



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0086046
(43) 공개일자 2018년07월30일

(51) 국제특허분류(Int. Cl.)
G08B 13/196 (2006.01) G10L 17/26 (2013.01)
H04N 5/232 (2006.01) H04N 7/18 (2006.01)

(52) CPC특허분류
G08B 13/1963 (2013.01)
G08B 13/19654 (2013.01)

(21) 출원번호 10-2017-0009930
(22) 출원일자 2017년01월20일
심사청구일자 없음

(71) 출원인
한화에어로스페이스 주식회사
경상남도 창원시 성산구 창원대로 1204 (성주동)
광주과학기술원
광주광역시 북구 첨단과기로 123 (오룡동)

(72) 발명자
박진수
경상남도 창원시 성산구 창원대로 1204 (성주동)
임정은
경상남도 창원시 성산구 창원대로 1204 (성주동)
(뒷면에 계속)

(74) 대리인
리엔목특허법인

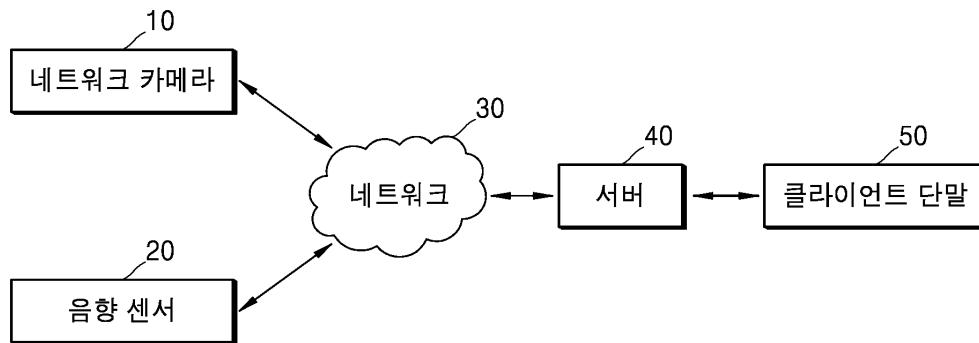
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **감시 시스템 및 그 동작 방법**

(57) 요약

본 발명이 해결하고자 하는 과제를 해결하기 위한 일 실시예에 따른 감시 시스템은 적어도 두 개의 음향 센서들로부터 음향 신호를 수신하는 통신 인터페이스; 및 상기 음향 신호를 버퍼링하고, 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우에 상기 음향 신호의 특징을 추출하고, 음향 모델에 기초하여 상기 음향 신호의 특징으로부터 상기 음향 신호를 판별하고, 판별된 음향 신호가 이벤트에 해당하면 상기 음향 신호로부터 음원의 위치를 검출하고, 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하는 프로세서;를 포함한다.

대표도 - 도1



(52) CPC특허분류

G10L 17/26 (2013.01)

H04N 5/23296 (2013.01)

H04N 7/18 (2013.01)

(72) 발명자

김홍국

광주광역시 광산구 첨단중앙로181번길 42-4, 101동
1101호(월계동, 첨단과학산업단지 금광아파트)

전광명

광주광역시 북구 첨단과기로 123, 생활관 2209호
(오룡동, 광주과학기술원)

김남균

광주광역시 남구 백운로 78 (백운동)

명세서

청구범위

청구항 1

적어도 두 개의 음향 센서들로부터 음향 신호를 수신하는 통신 인터페이스; 및

상기 음향 신호를 버퍼링하고, 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우에 상기 음향 신호의 특징을 추출하고, 음향 모델에 기초하여 상기 음향 신호의 특징으로부터 상기 음향 신호를 판별하고, 판별된 음향 신호가 이벤트에 해당하면 상기 음향 신호로부터 음원의 위치를 검출하고, 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하는 프로세서;를 포함하는, 감시 시스템.

청구항 2

청구항 1에 있어서,

상기 음향 신호의 특징은 상기 음향 신호의 MFCC(Mel-frequency cepstral coefficients)를 포함하고,

상기 음향 모델은 GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 하는, 감시 시스템.

청구항 3

청구항 1에 있어서,

카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고,

상기 프로세서는 상기 사용자 입력에 대응하여 상기 음향 신호의 특징을 추출하기 위한 상기 기준 값을 결정하는, 감시 시스템.

청구항 4

청구항 1에 있어서,

카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고,

상기 프로세서는 상기 사용자 입력에 대응하여 상기 음향 신호를 버퍼링함에 있어서, 버퍼링되는 상기 음향 신호의 프레임 개수를 결정하는, 감시 시스템.

청구항 5

청구항 1에 있어서,

상기 이벤트를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고,

상기 프로세서는 상기 사용자 입력에 대응하여 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단하는, 감시 시스템.

청구항 6

청구항 1에 있어서,

상기 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 상기 프로세서는 상기 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 상기 복수의 음원의 위치들에 대응하는 상기 복수의 각도들의 상기 평균 각도를 산출하는, 감시 시스템.

청구항 7

청구항 6에 있어서,

상기 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신하는 사용자 인터페이스;를 더

포함하고,

상기 프로세서는 상기 사용자 입력에 대응하여 상기 평균 각도를 산출하는, 감시 시스템.

청구항 8

청구항 1에 있어서,

상기 프로세서는 상기 음향 신호를 STFT(Short-Time Fourier Transform)하고, STFT된 음향 신호를 PHAT(phase transform)하고, PHAT된 음향 신호의 SRP(steered response power)를 산출하고, SRP가 피크인 위치를 상기 음원의 위치로 검출하는, 감시 시스템.

청구항 9

청구항 8에 있어서,

소정 SRP를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고,

상기 프로세서는 상기 사용자 입력에 대응하여, 상기 소정 SRP 이상의 상기 SRP가 피크인 위치에 한하여 상기 음원의 위치로 산출하는, 감시 시스템.

청구항 10

청구항 1에 있어서,

상기 이벤트는 비명, 총소리, 폭발음, 충돌음, 및 소정 주파수 대역의 음향 신호 중 적어도 하나를 포함하는, 감시 시스템.

청구항 11

통신 인터페이스 및 프로세서를 포함하는 감시 시스템의 동작 방법에 있어서,

상기 통신 인터페이스에 의해, 적어도 두 개의 음향 센서들로부터 음향 신호를 수신하는 단계;

상기 프로세서에 의해, 상기 음향 신호를 버퍼링하는 단계;

상기 프로세서에 의해, 버퍼링된 음향 신호의 에너지를 기준 값과 비교하는 단계;

상기 프로세서에 의해, 상기 버퍼링된 음향 신호의 에너지가 상기 기준 값 이상인 경우에 상기 음향 신호의 특징을 추출하는 단계;

상기 프로세서에 의해, 음향 모델에 기초하여 상기 음향 신호의 특징으로부터 상기 음향 신호를 판별하는 단계;

상기 프로세서에 의해, 판별된 음향 신호가 이벤트에 해당하는지 여부를 판단하는 단계;

상기 프로세서에 의해, 상기 판별된 음향 신호가 상기 이벤트에 해당하면 상기 음향 신호로부터 음원의 위치를 검출하는 단계;

상기 프로세서에 의해, 검출된 음원의 위치의 개수를 판단하는 단계; 및

상기 프로세서에 의해, 상기 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하는 단계;를 포함하는, 감시 시스템의 동작 방법.

청구항 12

청구항 11에 있어서,

상기 음향 신호의 특징을 추출하는 단계는,

상기 음향 신호의 MFCC(Mel-frequency cepstral coefficients)를 추출하는 단계이고,

상기 음향 신호를 판별하는 단계 이전에,

상기 프로세서에 의해, GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 하는 상기 음향 모델을 생성하는 단계;를 더 포함하는, 감시 시스템의 동작 방법.

청구항 13

청구항 11에 있어서,
 상기 음향 신호의 특징을 추출하는 단계 이전에,
 사용자 인터페이스에 의해, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 단계; 및
 상기 프로세서에 의해, 상기 사용자 입력에 대응하여 상기 음향 신호의 특징을 추출하기 위한 상기 기준 값을 결정하는 단계;를 더 포함하는, 감시 시스템의 동작 방법.

청구항 14

청구항 11에 있어서,
 사용자 인터페이스에 의해, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고,
 상기 음향 신호를 버퍼링하는 단계는,
 상기 사용자 입력에 대응하여 버퍼링되는 상기 음향 신호의 프레임 개수를 결정하는 단계; 및
 결정된 프레임 개수만큼 상기 음향 신호를 버퍼링하는 단계;를 포함하는, 감시 시스템의 동작 방법.

청구항 15

청구항 11에 있어서,
 사용자 인터페이스에 의해, 상기 이벤트를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고,
 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단하는 단계는,
 상기 사용자 입력에 대응하여 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단하는 단계인, 감시 시스템의 동작 방법.

청구항 16

청구항 11에 있어서,
 상기 평균 각도를 산출하는 단계는,
 상기 복수의 음원의 위치들에서 발생한 각 이벤트의 동일성을 판단하는 단계; 및
 상기 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 상기 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 상기 복수의 음원의 위치들에 대응하는 상기 복수의 각도들의 상기 평균 각도를 산출하는 단계;를 포함하는, 감시 시스템의 동작 방법.

청구항 17

청구항 16에 있어서,
 사용자 인터페이스에 의해, 상기 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신하는 단계;를 더 포함하고,
 상기 평균 각도를 산출하는 단계는,
 상기 사용자 입력에 대응하여 상기 평균 각도를 산출하는 단계인, 감시 시스템의 동작 방법.

청구항 18

청구항 11에 있어서,
 상기 음원의 위치를 검출하는 단계는,
 상기 음향 신호를 STFT(Short-Time Fourier Transform)하는 단계;
 STFT된 음향 신호를 PHAT(phase transform)하는 단계;

PHAT된 음향 신호의 SRP(steered response power)를 산출하는 단계; 및
SRP가 피크인 위치를 상기 음원의 위치로 검출하는 단계;를 포함하는, 감시 시스템의 동작 방법.

청구항 19

청구항 18에 있어서,
사용자 인터페이스에 의해, 소정 SRP를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고,
상기 SRP가 피크인 위치를 상기 음원의 위치로 검출하는 단계는,
상기 사용자 입력에 대응하여, 상기 소정 SRP 이상의 상기 SRP가 피크인 위치에 한하여 상기 음원의 위치로 산출하는 단계인, 감시 시스템의 동작 방법.

청구항 20

청구항 11에 있어서,
상기 프로세서에 의해, 카메라의 중앙 각도가 상기 평균 각도와 일치하도록 카메라 패닝을 제어하는 단계;를 더 포함하는, 감시 시스템의 동작 방법.

발명의 설명

기술 분야

[0001] 본 발명은 음향 센서에 의해 감지되는 음향 신호를 이용하는 감시 시스템 및 그 동작 방법에 관한 것이다.

배경 기술

[0002] 최근 들어, 방법, 보안 등 다양한 목적으로 건물의 내부, 외부 또는 길거리 등에 카메라들을 설치하는 경우가 증가하고 있다. 카메라들은 네트워크를 통하여 서로 연결되어 네트워크 카메라로서의 기능을 수행할 수 있다.

[0003] 또한 카메라가 설치된 장소를 관리하는 관리자는 개인용 컴퓨터 등을 통하여 카메라에 접속하여 원격으로 건물이나 매장 등의 원격지를 관리할 수 있다.

[0004] 한편, 카메라에 의해 촬영되는 영상 신호를 이용하는 감시 시스템은 영상 신호 상에서 표적 객체가 비표적 객체에 의해 가려진 경우 또는 표적 객체가 카메라의 화각을 벗어난 경우에 표적 객체를 제대로 감시할 수 없게 된다.

[0005] 카메라에 의해 촬영되는 영상 신호와 음향 센서에 의해 감지되는 음향 신호를 함께 이용하는 감시 시스템은 많은 계산량으로 인해 감시 동작을 수행하기까지 오랜 시간이 걸리거나, 많은 리소스를 필요로 하기 때문에 비용이 많이 드는 단점이 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 국내 등록특허공보 제1243897호

발명의 내용

해결하려는 과제

[0007] 본 발명이 해결하고자 하는 과제는 이벤트가 발생한 위치를 향하여 카메라 패닝을 수행하는 감시 시스템 및 그 동작 방법을 제공하는데 있다.

과제의 해결 수단

[0008] 본 발명이 해결하고자 하는 과제를 해결하기 위한 일 실시예에 따른 감시 시스템은 적어도 두 개의 음향 센서들

로부터 음향 신호를 수신하는 통신 인터페이스; 및 상기 음향 신호를 버퍼링하고, 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우에 상기 음향 신호의 특징을 추출하고, 음향 모델에 기초하여 상기 음향 신호의 특징으로부터 상기 음향 신호를 판별하고, 판별된 음향 신호가 이벤트에 해당하면 상기 음향 신호로부터 음원의 위치를 검출하고, 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하는 프로세서;를 포함한다.

- [0009] 본 실시예에서, 상기 음향 신호의 특징은 상기 음향 신호의 MFCC(Mel-frequency cepstral coefficients)를 포함하고, 상기 음향 모델은 GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 할 수 있다.
- [0010] 본 실시예에서, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고, 상기 프로세서는 상기 사용자 입력에 대응하여 상기 음향 신호의 특징을 추출하기 위한 상기 기준 값을 결정할 수 있다.
- [0011] 본 실시예에서, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고, 상기 프로세서는 상기 사용자 입력에 대응하여 상기 음향 신호를 버퍼링함에 있어서, 버퍼링되는 상기 음향 신호의 프레임 개수를 결정할 수 있다.
- [0012] 본 실시예에서, 상기 이벤트를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고, 상기 프로세서는 상기 사용자 입력에 대응하여 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단할 수 있다.
- [0013] 본 실시예에서, 상기 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 상기 프로세서는 상기 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 상기 복수의 음원의 위치들에 대응하는 상기 복수의 각도들의 상기 평균 각도를 산출할 수 있다.
- [0014] 본 실시예에서, 상기 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고, 상기 프로세서는 상기 사용자 입력에 대응하여 상기 평균 각도를 산출할 수 있다.
- [0015] 본 실시예에서, 상기 프로세서는 상기 음향 신호를 STFT(Short-Time Fourier Transform)하고, STFT된 음향 신호를 PHAT(phase transform)하고, PHAT된 음향 신호의 SRP(steered response power)를 산출하고, SRP가 피크인 위치를 상기 음원의 위치로 검출할 수 있다.
- [0016] 본 실시예에서, 소정 SRP를 선택하는 사용자 입력을 수신하는 사용자 인터페이스;를 더 포함하고, 상기 프로세서는 상기 사용자 입력에 대응하여, 상기 소정 SRP 이상의 상기 SRP가 피크인 위치에 한하여 상기 음원의 위치로 산출할 수 있다.
- [0017] 본 실시예에서, 상기 이벤트는 비명, 총소리, 폭발음, 충돌음, 및 소정 주파수 대역의 음향 신호 중 적어도 하나를 포함할 수 있다.
- [0018] 본 발명이 해결하고자 하는 과제를 해결하기 위한 일 실시예에 따른 감시 시스템의 동작 방법은 통신 인터페이스 및 프로세서를 포함하는 감시 시스템의 동작 방법에 있어서, 상기 통신 인터페이스에 의해, 적어도 두 개의 음향 센서들로부터 음향 신호를 수신하는 단계; 상기 프로세서에 의해, 상기 음향 신호를 버퍼링하는 단계; 상기 프로세서에 의해, 버퍼링된 음향 신호의 에너지를 기준 값과 비교하는 단계; 상기 프로세서에 의해, 상기 버퍼링된 음향 신호의 에너지가 상기 기준 값 이상인 경우에 상기 음향 신호의 특징을 추출하는 단계; 상기 프로세서에 의해, 음향 모델에 기초하여 상기 음향 신호의 특징으로부터 상기 음향 신호를 판별하는 단계; 상기 프로세서에 의해, 판별된 음향 신호가 이벤트에 해당하는지 여부를 판단하는 단계; 상기 프로세서에 의해, 상기 판별된 음향 신호가 상기 이벤트에 해당하면 상기 음향 신호로부터 음원의 위치를 검출하는 단계; 상기 프로세서에 의해, 검출된 음원의 위치의 개수를 판단하는 단계; 및 상기 프로세서에 의해, 상기 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하는 단계;를 포함한다.
- [0019] 본 실시예에서, 상기 음향 신호의 특징을 추출하는 단계는, 상기 음향 신호의 MFCC(Mel-frequency cepstral coefficients)를 추출하는 단계이고, 상기 음향 신호를 판별하는 단계 이전에, 상기 프로세서에 의해, GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 하는 상기 음향 모델을 생성하는 단계;를 더 포함할 수 있다.
- [0020] 본 실시예에서, 상기 음향 신호의 특징을 추출하는 단계 이전에, 사용자 인터페이스에 의해, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 단계; 및 상기 프로세서에 의해, 상기 사용자 입력에 대응하여 상기 음

향 신호의 특징을 추출하기 위한 상기 기준 값을 결정하는 단계;를 더 포함할 수 있다.

- [0021] 본 실시예에서, 사용자 인터페이스에 의해, 카메라 패닝 민감도를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고, 상기 음향 신호를 버퍼링하는 단계는, 상기 사용자 입력에 대응하여 버퍼링되는 상기 음향 신호의 프레임 개수를 결정하는 단계; 및 결정된 프레임 개수만큼 상기 음향 신호를 버퍼링하는 단계;를 포함할 수 있다.
- [0022] 본 실시예에서, 사용자 인터페이스에 의해, 상기 이벤트를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고, 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단하는 단계는, 상기 사용자 입력에 대응하여 상기 판별된 음향 신호가 상기 이벤트에 해당하는지 여부를 판단하는 단계일 수 있다.
- [0023] 본 실시예에서, 상기 평균 각도를 산출하는 단계는, 상기 복수의 음원의 위치들에서 발생한 각 이벤트의 동일성을 판단하는 단계; 및 상기 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 상기 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 상기 복수의 음원의 위치들에 대응하는 상기 복수의 각도들의 상기 평균 각도를 산출하는 단계;를 포함할 수 있다.
- [0024] 본 실시예에서, 사용자 인터페이스에 의해, 상기 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신하는 단계;를 더 포함하고, 상기 평균 각도를 산출하는 단계는, 상기 사용자 입력에 대응하여 상기 평균 각도를 산출하는 단계일 수 있다.
- [0025] 본 실시예에서, 상기 음원의 위치를 검출하는 단계는, 상기 음향 신호를 STFT(Short-Time Fourier Transform) 하는 단계; STFT된 음향 신호를 PHAT(phase transform)하는 단계; PHAT된 음향 신호의 SRP(steered response power)를 산출하는 단계; 및 SRP가 피크인 위치를 상기 음원의 위치로 검출하는 단계;를 포함할 수 있다.
- [0026] 본 실시예에서, 사용자 인터페이스에 의해, 소정 SRP를 선택하는 사용자 입력을 수신하는 단계;를 더 포함하고, 상기 SRP가 피크인 위치를 상기 음원의 위치로 검출하는 단계는, 상기 사용자 입력에 대응하여, 상기 소정 SRP 이상의 상기 SRP가 피크인 위치에 한하여 상기 음원의 위치로 산출하는 단계일 수 있다.
- [0027] 본 실시예에서, 상기 프로세서에 의해, 카메라의 중앙 각도가 상기 평균 각도와 일치하도록 카메라 패닝을 제어하는 단계;를 더 포함할 수 있다.

발명의 효과

- [0028] 본 발명의 실시예들에 따르면, 이벤트가 발생한 위치를 향하여 카메라 패닝을 수행함으로써, 표적 객체를 제대로 감시할 수 있다.
- [0029] 또한, 적은 계산량 또는 적은 리소스로도 표적 객체를 효과적으로 감시할 수 있다.

도면의 간단한 설명

- [0030] 도 1은 일 실시예에 따른 감시 시스템을 설명하기 위한 도면이다.
- 도 2는 일 실시예에 따른 감시 시스템의 구성을 나타내는 블록도이다.
- 도 3은 일 실시예에 따른 감시 시스템의 동작 방법을 설명하기 위한 흐름도이다.
- 도 4는 일 실시예에 따른 기준 값 결정 방법을 설명하기 위한 흐름도이다.
- 도 5는 일 실시예에 따른 음향 모델 생성 방법을 설명하기 위한 흐름도이다.
- 도 6은 일 실시예에 따른 이벤트 판단 방법을 설명하기 위한 흐름도이다.
- 도 7a 및 도 7b는 일 실시예에 따른 카메라 패닝을 위해 산출된 각도에 대하여 설명하기 위한 도면이다.
- 도 8은 일 실시예에 따른 복수의 각도들의 평균 각도 산출 방법을 설명하기 위한 흐름도이다.
- 도 9a 및 도 9b는 일 실시예에 따른 카메라 패닝을 위해 산출된 각도에 대하여 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0031] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시 예를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되

어야 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

- [0032] 이하의 실시예에서, 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들은 용어들에 의해 한정되어서는 안 된다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0033] 이하의 실시예에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 다수의 표현을 포함한다. 이하의 실시예에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0034] 본 발명의 실시예들은 기능적인 블록 구성들 및 다양한 처리 단계들로 나타내어질 수 있다. 이러한 기능 블록들은 특정 기능들을 실행하는 다양한 개수의 하드웨어 또는/및 소프트웨어 구성들로 구현될 수 있다. 예를 들어, 본 발명의 실시예들은 하나 이상의 마이크로프로세서들의 제어 또는 다른 제어 장치들에 의해서 다양한 기능들을 실행할 수 있는, 메모리, 프로세싱, 로직(logic), 룩업 테이블(look-up table) 등과 같은 직접 회로 구성들을 채용할 수 있다. 본 발명의 실시예의 구성 요소들이 소프트웨어 프로그래밍 또는 소프트웨어 요소들로 실행될 수 있는 것과 유사하게, 본 발명의 실시예는 데이터 구조, 프로세스들, 루틴들 또는 다른 프로그래밍 구성들의 조합으로 구현되는 다양한 알고리즘을 포함하여, C, C++, 자바(Java), 어셈블러(assembler) 등과 같은 프로그래밍 또는 스크립팅 언어로 구현될 수 있다. 기능적인 측면들은 하나 이상의 프로세서들에서 실행되는 알고리즘으로 구현될 수 있다. 또한, 본 발명의 실시예들은 전자적인 환경 설정, 신호 처리, 및/또는 데이터 처리 등을 위하여 종래 기술을 채용할 수 있다. 매킨리즘, 요소, 수단, 구성과 같은 용어는 넓게 사용될 수 있으며, 기계적이고 물리적인 구성들로서 한정되는 것은 아니다. 상기 용어는 프로세서 등과 연계하여 소프트웨어의 일련의 처리들(routines)의 의미를 포함할 수 있다.
- [0035] 이하 첨부된 도면들을 참조하여 본 발명의 실시예가 상세히 설명된다.
- [0036] 도 1은 일 실시예에 따른 감시 시스템을 설명하기 위한 도면이다.
- [0037] 도 1을 참조하면, 일 실시예에 따른 감시 시스템은 네트워크 카메라(10), 음향 센서(20), 네트워크(30), 서버(40), 클라이언트 단말(50)을 포함한다.
- [0038] 감시 시스템은 네트워크 카메라(10)의 정보 및 음향 센서(20)의 정보 중 적어도 하나가 네트워크(30)를 통해 서버(40)로 전송되면, 관리자가 클라이언트 단말(50)을 이용하여 서버(40)에 전송된 정보를 모니터링하는 구성과, 관리자가 클라이언트 단말(50)을 이용하여 서버(40)에 명령을 입력하면, 네트워크 카메라(10)가 네트워크(30)를 통해 서버(40)로부터 전송된 명령에 대응하는 동작을 수행하는 구성을 제공할 수 있다.
- [0039] 네트워크 카메라(10)는 패닝(panning)과 틸팅(tilting)이 가능하며 렌즈의 줌(zoom) 배율이 조절 가능한 PTZ 카메라일 수 있다. 패닝 또는 틸팅됨에 따라 감시 영역을 변경할 수 있다.
- [0040] 네트워크 카메라(10)는 배터리로 구동되는 저전력 카메라일 수 있다. 네트워크 카메라(10)는 평상시 슬립 모드를 유지하고, 주기적으로 깨어나 이벤트가 발생하였는지 여부를 체크할 수 있다. 네트워크 카메라(10)는 이벤트가 발생한 경우 액티브 모드로 전환되고, 이벤트가 발생하지 않은 경우 다시 슬립 모드로 복귀할 수 있다. 이와 같이, 이벤트가 발생한 경우에만 액티브 모드를 유지함으로써 네트워크 카메라(10)의 전력 소모를 줄일 수 있다.
- [0041] 네트워크 카메라(10)는 이벤트를 감지하는 센서를 포함할 수 있다. 센서는 감시 또는 보안의 목적으로 센싱 영역에서 이벤트가 발생하는지 여부를 실시간으로 감지할 수 있다. 센서는 음향 센서, 적외선 센서, 모션 센서, 가스 센서, 누수 센서, 온도 센서, 습도 센서, 가속도 센서, 자이로 센서, 촉각 센서, 압력 센서, 진동 센서 등을 포함할 수 있다.
- [0042] 네트워크 카메라(10)는 유무선 LAN(Local Area Network), 와이파이(Wi-Fi), 지그비(ZigBee), 블루투스(Bluetooth), 근거리 통신(Near Field Communication) 등 다양한 통신 방식을 이용하여 게이트웨이(미도시)와 통신할 수 있다. 예를 들어, 네트워크 카메라(10)는 ISM 대역(Industrial Scientific Medical band)의 무선 주파수(Radio Frequency)를 사용하는 저전력 무선 통신 프로토콜에 따라 게이트웨이(미도시)와 통신할 수 있다.

- [0043] 네트워크 카메라(10)는 하나 이상 구비될 수 있다.
- [0044] 음향 센서(20)는 음원으로부터 발생하는 음향 신호를 감지할 수 있다. 음향 센서(20)는 하나 이상 구비될 수 있다. 본 실시예에 따른 감시 시스템은 적어도 두 개의 음향 센서들을 구비할 수 있다. 음향 센서(20)는 네트워크 카메라(10)에 내장될 수 있다.
- [0045] 네트워크(30)는 유선 네트워크 또는 무선 네트워크를 포함할 수 있다. 무선 네트워크는 2G(Generation) 또는 3G 셀룰러 통신 시스템, 3GPP(3rd Generation Partnership Project), 4G 통신 시스템, LTE(Long-Term Evolution), WiMAX(World Interoperability for Microwave Access) 등일 수 있다.
- [0046] 서버(40)는 네트워크(30)를 통해, 네트워크 카메라(10), 음향 센서(20) 또는 게이트웨이(미도시)로부터 전송된 정보에 기초하여 클라이언트 단말(50)에 알림을 전송할 수 있고, 클라이언트 단말(50)로부터 전송된 명령을 네트워크 카메라(10) 또는 게이트웨이(미도시)에 전송할 수 있다.
- [0047] 클라이언트 단말(50)은 서버(40)로부터 전송된 정보를 디스플레이할 수 있고, 저장할 수도 있다. 예를 들어, 클라이언트 단말(50)은 서버(40)로부터 전송된 알림을 디스플레이할 수 있다. 클라이언트 단말(50)은 적어도 하나 이상의 프로세서를 포함할 수 있다. 클라이언트 단말(50)은 마이크로 프로세서나 범용 컴퓨터 시스템과 같은 다른 하드웨어 장치에 포함된 형태로 구동될 수 있다. 클라이언트 단말(50)은 개인용 컴퓨터 또는 이동 단말일 수 있다.
- [0048] 클라이언트 단말(50)은 네트워크 카메라(10) 또는 서버(40)의 동작을 제어하기 위한 사용자 인터페이스를 포함할 수 있다.
- [0049] 본 발명의 실시예들에 따른 감시 시스템(100, 도 2)는 하나의 물리적 장치로 구현될 수도 있고, 복수의 물리적 장치가 유기적으로 결합되어 구현될 수도 있다. 이를 위해 감시 시스템(100)에 포함된 구성 중 일부는 어느 하나의 물리적 장치로 구현되거나 설치되고, 나머지 일부는 다른 물리적 장치로 구현되거나 설치될 수도 있다. 이때, 어느 하나의 물리적 장치는 서버(40)의 일부로 구현될 수 있고, 다른 물리적 장치는 클라이언트 단말(50)의 일부로 구현될 수 있다. 한편, 감시 시스템(100)은 네트워크 카메라(10), 서버(40) 또는 클라이언트 단말(50)에 내장될 수도 있고, 네트워크 카메라(10), 서버(40) 또는 클라이언트 단말(50)과 별개로 구비된 장치에 적용될 수도 있다.
- [0050] 도 2는 일 실시예에 따른 감시 시스템의 구성을 나타내는 블록도이다.
- [0051] 도 2를 참조하면, 일 실시예에 따른 감시 시스템(100)은 통신 인터페이스(110), 프로세서(130), 디스플레이 모듈(150), 및 사용자 인터페이스(170)를 포함한다.
- [0052] 통신 인터페이스(110)는 네트워크 장치와 통신을 수행한다. 통신 인터페이스(110)는 유무선 LAN(Local Area Network), 와이파이(Wi-Fi), 지그비(ZigBee), 블루투스(Bluetooth), 근거리 통신(Near Field Communication) 중 적어도 하나의 통신 기능을 수행할 수 있다.
- [0053] 통신 인터페이스(110)는 적어도 두 개의 음향 센서(20)로부터 음향 신호를 수신한다. 이때, 통신 인터페이스(110)는 네트워크 카메라(10)에 내장된 음향 센서(20) 및 네트워크 카메라(10)와 별개로 구비된 음향 센서(20) 중 적어도 두 개로부터 음향 신호를 수신할 수 있다.
- [0054] 프로세서(130)는 감시 시스템(100)의 전반적인 동작을 제어한다.
- [0055] 프로세서(130)는 음향 신호를 버퍼링하고, 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우에 음향 신호의 특징을 추출하고, 음향 모델에 기초하여 음향 신호의 특징으로부터 음향 신호를 판별하고, 판별된 음향 신호가 이벤트에 해당하면 음향 신호로부터 음원의 위치를 검출하고, 검출된 음원의 위치가 복수이면 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하고, 카메라의 중앙 각도가 평균 각도와 일치하도록 카메라 패닝을 제어한다.
- [0056] 프로세서(130)는 음향 신호를 버퍼링하고, 버퍼링된 음향 신호의 에너지를 기준 값과 비교할 수 있다.
- [0057] 이때, 프로세서(130)는 N개의 음향 신호의 프레임들을 버퍼링하고, N개의 음향 신호의 프레임들의 에너지의 평균 값을 기준 값과 비교할 수 있다.
- [0058] 한편, 프로세서(130)는 카메라 패닝 민감도를 선택하는 사용자 입력에 대응하여 버퍼링되는 음향 신호의 프레임 개수를 결정할 수 있다. 예를 들어, 프로세서(130)는 카메라 패닝 민감도를 높게 설정하는 사용자 입력에 대응

하여 버퍼링되는 음향 신호의 프레임 개수를 줄이거나, 카메라 패닝 민감도를 낮게 설정하는 사용자 입력에 대응하여 버퍼링되는 음향 신호의 프레임 개수를 늘릴 수 있다.

- [0059] 한편, 프로세서(130)는 카메라 패닝 민감도를 선택하는 사용자 입력에 대응하여 음향 신호의 특징을 추출하기 위한 기준 값을 결정할 수 있다. 예를 들어, 프로세서(130)는 카메라 패닝 민감도를 높게 설정하는 사용자 입력에 대응하여 기준 값을 낮추거나, 카메라 패닝 민감도를 낮게 설정하는 사용자 입력에 대응하여 기준 값을 높일 수 있다.
- [0060] 프로세서(130)는 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우에 한하여 음향 신호의 MFCC(Mel-frequency cepstral coefficients) 및 스펙트럼 특징 중 적어도 하나를 추출할 수 있다. 스펙트럼 특징(spectral feature)은 음향 신호의 SpectralCentroid, SpectralCrestFactor, SpectralEntropy, SpectralFlatness, SpectralFlux, SpectralRollOff, ZCR(Zero Crossing Rate), RMS(root mean square) 등을 포함할 수 있다.
- [0061] 한편, 프로세서(130)는 음향 신호로부터 복수의 음향 신호의 특징들을 추출할 수 있다.
- [0062] 프로세서(130)는 GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 미리 학습된 음향 모델에 기초하여, 음향 신호의 특징으로부터 음향 신호를 판별할 수 있다. 음향 모델은 예를 들어, 비명, 총소리, 폭발음, 충돌음 등으로 분류될 수 있다.
- [0063] 프로세서(130)는 음향 신호의 특징과 유사도(likelihood)가 가장 높은 음향 모델을 음향 신호로 추측할 수 있다. 즉, 음향 신호의 특징과 유사도가 가장 높은 음향 모델이 음향 신호로서 판별될 수 있다. 복수의 음향 신호의 특징들이 추출된 경우, 프로세서(130)는 복수의 음향 신호들을 판별할 수 있다.
- [0064] 프로세서(130)는 판별된 음향 신호가 이벤트에 해당하는지 여부를 판단할 수 있다. 이벤트는 비명, 총소리, 폭발음, 충돌음, 및 소정 주파수 대역의 음향 신호 중 적어도 하나를 포함할 수 있다.
- [0065] 한편, 프로세서(130)는 이벤트를 선택하는 사용자 입력에 대응하여 판별된 음향 신호가 이벤트에 해당하는지 여부를 판단할 수 있다.
- [0066] 프로세서(130)는 판별된 음향 신호가 이벤트에 해당하면, 음향 신호를 STFT(Short-Time Fourier Transform)하고, STFT된 음향 신호를 PHAT(phase transform)하고, PHAT된 음향 신호의 SRP(steered response power)를 산출하고, SRP가 피크(peak)인 위치를 음원의 위치로 검출할 수 있다. SRP 피크인 위치가 복수인 경우, 프로세서(130)는 복수의 음원의 위치들을 검출할 수 있다.
- [0067] 한편, 프로세서(130)는 소정 SRP를 선택하는 사용자 입력에 대응하여 소정 SRP 이상의 SRP가 피크인 위치에 한하여 음원의 위치로 산출할 수 있다. SRP 피크인 위치가 복수인 경우, 프로세서(130)는 소정 SRP 이상의 SRP가 피크인 위치를 음원의 위치로 산출할 수 있다.
- [0068] 프로세서(130)는 검출된 음원의 위치의 개수를 판단할 수 있다. 검출된 음원의 위치가 복수인 경우, 프로세서(130)는 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출할 수 있다. 평균 각도는 수학식 1과 같이 산출될 수 있다.

수학식 1

$$\theta_{av1} = \frac{1}{M} \sum_{m=1}^M \theta_m$$

[0069]

[0070] 여기서, θ_{av1} 은 동일한 이벤트가 발생한 복수의 음원의 위치들에 대응하는 평균 각도, M은 음원의 위치

의 개수, θ_m 은 m번째 음원의 위치에 대응하는 각도이다.

[0071] 프로세서(130)는 복수의 음원의 위치들에서 발생한 이벤트의 동일성을 판단할 수 있다.

[0072] 복수의 음원의 위치들에서 동일한 이벤트가 발생한 경우, 프로세서(130)는 수학식 1을 이용하여 평균 각도를 산출할 수 있다. 이때, 동일한 이벤트는 음향 모델이 동일한 이벤트, 우선순위가 동일한 이벤트 등을 의미할 수 있으며, 이에 한정하지 않는다.

[0073] 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 프로세서(130)는 수학식 2와 같이 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출할 수 있다.

수학식 2

$$\theta_{av2} = \frac{1}{M} \sum_{m=1}^M w_p \theta_m$$

[0074]

여기서, θ_{av2} 은 서로 다른 이벤트가 발생한 복수의 음원의 위치들에 대응하는 평균 각도, M은 음원의 위치의 개수,

θ_m 은 m번째 음원의 위치에 대응하는 각도, w_p 는 이벤트의 우선순위에 비례하는 가중치이다.

[0076] 한편, 프로세서(130)는 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력에 대응하여, 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치가 부가된 복수의 각도들의 평균 각도를 산출할 수 있다. 이때, 산출된 평균 각도는 우선순위가 낮은 음원의 위치보다 우선순위가 높은 음원의 위치와 가까울 수 있다.

[0077] 디스플레이 모듈(150)은 음향 신호 판별 결과, 이벤트 발생 여부, 카메라 패닝 민감도 선택 화면, 이벤트 선택 화면, 이벤트 우선순위 선택 화면, SRP 선택 화면 중 적어도 하나를 출력할 수 있다.

[0078] 사용자 인터페이스(170)는 사용자 입력을 수신한다.

[0079] 사용자 인터페이스(170)는 카메라 패닝 민감도를 선택하는 사용자 입력을 수신할 수 있다. 예를 들어, 사용자 인터페이스(170)는 카메라 패닝 민감도를 높게 설정하는 사용자 입력 또는 카메라 패닝 민감도를 낮게 설정하는 사용자 입력을 수신할 수 있다.

[0080] 사용자 인터페이스(170)는 이벤트를 선택하는 사용자 입력을 수신할 수 있다. 예를 들어, 사용자 인터페이스(170)는 비명, 총소리, 폭발음, 충돌음, 및 소정 주파수 대역의 음향 신호 중 적어도 하나를 이벤트로서 선택하는 사용자 입력을 수신할 수 있다.

[0081] 사용자 인터페이스(170)는 소정 SRP를 선택하는 사용자 입력을 수신할 수 있다.

[0082] 사용자 인터페이스(170)는 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신할 수 있다. 예를 들어, 사용자 인터페이스(170)는 비명, 총소리, 폭발음, 충돌음, 및 소정 주파수 대역의 음향 신호 순으로 우선순위를 선택하는 사용자 입력을 수신할 수 있다.

[0083] 이처럼, 본 실시예에 따르면, 음향 신호의 파라미터가 기준 값 이상인 경우에 한하여 카메라 패닝을 수행함으로써, 신뢰성이 높은 감시 시스템을 제공할 수 있다.

[0084] 또한, 본 실시예에 따르면, 사용자 선택에 따른 카메라 패닝 민감도 또는 이벤트에 대응하여 카메라 패닝을 수

행함으로써, 사용자의 니즈에 부합하면서 감시가 강화된 감시 시스템을 제공할 수 있다.

- [0085] 또한, 본 실시예에 따르면, 우선순위가 보다 높은 이벤트가 발생한 영역을 보다 상세하게 감시함으로써, 감시 시스템의 감시 효율을 높일 수 있다.
- [0086] 이하에서, 도 3 내지 도 9a 및 도 9b를 참조하여, 실시예들에 따른 감시 시스템의 동작 방법에 대하여 상세하게 설명한다.
- [0087] 도 3은 일 실시예에 따른 감시 시스템의 동작 방법을 설명하기 위한 흐름도이다.
- [0088] 도 4는 일 실시예에 따른 기준 값 결정 방법을 설명하기 위한 흐름도이다.
- [0089] 도 5는 일 실시예에 따른 음향 모델 생성 방법을 설명하기 위한 흐름도이다.
- [0090] 도 6은 일 실시예에 따른 이벤트 판단 방법을 설명하기 위한 흐름도이다.
- [0091] 도 7a 및 도 7b는 일 실시예에 따른 카메라 패닝을 위해 산출된 각도에 대하여 설명하기 위한 도면이다.
- [0092] 도 8은 일 실시예에 따른 복수의 각도들의 평균 각도 산출 방법을 설명하기 위한 흐름도이다.
- [0093] 도 9a 및 도 9b는 일 실시예에 따른 카메라 패닝을 위해 산출된 각도에 대하여 설명하기 위한 도면이다.
- [0094] 도 3을 참조하면, 통신 인터페이스(110)는 적어도 두 개의 음향 센서(20)들로부터 음향 신호를 수신한다(S110).
- [0095] 도시하지 않았으나, 사용자 인터페이스(170)는 카메라 패닝 민감도를 선택하는 사용자 입력을 수신할 수 있다.
- [0096] 이어서, 프로세서(130)는 수신된 음향 신호를 버퍼링한다(S120). 이때, 프로세서(130)는 N개의 음향 신호의 프레임들을 버퍼링할 수 있다. N개의 음향 신호의 프레임들은 카메라 패닝 민감도를 선택하는 사용자 입력에 따라 결정될 수 있다.
- [0097] 도 4를 참조하면, 사용자 인터페이스(170)는 카메라 패닝 민감도를 선택하는 사용자 입력을 수신한다(S31).
- [0098] 이어서, 프로세서(130)는 카메라 패닝 민감도를 선택하는 사용자 입력에 대응하여 음향 신호의 특징을 추출하기 위한 기준 값을 결정한다(S33).
- [0099] 다시 도 3을 참조하면, 프로세서(130)는 버퍼링된 음향 신호의 에너지를 기준 값과 비교한다(S130). 기준 값은 카메라 패닝 민감도를 선택하는 사용자 입력에 따라 결정될 수 있다.
- [0100] 버퍼링된 음향 신호의 에너지가 기준 값 이상인 경우, 프로세서(130)는 음향 신호의 특징을 추출한다(S140). 예를 들어, 프로세서(130)는 2 개의 음향 신호의 특징들을 추출할 수 있다.
- [0101] 한편, 도 5를 참조하면, 통신 인터페이스(110)에서 음향 신호를 수신하면(S51), 프로세서(130)는 음향 신호의 특징을 추출하여(S53), 음향 신호를 모델링함으로써(S55), 음향 모델을 생성한다. 이때, 프로세서(130)는 GMM(Gaussian mixture model) 및 HMM(Hidden Markov Model) 중 적어도 하나를 기반으로 음향 신호를 모델링할 수 있다. 그 결과, 음향 모델은 예를 들어, 비명, 총소리, 폭발음, 충돌음 등으로 분류될 수 있다.
- [0102] 다시 도 3을 참조하면, 프로세서(130)는 음향 모델에 기초하여 음향 신호의 특징으로부터 음향 신호를 판별한다(S150). 예를 들어, 프로세서(130)는 2개의 음향 신호의 특징들에 대응하는 하나의 음향 모델 즉, 비명을 음향 신호로서 판별할 수 있다. 다른 예를 들어, 프로세서(130)는 2개의 음향 신호의 특징들에 대응하는 2개의 음향 모델들 즉, 비명과 총소리를 음향 신호로서 판별할 수 있다.
- [0103] 도 6을 참조하면, 사용자 인터페이스(170)는 이벤트를 선택하는 사용자 입력을 수신한다(S61). 예를 들어, 사용자 인터페이스(170)는 비명과 총소리를 이벤트로서 선택하는 사용자 입력을 수신할 수 있다.
- [0104] 다시 도 3을 참조하면, 프로세서(130)는 판별된 음향 신호가 이벤트에 해당하는지 여부를 판단한다(S160). 예를 들어, 프로세서(130)는 비명과 총소리를 이벤트로 선택하는 사용자 입력에 대응하여 음향 신호로서 판별된 비명과 총소리가 이벤트에 해당한다고 판단할 수 있다.
- [0105] 이어서, 프로세서(130)는 판별된 음향 신호가 이벤트에 해당하면, 음향 신호로부터 음원의 위치를 검출한다(S170).
- [0106] 예를 들어, 사용자 인터페이스(170)가 소정 SRP를 선택하는 사용자 입력을 수신하면, 프로세서(130)는 음향 신호를 STFT(Short-Time Fourier Transform)하고, STFT된 음향 신호를 PHAT(phase transform)하고, PHAT된 음향 신호의 SRP(steered response power)를 산출하고, 사용자 입력에 대응하여 소정 SRP 이상의 적어도 하나의 SRP

가 피크인 위치를 음원의 위치로 검출할 수 있다.

[0107] 도 7a를 참조하면, 프로세서(130)는 각각 비명으로 판별된 제1 음원의 위치(S1) 및 제2 음원의 위치(S2)를 검출할 수 있다.

[0108] 제1 음원의 위치(S1) 및 제2 음원의 위치(S2)는 네트워크 카메라(10)의 화각(11) 내에 존재하며, 네트워크 카메라(10)의 현재 중심 각(θ_r)은 화각(11)의 절반(half)일 수 있다.

[0109] 제1 각도(θ_1)는 제1 음원의 위치(S1)에 대응하고, 제2 각도(θ_2)는 제2 음원의 위치(S2)에 대응할 수 있다.

[0110] 다시 도 3을 참조하면, 프로세서(130)는 검출된 음원의 위치의 개수가 2 이상인지 여부를 판단한다(S180).

[0111] 검출된 음원의 위치가 복수이면, 프로세서(130)는 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출한다(S190).

[0112] 도시하지 않았으나, 프로세서(130)는 복수의 음원의 위치들에서 발생한 이벤트의 동일성을 판단할 수 있다.

[0113] 복수의 음원의 위치들에서 동일한 이벤트가 발생한 경우, 프로세서(130)는 수학적 1을 이용하여 평균 각도를 산출할 수 있다.

[0114] 이어서, 프로세서(130)는 카메라의 중앙 각도가 평균 각도와 일치하도록 카메라 패닝을 제어한다(S200).

[0115] 도 7b를 참조하면, 프로세서(130)는 네트워크 카메라(10)의 목표 중심 각(θ'_r)이 수학적 1을 통해 산출된 동일한 이벤트가 발생한 복수의 음원의 위치들에 대응하는 평균 각도(θ_{av1})와 일치하도록 카메라 패닝을 제어할 수 있다. 그 결과, 네트워크 카메라(10)의 목표 중심 각(θ'_r)이 제1 음원의 위치(S1)와 제2 음원의 위치(S2)의 사이에 존재할 수 있다.

[0116] 한편, 도 8을 참조하면, 사용자 인터페이스(170)는 서로 다른 이벤트들 간의 우선순위를 결정하는 사용자 입력을 수신한다(S191).

[0117] 복수의 음원의 위치들에서 각각 서로 다른 이벤트가 발생한 경우, 프로세서(130)는 사용자 입력에 대응하여, 수학적 2를 이용하여 서로 다른 이벤트들 간의 우선순위에 비례하는 가중치를 부과하여 복수의 음원의 위치들에 대응하는 복수의 각도들의 평균 각도를 산출하고(S193), 카메라의 중앙 각도가 평균 각도와 일치하도록 카메라 패닝을 제어한다(S200, 도 3).

[0118] 도 9a를 참조하면, 프로세서(130)는 비명으로 판별된 제1 음원의 위치(S1) 및 총소리로 판별된 제2 음원의 위치(S2)를 검출할 수 있다.

[0119] 제1 음원의 위치(S1) 및 제2 음원의 위치(S2)는 네트워크 카메라(10)의 화각(11) 내에 존재하며, 네트워크 카메라(10)의 현재 중심 각(θ_r)은 화각(11)의 절반(half)일 수 있다.

[0120] 제1 각도(θ_1)는 제1 음원의 위치(S1)에 대응하고, 제2 각도(θ_2)는 제2 음원의 위치(S2)에 대응할 수 있다.

[0121] 도 9b를 참조하면, 프로세서(130)는 네트워크 카메라(10)의 목표 중심 각(θ'_r)이 수학적 2를 통해 산출된 서

로 다른 이벤트가 발생한 복수의 음원의 위치들에 대응하는 평균 각도(θ_{av2})와 일치하도록 카메라 패닝을 제어할 수 있다. 그 결과, 네트워크 카메라(10)의 목표 중심 각(θ'_r)은 제1 음원의 위치(S1)와 제2 음원의 위치(S2)의 사이에 존재하되, 우선순위가 낮은 제2 음원의 위치(S2)보다 우선순위가 높은 제1 음원의 위치(S1)와 가까울 수 있다.

[0122] 다시 도 3을 참조하면, 프로세서(130)는 검출된 음원의 위치의 개수가 하나인 경우, 검출된 음원의 위치를 향하도록 카메라 패닝을 제어한다(S210). 다만, S170 단계에서 음원의 위치가 검출되지 않는 경우, 프로세서(130)는 카메라 패닝 제어를 중단할 수 있다.

[0123] 이제까지 본 발명에 대하여 바람직한 실시예를 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 본 발명을 구현할 수 있음을 이해할 것이다.

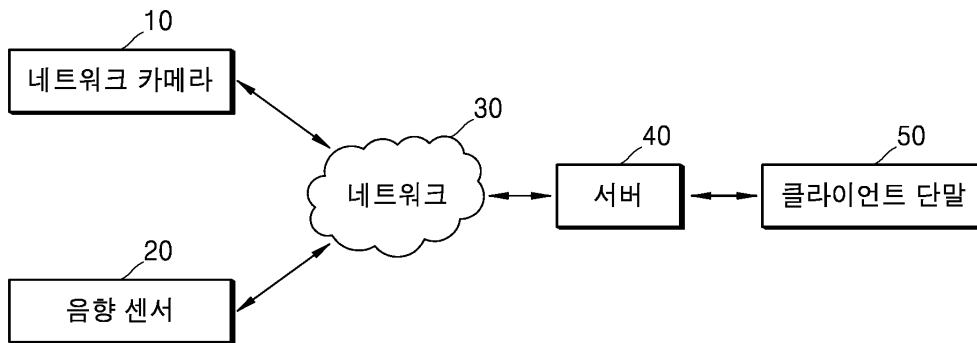
[0124] 그러므로 상기 개시된 실시예는 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 특허청구범위에 의해 청구된 발명 및 청구된 발명과 균등한 발명들은 본 발명에 포함된 것으로 해석되어야 한다.

부호의 설명

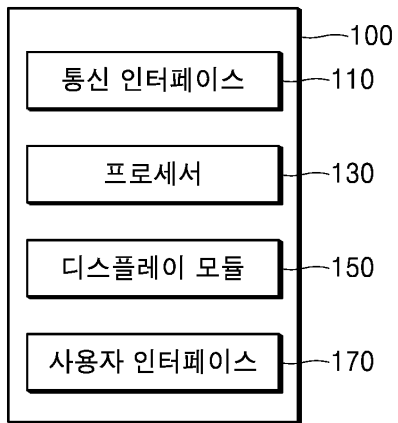
- [0125] 10: 네트워크 카메라
- 20: 음향 센서
- 30: 네트워크
- 40: 서버
- 50: 클라이언트 단말

도면

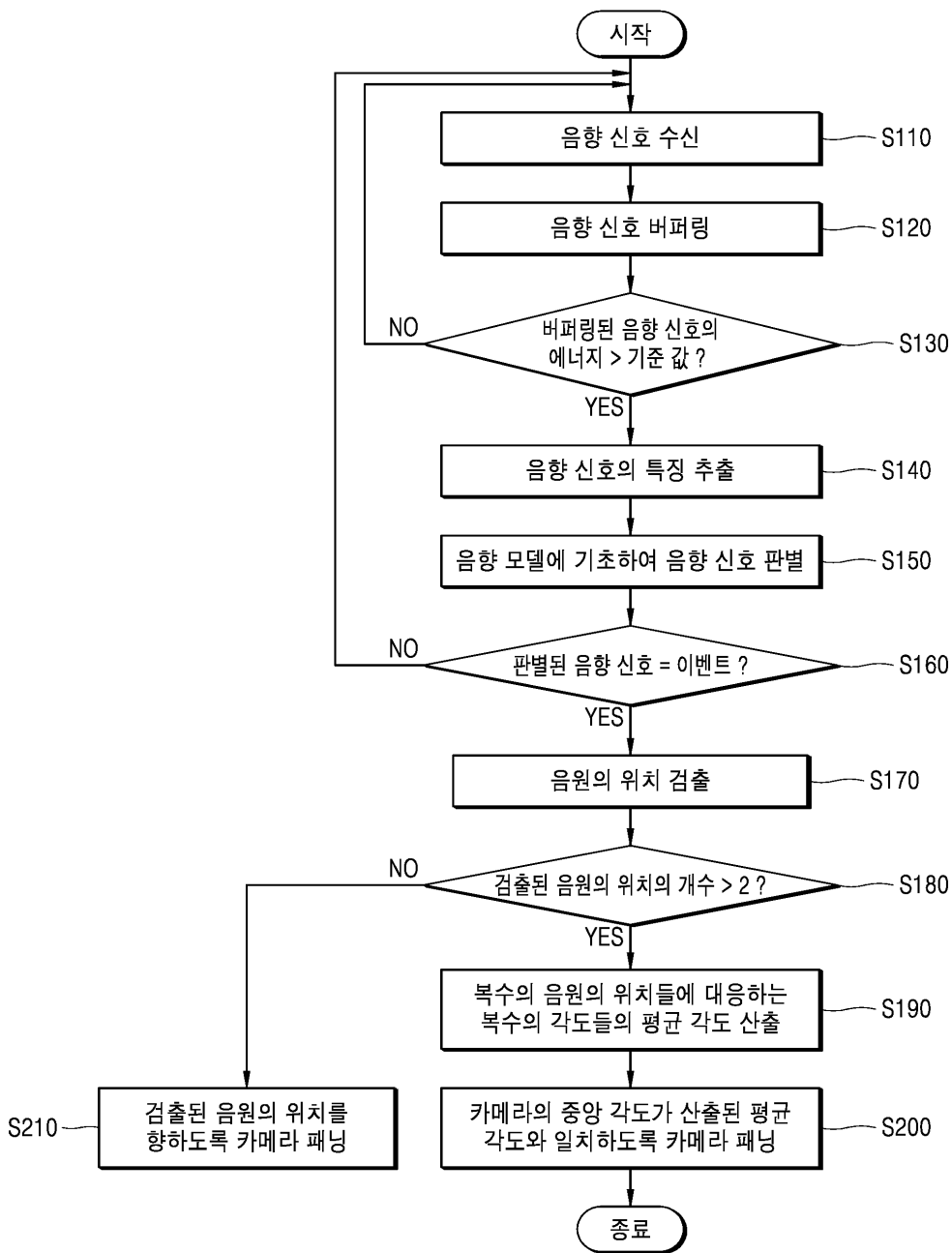
도면1



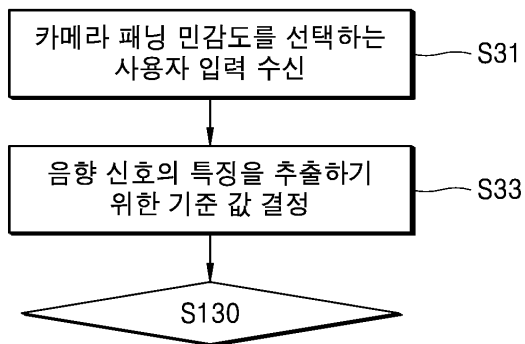
도면2



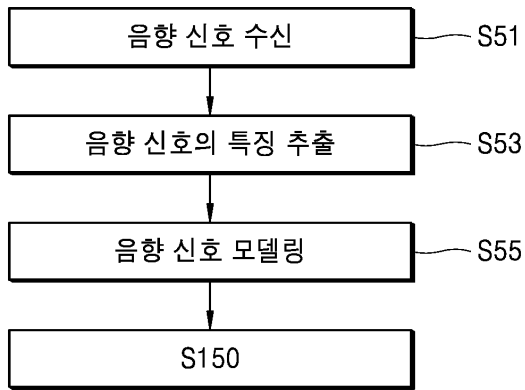
도면3



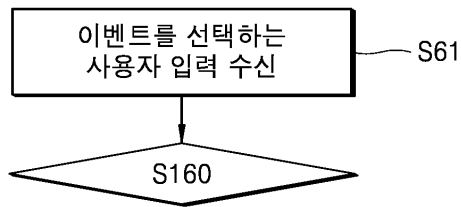
도면4



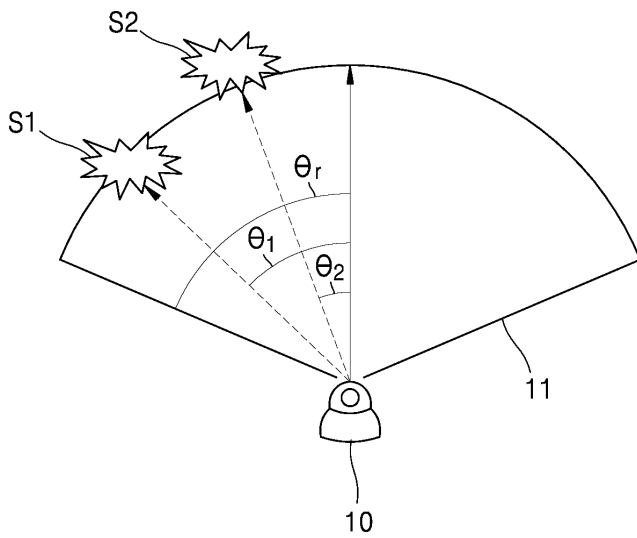
도면5



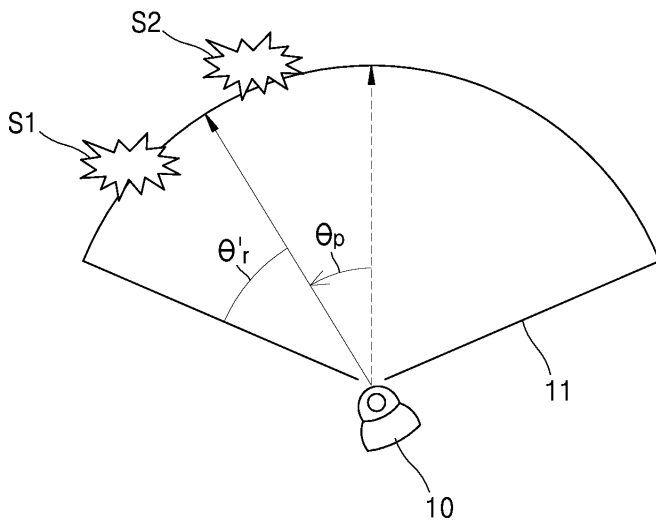
도면6



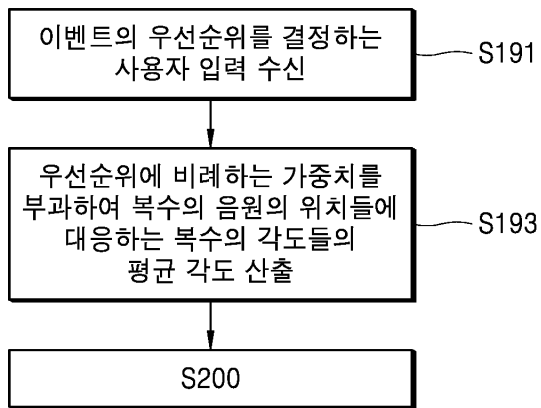
도면7a



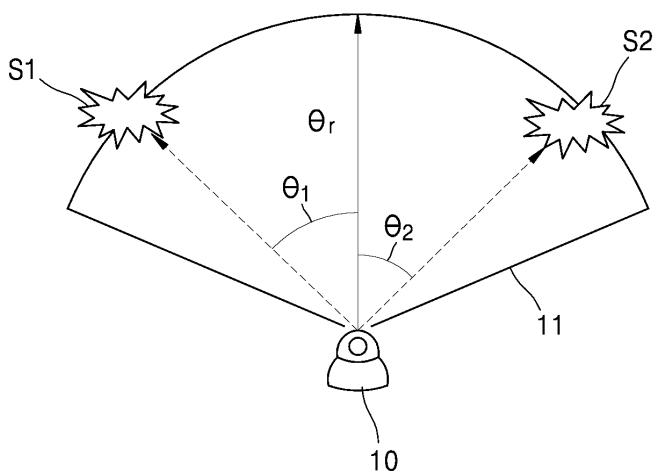
도면7b



도면8



도면9a



도면9b

