역에 있는 음성 신호는 조정된 게인을 적용하여 최적의 영역 즉, 제 3 영역에 위치하는 신호로 변경가능하기 때문에, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.
- [0044] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 최고 한계 크기( $L_{upper}$ )보다 커서 제 5 영역에 있다고 판단되면, 게인을 낮추고 이후에 입력되는 신호에 조정된 게인을 적용하여 증폭시킨다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.
- [0045] 음성처리부(40)는 자동 게인 제어부(30)로부터 다수의 디지털 신호를 입력받아 음성신호로 처리한다.
- [0046] 음성인식부(50)는 음성처리부(40)에서 출력되는 신호에 의해 음성신호를 인식할 수 있다.
- [0047] 도 3을 참조하면, 자동 게인 제어부(30)는 음성 신호의 크기 영역을 5개의 영역으로 나누고, 음성 신호의 크기 영역의 중간 지점을 추가적으로 설정한다.
- [0048] 제 1 영역 :  $M_{avg} \leq L_{lower}$
- [0049] 제 2 영역 :  $L_{lower} < M_{avg} \leq M_{min}$
- [0050] 제 3 영역 :  $M_{min} < M_{avg} \leq M_{max}$
- [0051] 제 4 영역 :  $M_{max} < M_{avg} \leq L_{upper}$

- [0052] 제 5 영역 :  $M_{avg} > L_{upper}$
- [0053] 중간 지점 :  $L_{med}$
- [0054]  $M_{avg}$ 는 마이크(10)로 입력되는 음성 신호의 프레임 단위의 크기를 평균한 값이며(단, 프레임의 개수는 미리 정해짐),  $L_{lower}$ 는 마이크(10)에서 감지할 수 있는 다이내믹 레인지(Dynamic Range)의 최저 한계 크기이며,  $L_{upper}$ 는 마이크(10)에서 감지할 수 있는 다이내믹 레인지(Dynamic Range)의 최고 한계 크기이며,  $M_{min}$ 과  $M_{max}$ 는 다이내믹 레인지(Dynamic Range)내에서 신호가 처리될 수 있는 최적의 영역의 로우값 및 하이값을 의미하며,  $L_{med}$ 는 음성 신호 대역의 중간지점이다.
- [0055] 자동 게인 제어부(30)는 입력되는 음성 신호의 크기를 프레임 단위로 계산하여 평균하고, 평균값( $M_{avg}$ )이 위치하는 영역을 확인하여 게인을 조정한다. 자동 게인 제어부(30)는 입력되는 음성 신호의 크기의 평균값에 조정된 게인을 적용하여 그 영역이 다이내믹 레인지의 음성 신호 대역의 중간지점( $L_{med}$ )에 위치하도록 제어한다.
- [0056] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 작아서 제 1 영역에 있다고 판단되면, 게인을 높이고 이후에 입력되는 신호에 조정된 게인을 적용하여 증폭시킨다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.
- [0057] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 크지만 다이내믹 레인지(Dynamic Range)내의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )보다 작아서 제 2 영역에 있다고 판단되면, 게인을 높이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 제 2 영역에 있는 음성 신호는 조정된 게인을 적용하여 다이내믹 레인지의 음성 신호 대역의 중간지점( $L_{med}$ )에 위치하는 신호로 변경가능하기 때문에, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.
- [0058] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )과 신호가 처리될 수 있는 최적의 영역의 하이값( $M_{max}$ ) 사이에 위치하게 되면 다이내믹 레인지의 음성 신호 대역의 중간지점( $L_{med}$ )으로 신호의 크기를 조정할 수 있도록 게인을 조정한다.
- [0059] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최적의 영역의 하이값( $M_{max}$ )보다 크고 최고 한계 크기( $L_{upper}$ )보다 작아서 제 4 영역에 있다고 판단되면, 게인을 줄이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 제 4 영역에 있는 음성 신호는 조정된 게인을 적용하여 다이내믹 레인지의 음성 신호 대역의 중간지점( $L_{med}$ )에 위치하는 신호로 변경가능하기 때문에, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.
- [0060] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 최고 한계 크기( $L_{upper}$ )보다 커서 제 5 영역에 있다고 판단되면, 게인을 낮추고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.
- [0061] 음성처리부(40)는 자동 게인 제어부(30)로부터 다수의 디지털 신호를 입력받아 음성신호로 처리한다.
- [0062] 음성인식부(50)는 음성처리부(40)에서 출력되는 신호에 의해 음성신호를 인식할 수 있다.
- [0063] 한편, 자동 게인 제어부(30)는 음성처리부(40)의 기능 및 사양을 고려해  $M_{min}$ 과  $M_{max}$ 를 정할 수 있다. 자동 게인 제어부(30)는 수식 1과 같이  $M_{min}$ 과  $M_{max}$ 를 정할 때 음성처리부(40)의 음성 크기 변형 게인( $A_{max}$ ,  $A_{min}$ )을 고려하여 정할 수 있다.
- [0064] 수식 1

- [0065]  $M_{max} = k * L_{med} * A_{max}$
- [0066]  $M_{min} = k * L_{med} * A_{min}$
- [0067] 여기서,  $M_{min}$ 과  $M_{max}$ 는 다이내믹 레인지(Dynamic Range)내에서 신호가 처리될 수 있는 최적의 영역의 로우값 및 하이값이며,  $k$  ( $k \geq 1$ )는 실험에 의해 설계자가 정하는 임의의 상수이고,  $L_{med}$ 는 다이내믹 레인지의 음성 신호 대역의 중간지점이고,  $A_{max}$ ,  $A_{min}$ 는 음성처리부(40)의 음성 크기 변형 계인으로서 각각 최대 음성 크기 변형 계인, 최소 음성 크기 변형 계인이다. 즉, 음성처리부(40)는 입력되는 음성 신호의 크기를 변형시킬 수 있으며, 크기를 변형시키기 위해 사용되는 최대 계인( $A_{max}$ )과 최소 계인( $A_{min}$ )을 나타낸다.
- [0068] 수식 1의 의미를 살펴보면, 음성처리부(40)에 입력된 신호의 세기( $L_{med}$ )에 음성처리부(40)의 최대 계인( $A_{max}$ )을 적용하여도  $M_{max}$ 보다 높지 않도록 하고, 음성처리부(40)에 입력된 신호의 세기( $L_{med}$ )에 음성처리부(40)의 최소 계인( $A_{min}$ )을 적용하여도  $M_{min}$ 보다 낮지 않도록 한다. 즉, 음성처리부(40)에 전송된 신호는 음성처리부(40)의 계인이 적용되어도 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )과 신호가 처리될 수 있는 최적의 영역의 하이값( $M_{max}$ ) 사이의 크기에 있게 된다.
- [0069] 그리고, 자동 계인 제어부(30)는 다음 수식 2의 조건에 만족하도록  $M_{min}$ 과  $M_{max}$ 를 정한다.
- [0070] 수식 2
- [0071]  $A_{max} * M_{max} < L_{upper}$
- [0072]  $A_{min} * M_{min} > L_{lower}$
- [0073] 수식 2의 의미를 살펴보면, 다이내믹 레인지(Dynamic Range)내에서 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ ) 및 하이값( $M_{max}$ )이 음성처리부(40)에 전송된 후, 계인( $A_{max}$ ,  $A_{min}$ )이 적용되어 변환되어도 마이크(10)가 음성을 인식할 수 있는 한계영역을 벗어나지 않도록 하기 위함이다. 여기서, 한계영역을 벗어난다는 것은 최저 한계 크기( $L_{lower}$ )보다 작거나 최고 한계 크기( $L_{upper}$ )보다 큰 것을 의미한다.
- [0074] 도 4는 본 발명의 일실시예에 의한 음성처리장치의 제어흐름도이다.
- [0075] 마이크(10)에 아날로그 음성 신호가 입력되면, 아날로그 음성 신호를 디지털 음성 신호로 변환한다.(100,110)
- [0076] 다음으로, 자동 계인 제어부(30)는 아날로그/디지털 변환부(20)로부터 디지털 음성 신호가 전송되면  $M_{avg}$ 를 계산한다.  $M_{avg}$ 는 마이크(10)로 입력되는 음성 신호의 복수 개 프레임의 크기를 평균한 값이다.(120)
- [0077] 다음으로, 자동 계인 제어부(30)는 계산된  $M_{avg}$ 가 미리 정해진 음성 신호 영역 중 어느 영역에 속하는지 확인한다. 음성 신호 영역은 복수 개의 영역으로 정해지며, 본 실시예에서는 도 2 또는 도 3에 도시한 것과 같이 5개의 영역이 있는 것을 그 예로 든다.(130)
- [0078] 다음으로, 자동 계인 제어부(30)는  $M_{avg}$ 가 속하는 영역에 따른 계인 조절을 수행한다.(140)
- [0079] 도 5는 본 발명의 일실시예에 의한 음성처리장치의 계인 조절을 설명하기 위한 제어흐름도이다.
- [0080] 자동 계인 제어부(30)는 입력되는 음성 신호의  $M_{avg}$ 를 계산한다.(200)
- [0081] 자동 계인 제어부(30)는 계산된  $M_{avg}$ 가  $L_{lower}$ 보다 큰지 확인한다. 자동 계인 제어부(30)는  $M_{avg}$ 가  $L_{lower}$ 보다 작으면 계인을 상승시킨다. 즉, 자동 계인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 작다고 판단되면, 계인을 높이고 이후에 입력되는 신호에 조정된 계인을 적용하여 증폭시킨다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.(210,220,230)
- [0082] 자동 계인 제어부(30)는  $M_{avg}$ 가  $L_{lower}$ 보다 크면  $M_{min}$ 보다 큰지 확인한다. 자동 계인 제어부(30)는 입력되는 음성

신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최저 한계 크기( $L_{lower}$ )보다 크지만 다이내믹 레인지(Dynamic Range)내의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )보다 작다고 판단되면, 게인을 높이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.(240 내지 270)

[0083] 자동 게인 제어부(30)는  $M_{avg}$ 가  $M_{min}$ 보다 크면  $M_{max}$ 보다 작은지 확인한다. 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 신호가 처리될 수 있는 최적의 영역의 로우값( $M_{min}$ )과 신호가 처리될 수 있는 최적의 영역의 하이값( $M_{max}$ ) 사이에 위치하게 되면 게인 조정을 안하거나, 다이내믹 레인지의 음성 신호 대역의 중간지점( $L_{med}$ )으로 신호의 크기를 조정할 수 있도록 게인을 조정한다.(240, 280, 290)

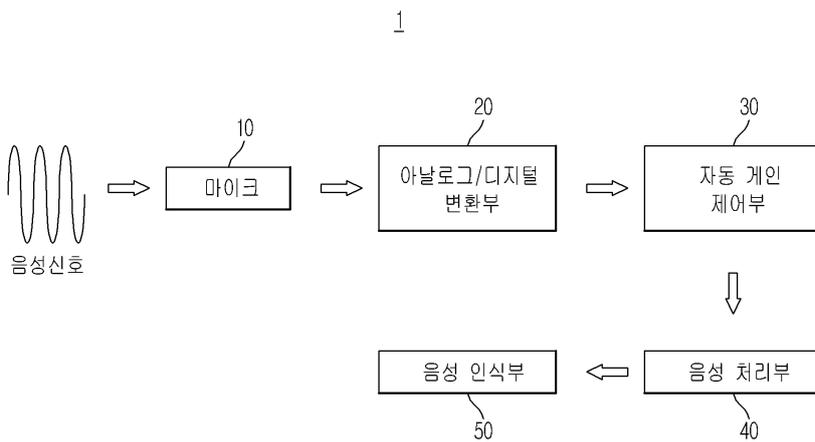
[0084] 자동 게인 제어부(30)는  $M_{avg}$ 가  $M_{max}$ 보다 크면  $L_{upper}$ 보다 작은지 확인한다. 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 다이내믹 레인지(Dynamic Range)의 최적의 영역의 하이값( $M_{max}$ )보다 크고 최고 한계 크기( $L_{upper}$ )보다 작다고 판단되면, 게인을 줄이고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호에도 조정된 게인을 적용하여 음성처리부(40)에서 처리할 수 있도록 한다.(300 내지 330)

[0085] 자동 게인 제어부(30)는 입력되는 음성 신호의 평균값( $M_{avg}$ )이 최고 한계 크기( $L_{upper}$ )보다 크다고 판단되면, 게인을 낮추고 이후에 입력되는 신호에 조정된 게인을 적용한다. 이 때, 최초 음성 신호의 평균값( $M_{avg}$ )을 계산하는데 사용되는 프레임에 속한 신호는 음성처리부(40)에서 사용되지 않는다.(300,340,350)

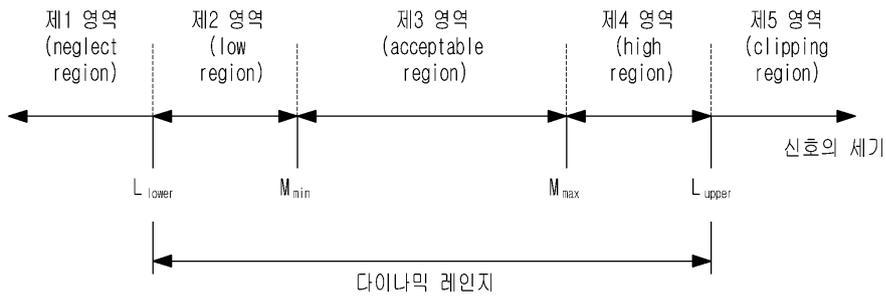
[0086] 한편, 상술한 과정에서 게인을 조정하는 것은 설계에 따라 도 2에 도시된 제 3 영역 즉, 신호가 처리될 수 있는 최적의 영역( $M_{min}$ 과  $M_{max}$ 의 사이 영역)에 음성신호의 세기가 존재하도록 조정하거나, 도 3에 도시된 음성 신호 대역의 중간지점( $L_{med}$ )에 음성신호의 세기가 존재하도록 조정하는 것을 의미한다.

도면

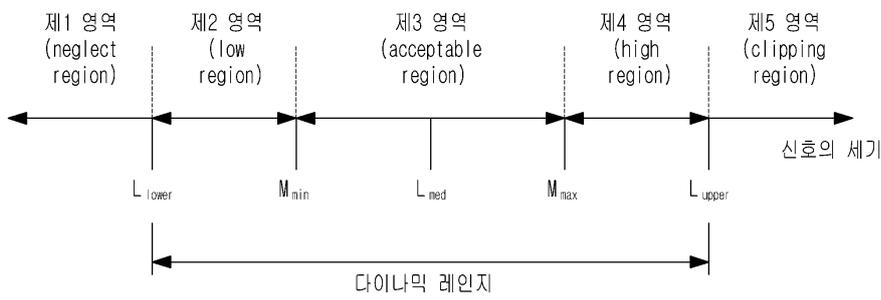
도면1



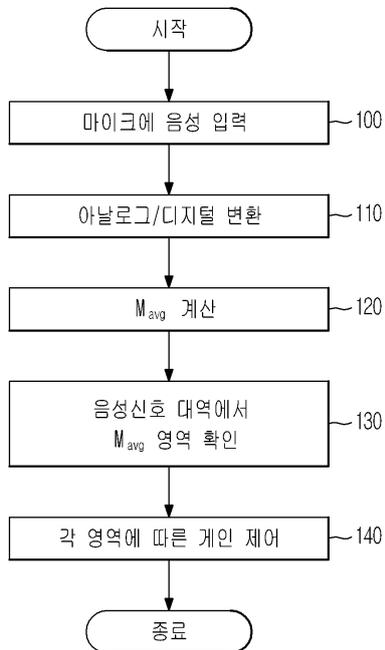
도면2



도면3



도면4



도면5

