



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0121168
(43) 공개일자 2014년10월15일

(51) 국제특허분류(Int. Cl.)

G10L 25/84 (2013.01)

(21) 출원번호 10-2013-0037586

(22) 출원일자 2013년04월05일

심사청구일자 없음

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

고한석

서울 용산구 이촌로88길 30, 103동 1602호 (이촌동, 삼성리버스위트아파트)

윤종성

서울 노원구 공릉로59길 25, 702동 313호 (하계동, 시영7단지청솔아파트)

(뒷면에 계속)

(74) 대리인

송인호, 민영준, 최관락

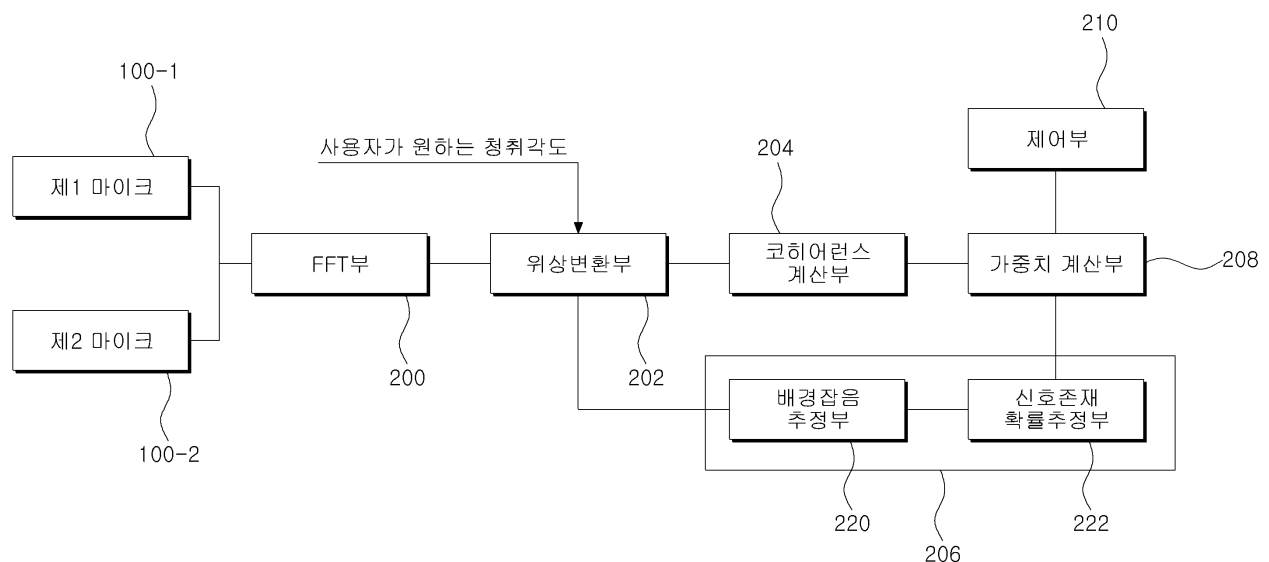
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **방향성 음향 신호 처리 장치**

(57) 요약

본 발명은 방향성 음향 처리 장치를 개시한다. 본 발명에 따르면, 주변 음향을 수신하는 복수의 마이크; 상기 복수의 마이크에서 입력된 음향 신호의 각 주파수 대역별 코히어런스를 계산하는 코히어런스 계산부; 상기 복수의 마이크 각각에서 입력된 음향 신호를 통해 음성 신호의 존재 확률을 추정하는 추정부; 상기 추정된 음성 신호 존재 확률에 따라 각 주파수 대역별로 서로 다른 가중치를 부여하며, 상기 부여된 가중치를 상기 코히어런스에 반영하여 결과값을 출력하는 가중치 계산부; 및 상기 결과값과 미리 설정된 임계치를 비교하여 상기 복수의 마이크로 입력된 음향 신호의 처리 여부를 결정하는 제어부를 포함하는 방향성 음향 신호 처리 장치가 제공된다.

대표도 - 도2



(72) 발명자

금민석

서울 광진구 군자로5길 18-3, 302호 (화양동)

박진수

서울 강남구 도곡로 404, 101동 803호 (대치동, 롯
데캐슬리베아파트)

문성규

경기 안양시 동안구 부림로 52, 502동 301호 (평촌
동, 초원럭키아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호 A11118912020000100

부처명 한국보건산업진흥원

연구관리전문기관 연구지원부

연구사업명 보건의료연구개발사업

연구과제명 보급형 모바일 장치를 이용한 청각장애자용 상황대처 음성 자막 변환기

기 여 율 1/1

주관기관 고려대학교 산학협력단

연구기간 2012.04.01 ~ 2013.03.31

특허청구의 범위

청구항 1

방향성 음향 신호 처리 장치로서,

주변 음향을 수신하는 복수의 마이크; 및

상기 복수의 마이크에서 입력된 음향 신호의 각 주파수 대역별 코히어런스를 계산하는 코히어런스 계산부;

상기 복수의 마이크 각각에서 입력된 음향 신호를 통해 음성 신호의 존재 확률을 추정하는 추정부;

상기 추정된 음성 신호 존재 확률에 따라 각 주파수 대역별로 서로 다른 가중치를 부여하며, 상기 부여된 가중치를 상기 코히어런스에 반영하여 결과값을 출력하는 가중치 계산부; 및

상기 결과값과 미리 설정된 임계치를 비교하여 상기 복수의 마이크로 입력된 음향 신호의 처리 여부를 결정하는 제어부를 포함하는 방향성 음향 신호 처리 장치.

청구항 2

제1항에 있어서,

사용자가 청취를 원하는 각도를 입력 받는 인터페이스부; 및

상기 인터페이스부를 통해 입력된 청취 각도에 따라 상기 복수의 마이크에서 입력된 음향 신호의 위상을 변환하는 위상 변환부를 더 포함하는 방향성 음향 신호 처리 장치.

청구항 3

제1항에 있어서,

상기 추정부는,

상기 복수의 마이크 각각에서 입력된 음향 신호를 통해 배경 잡음을 추정하는 배경 잡음 추정부; 및

상기 배경 잡음에 기초하여 상기 각 주파수 대역별 음성 신호 존재 확률을 추정하는 신호 존재 확률 추정부를 포함하는 방향성 음향 신호 처리 장치.

청구항 4

제1항에 있어서,

상기 가중치 계산부는,

상기 신호 존재 확률이 미리 설정된 값 이상인 경우에는 가중치를 1로 부여하고, 미리 설정된 값 미만인 경우에는 상기 신호 존재 확률을 가중치로 부여하는 방향성 음향 신호 처리 장치.

청구항 5

제1항에 있어서,

상기 결과값은 각 주파수 대역별 코히어런스에 가중치를 곱한 값을 모두 합산한 값인 방향성 음향 신호 처리 장치.

청구항 6

제1항에 있어서,

상기 제어부의 판단에 따라 상기 복수의 마이크로 입력된 신호를 스피커, 음성 인식부 및 전송부 중 적어도 하나로 출력되도록 선택하는 음향 출력 선택부를 더 포함하는 방향성 음향 신호 처리 장치.

명세서

기술 분야

[0001] 본 발명은 방향성 음향 신호 처리 장치에 관한 것으로서, 더욱 상세하게는 소음 환경에서 사용자가 원하는 방향의 소리만을 선택적으로 획득하여 처리할 수 있도록 하는 장치에 관한 것이다.

배경 기술

- [0002] 마이크는 음파 또는 초음파를 받아서 그 진동에 따른 전기신호를 발생하는 장치를 의미한다.
- [0003] 최근 로봇 관련 기술의 발전에 따라, 로봇과 사용자의 의사 소통을 위한 인터페이스로 마이크가 사용되고 있으며, 로봇은 마이크를 통해 입력되는 음향 신호를 전기신호로 변환하여 사용자의 음향을 인식하게 된다.
- [0004] 이러한 인식 과정에는 마이크의 주변 환경에서 발생하는 소음(잡음)이 항상 수반되며, 보다 다양한 명령을 정확하게 인식하여 사용자의 편의성을 향상시키기 위해서는 소음을 제외한 음향 신호만을 수신하는 것이 필수적이다.
- [0005] 또한, 이동 통신 단말기 및 네비게이션 장치 등 많은 수의 장치들이 사용자의 음성 인식 기능을 구비하고 있으며, 이러한 장치에서도 정확한 음성 인식에 기반한 동작을 위해서는 원하는 방향에서의 음성 신호만을 수신 처리하는 것이 필수적이다.
- [0006] 종래의 다수의 음향기기 업체에서 능동형 소음제거(Active Noise Cancelling) 기술을 적용하여 주변의 소음을 차단하는 장치를 제안하고 있다.
- [0007] 그러나 종래의 장치는 주변의 모든 소음을 제거하는 정숙한 환경을 조성하기 위한 목적으로 사용된다.
- [0008] 또한 멀티 채널 마이크를 사용하여 특정 방향에서 들어오는 소리만을 획득하는 종래의 기술이 헤드폰 및 보청기에 적용된 사례가 있으나, 사람의 시선이 향하는 방향의 소리만을 받아들이기 때문에 활용성에 제약이 있다.
- [0009] 보다 상세하게, 종래의 잡음 제거 방식으로는 GSC(Generalized Sidelobe Canceller)를 이용한 방법과 두 개의 마이크로 입력되는 신호의 위상 차를 이용하는 방법이 있다.
- [0010] 전자의 방법은 GSC 내부의 적응 필터 수렴 문제로 인해 출력이 발산하거나 수렴 속도가 느려지는 문제점을 갖는다.
- [0011] 그리고, 후자의 방법은 적응 필터를 사용하지 않아 수렴 문제는 갖지 않으나, 두 개의 마이크간의 위상 정보만을 사용하는 한계로 인하여 위상의 작은 변화에도 시스템이 민감하게 반응하여 일시적인 오류가 자주 발생하는 문제점을 갖는다.
- [0012] 대한민국 공개특허 제2008-0000478호(발명명칭: 휴대 단말기에서 복수의 마이크들로 입력된 신호들의 잡음을 제거하는 방법 및 장치)는 휴대 단말기로 입력되는 위상 차가 있는 신호로부터 잡음을 제거하는 방법 및 장치를 개시하고 있다.
- [0013] 그러나, 앞서 설명한 바와 같이, 상기의 종래특허는 방향성 신호를 제거하기 위해 위상 차만을 이용함으로써 시스템의 안정성이 낮아 음성 신호에서 잡음을 효과적으로 제거할 수 없게 된다.

발명의 내용

해결하려는 과제

[0014] 상기한 종래기술의 문제점을 해결하기 위해, 본 발명은 사용자가 원하는 특정 방향의 음향만을 획득할 수 있는 방향성 음향 신호 처리 장치를 제안하고자 한다.

과제의 해결 수단

[0015] 상기한 기술적 과제를 해결하기 위해, 본 발명의 바람직한 일 실시예에 따르면, 방향성 음향 신호 처리 장치로서, 주변 음향을 수신하는 복수의 마이크; 상기 복수의 마이크에서 입력된 음향 신호의 각 주파수 대역별 코히어런스를 계산하는 코히어런스 계산부; 상기 복수의 마이크 각각에서 입력된 음향 신호를 통해 음성 신호의 존재 확률을 추정하는 추정부; 상기 추정된 음성 신호 존재 확률에 따라 각 주파수 대역별로 서로 다른 가중치를 부여하며, 상기 부여된 가중치를 상기 코히어런스에 반영하여 결과값을 출력하는 가중치 계산부; 및 상기 결과값과 미리 설정된 임계치를 비교하여 상기 복수의 마이크로 입력된 음향 신호의 처리 여부를 결정하는 제어부를

포함하는 방향성 음향 신호 처리 장치가 제공된다.

- [0016] 사용자가 청취를 원하는 각도를 입력 받는 인터페이스부; 및 상기 인터페이스부를 통해 입력된 청취 각도에 따라 상기 복수의 마이크에서 입력된 음향 신호의 위상을 변환하는 위상 변환부를 더 포함할 수 있다.
- [0017] 상기 추정부는, 상기 복수의 마이크 각각에서 입력된 음향 신호를 통해 배경 잡음을 추정하는 배경 잡음 추정부; 및 상기 배경 잡음에 기초하여 상기 각 주파수 대역별 음성 신호 존재 확률을 추정하는 신호 존재 확률 추정부를 포함할 수 있다.
- [0018] 상기 가중치 계산부는, 상기 신호 존재 확률이 미리 설정된 값 이상인 경우에는 가중치를 1로 부여하고, 미리 설정된 값 미만인 경우에는 상기 신호 존재 확률을 가중치로 부여할 수 있다.
- [0019] 상기 결과값은 각 주파수 대역별 코히어런스에 가중치를 곱한 값을 모두 합산한 값일 수 있다.
- [0020] 상기 제어부의 판단에 따라 상기 복수의 마이크로 입력된 신호를 스피커, 음성 인식부 및 전송부 중 적어도 하나로 출력되도록 선택하는 음향 출력 선택부를 더 포함할 수 있다.

발명의 효과

- [0021] 본 발명에 따르면, 배경 잡음을 추정하고, 이를 통해 신호 존재 확률을 추정하여 음향 신호의 처리 여부를 결정하기 때문에 소음을 제외한 실제 사용자의 음성만을 선택적으로 처리할 수 있는 장점이 있다.

도면의 간단한 설명

- [0022] 도 1은 본 발명의 바람직한 일 실시예에 따른 방향성 음향 처리 장치의 구성을 도시한 도면.
- 도 2는 본 발명에 따른 공간 음향 활성 감지부의 상세 구성을 도시한 도면.
- 도 3은 본 발명에 따른 음향 출력 선택부의 온/오프 과정을 도시한 도면.
- 도 4는 본 발명에 따른 음향 감쇠 및 처리 상태를 도시한 도면.

발명을 실시하기 위한 구체적인 내용

- [0023] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.
- [0024] 이하, 본 발명의 바람직한 실시예를 첨부한 도면들을 참조하여 상세히 설명하기로 한다. 본 발명을 설명함에 있어 전체적인 이해를 용이하게 하기 위하여 도면 번호에 상관없이 동일한 수단에 대해서는 동일한 참조 번호를 사용하기로 한다.
- [0025] 도 1은 본 발명의 바람직한 일 실시예에 따른 방향성 음향 처리 장치의 블록도이다.
- [0026] 도 1에 도시된 바와 같이, 본 발명에 따른 방향성 음향 처리 장치는 복수의 마이크(100-1 내지 100-n), 공간 음향 활성 감지부(Spatial Voice Activity Detection, 102), 인터페이스부(104) 및 음향 출력 선택부(106)를 포함할 수 있다.
- [0027] 도 1에 도시된 바와 같이, 본 발명에 따른 방향성 음향 처리 장치에는 정면 소리 및 측면 소리와 같이 서로 다른 방향으로부터의 음향이 복수의 마이크(100)를 통해 입력된다.
- [0028] 공간 음향 활성 감지부(102)는 미리 설정된 알고리즘에 따라 마이크(100)를 통해 입력된 음향을 스피커로 출력할지 또는 상대방 단말로 전송할지 여부를 결정한다.
- [0029] 바람직하게, 본 발명에 따른 공간 음향 활성 감지부(102)는 배경 잡음을 추정하고, 추정된 배경 잡음에 따라 음성 신호 존재 확률을 추정하며, 추정된 신호 존재 확률에 기초하여 최종적으로 마이크를 통해 입력된 신호를 처리할지 여부를 결정한다.
- [0030] 이때, 공간 음향 활성 감지부(102)는 미리 설정된 시간 간격(예를 들어, 20ms 또는 30ms 간격)으로 입력된 음향

의 신호 처리 여부를 결정한다.

[0031] 여기서, 마이크를 통해 입력된 신호의 처리는 입력된 음향 신호를 스피커를 통해 출력할지 여부, 장치가 음성 인식 장치일 경우, 음성 인식 처리부(미도시)로 출력할지 여부, 장치가 통화 장치인 경우, 상대방 장치에 전달할지 여부를 결정하는 것을 의미한다.

[0032] 바람직하게, 음향 출력 선택부(106)는 스피커, 음성 인식 처리부 또는 전송부와 연결되며, 공간 음향 활성 감지부(102)의 판단에 따라 음향의 출력 여부가 결정된다.

[0033] 도 2는 본 발명에 따른 공간 음향 활성 감지부의 상세한 구성을 도시한 도면이다.

[0034] 도 2에 도시된 바와 같이, 본 발명에 따른 공간 음향 활성 감지부(102)는 FFT부(200), 위상 변환부(202), 코히어런스 계산부(204), 추정부(206), 가중치 계산부(208) 및 제어부(210)를 포함할 수 있다.

[0035] 도 2에서는 설명의 편의를 위해 제1 마이크(100-1) 내지 제2 마이크(100-2)가 구비되는 것으로 설명하나, 반드시 이에 한정되지 않는다는 점을 당업자는 이해하여야 할 것이다.

[0036] 제1 마이크(100-1)를 통해 입력되는 음향 신호를 $X(t, w)$, 제2 마이크(100-2)를 통해 입력되는 음향 신호를 $Y(t, w)$ 라 가정한다.

[0037] 여기서, t 는 시간(프레임)이며, w 는 주파수 대역 인덱스를 의미한다.

[0038] FFT부(200)는 제1 및 제2 마이크(100-1 내지 100-2)를 통해 입력된 음향 신호를 주파수 도메인으로 변환한다. 바람직하게, FFT부(200)는 Short Time Fast Fourier Transform(STFFT)를 수행하여 마이크를 통해 입력된 음향 신호를 주파수 도메인으로 변환한다.

[0039] 위상 변환부(202)는 인터페이스부(104)를 통해 사용자가 미리 설정한 청취 각도에 따라 제1 마이크 및 제2 마이크의 주파수별 위상(phase) 값을 변환한다.

[0040] 여기서, 청취 각도는 사용자가 청취를 원하는 각도로 정의되며, 예를 들어, 마이크(100-1, 100-2)를 기준으로 정면에서 입력되는 음향의 각도를 0° 라 가정할 때, 일반적으로 청취 각도는 0° 로 설정될 수 있다. 그러나 경우에 따라 사용자가 측면에서 입력되는 음향을 청취하기를 원하는 경우 청취 각도는 예를 들어, 90° 로 설정될 수 있으며, 위상 변환부(202)는 90° 각도에서 입력되는 음향의 위상을 변환하여 0° , 즉 정면에서 입력되는 것과 마찬가지로 위상 값을 변환할 수 있다.

[0041] 위상 변환 후 코히어런스 계산부(204)는 아래의 수학식 1을 통해 주파수 대역별 코히어런스를 계산한다.

수학식 1

$$\Gamma(t, w_k) = \frac{G_{xy}(t, w_k)}{G_{xx}(t, w_k)G_{yy}(t, w_k)}$$

[0042]

[0043] 여기서, $\Gamma(t, w_k)$ 는 주파수 대역별 코히어런스, $G_{xy}(t, w_k)$ 는 제1 마이크(100-1) 및 제2 마이크(100-2) 입력 X, Y 의 상호상관(cross correlation) 값이며, $G_{xx}(t, w_k)$ 와 $G_{yy}(t, w_k)$ 는 자기상관(auto correlation) 값이다.

[0044] 본 발명의 바람직한 일 실시예에 따르면, 추정부(206)는 복수의 마이크 각각에서 입력된 신호를 통해 음성 신호 존재 확률을 추정한다.

[0045] 보다 상세하게 추정부(206)는 배경 잡음 추정부(220) 및 신호 존재 확률 추정부(222)를 포함할 수 있다.

[0046] 배경 잡음 추정부(220)에서는 MCRA(Minima Controlled Recursive Averaging) 방법 등을 이용하여 마이크 입력

신호로부터 배경 잡음 성분 $N(t, w_k)$ 을 추정한다.

[0047] 신호 존재 확률 추정부(222)는 추정된 배경 잡음 성분에 기초하여 각 주파수 대역의 음성 존재 확률인 $p(t, w)$ 를 아래의 수학적 2를 통해 계산한다.

수학적 2

$$p(t, \omega) = \left\{ 1 + \frac{q(t, \omega)}{1 - q(t, \omega)} (1 + \xi(t, \omega) * \exp(-\nu(t, \omega))) \right\}^{-1}$$

[0048]

$\lambda_x(t, \omega)$: 마이크 입력신호의 power $\lambda_y(t, \omega)$: 추정 잡음신호의 power

[0049] 여기서,

[0050] $q(t, w)$: Prior Speech Absent Probability,

$$\xi(t, \omega) = \frac{\lambda_x(t, \omega)}{\lambda_y(t, \omega)}, \gamma(t, \omega) = \frac{|X(t, \omega)|^2}{\lambda_y(t, \omega)}, \nu(t, \omega) = \frac{\gamma(t, \omega)\xi(t, \omega)}{1 + \xi(t, \omega)}$$

[0051] $q(t, \omega)$: Prior Speech Absent Probability, 이다.

[0052] 상기한 수학적 2를 통해 계산된 $p(t, w)$ 를 바탕으로 본 발명에 따른 가중치 계산부(208)는 가중치를 계산한다.

[0053] 본 발명에 일 실시예에 따르면, $p(t, w)$ 가 0.8 이상이면 가중치 $W(t, w_k)$ 에 1을 부여하고, 나머지 경우에는 $p(t, w)$ 값을 $W(t, w_k)$ 으로 부여한다.

[0054] 여기서, 0.8은 실험적으로 결정된 수치이며, 이에 한정됨이 없이 가중치 1에 부여되는 수치는 다양하게 적용될 수 있다.

[0055] 가중치 계산부(208)는 아래의 수학적 3과 같이 소정 주파수 대역마다 앞서 계산된 코히어런스에 상기와 같이 부여된 가중치를 곱한다.

수학적 3

$$W(t, w_k) * \Gamma(t, w_k) = \Gamma'(t, w_k)$$

[0056]

[0057] 또한, 가중치 계산부(208)는 각 주파수 대역의 코히어런스에 가중치가 곱해진 결과를 아래의 수학적 4와 같이

더한 후 이를 제어부(210)로 출력한다.

수학식 4

$$\sum_{k=1}^M \Gamma'(t, w_k)$$

[0058]

[0059]

본 발명에 따른 제어부(210)는 수학식 5와 같이 가중치 계산부(208)에서 출력된 결과값이 미리 설정된 임계치를 넘는지 여부를 판단하며, 임계치를 넘으면 사용자가 청취를 원하는 각도에서 음성이 입력된 것으로 판단한다. 즉, 제어부(210)는 결과값이 임계치 이상인 경우 현재 입력된 음향 신호를 처리하는 것으로 결정한다.

수학식 5

$$\sum_{k=1}^M \Gamma'(t, w_k) \geq T_{SVAD}$$

[0060]

[0061]

반면, 제어부(210)는 가중치 계산부(208)에서의 결과값이 임계치를 넘지 못할 경우 현재 입력된 음향 신호를 감쇠시키는 것으로 결정한다.

[0062]

본 발명에 따르면, 제어부(210)는 상기한 판단 결과에 따른 제어 신호를 음향 출력 선택부(106)로 출력한다.

[0063]

도 3에 도시된 바와 같이, 제어부(210)의 온/오프 제어 신호에 따라 음향 출력 선택부(106)는 마이크를 통해 입력된 음향 신호의 출력 여부를 선택한다.

[0064]

바람직하게 음향 출력 선택부(106)는 스위칭 소자로 구현될 수 있다.

[0065]

보다 상세하게, 도 3b와 같이 제어부(210)는 가중치 계산부(208)의 결과값이 임계치를 초과하는 경우에만 마이크로로부터 입력된 음향 신호가 스피커, 음성 인식부 또는 전송부로 입력되도록 제어한다.

[0066]

도 4에 도시된 바와 같이, 사용자가 원하지 않는 방향에서 음향이 입력될 때, 이러한 신호는 감쇠시키고, 사용자가 원하는 방향에서 입력된 음향만을 출력하기 때문에 소음 환경에서 효율적으로 사용될 수 있다.

[0067]

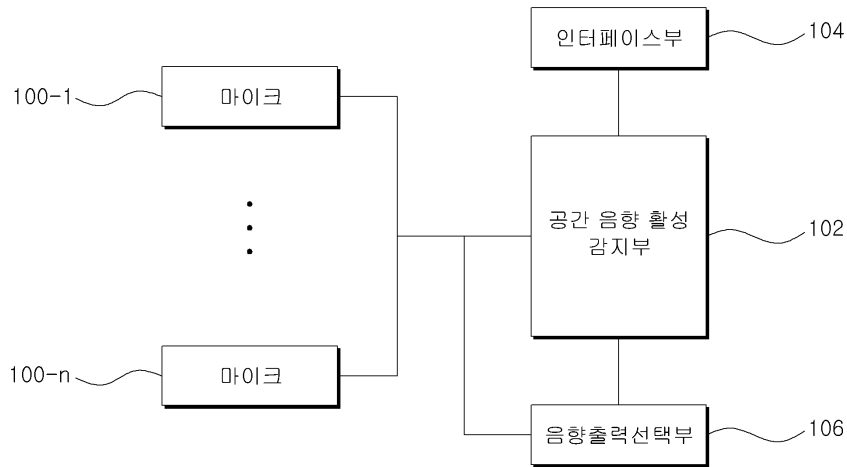
본 발명의 실시예들은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 본 발명의 일 실시예들의 동작을 수행하기 위해 적어도 하나의 소프트웨어 모듈로서 동작하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0068]

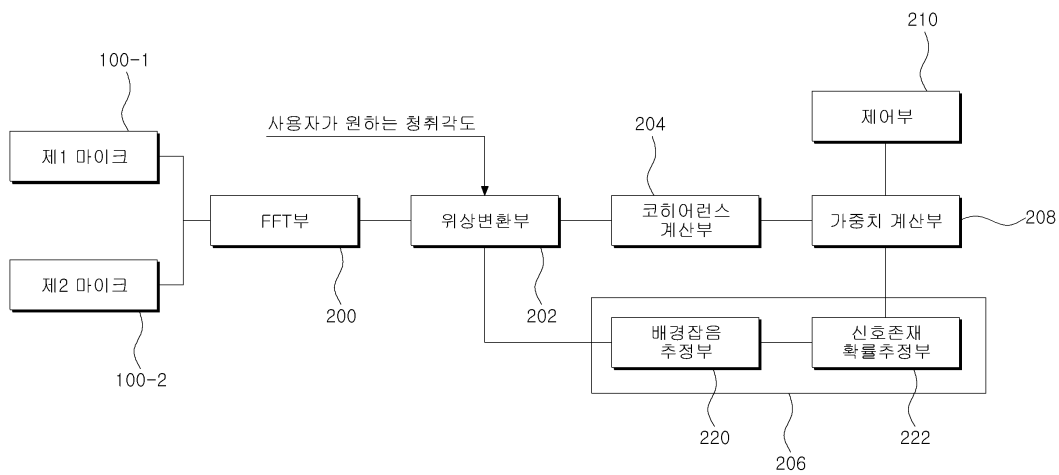
상기한 본 발명의 바람직한 실시예는 예시의 목적을 위해 개시된 것이고, 본 발명에 대해 통상의 지식을 가진 당업자라면 본 발명의 사상과 범위 안에서 다양한 수정, 변경, 부가가 가능할 것이며, 이러한 수정, 변경 및 부가는 하기의 특허청구범위에 속하는 것으로 보아야 할 것이다.

도면

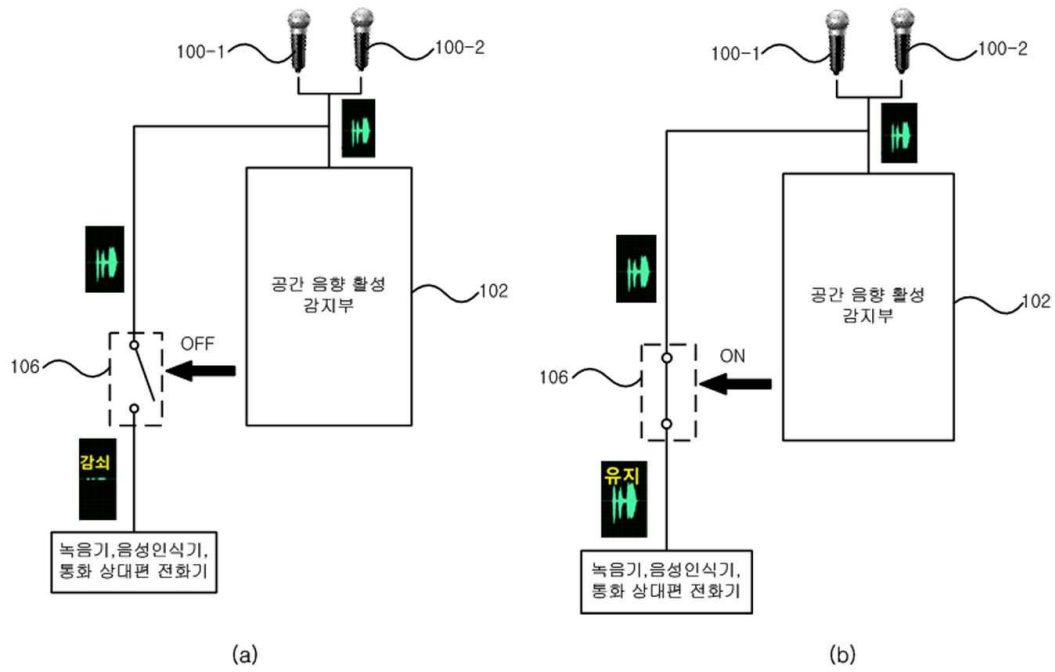
도면1



도면2



도면3



도면4

