(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2016년10월24일
(11) 등록번호 10-1664733
(24) 등록일자 2016년10월04일

(51) 국제특허분류(Int. Cl.)
*H04N 5/77* (2006.01) *G08B 13/196* (2006.01)
*H04N 5/232* (2006.01) *H04N 7/18* (2006.01)
(52) CPC특허분류
*H04N 5/77* (2013.01)
*G08B 13/19608* (2013.01)
(21) 출원번호 10-2015-0113100
(22) 출원일자 2015년08월11일
심사청구일자 2015년08월11일
(56) 선행기술조사문헌
KR1020130000660 A*
JP2002123878 A*
JP20060191411 A*
KR1020090118718 A*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
**고려대학교 산학협력단**
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
**육동석**
서울특별시 용산구 이촌로 347, 1동 102호(서빙고동, 신동아아파트)
**이태우**
서울특별시 중구 서애로1길 34, 711호 (묵정동, 묵정아파트)
(74) 대리인
**송인호, 최관락**

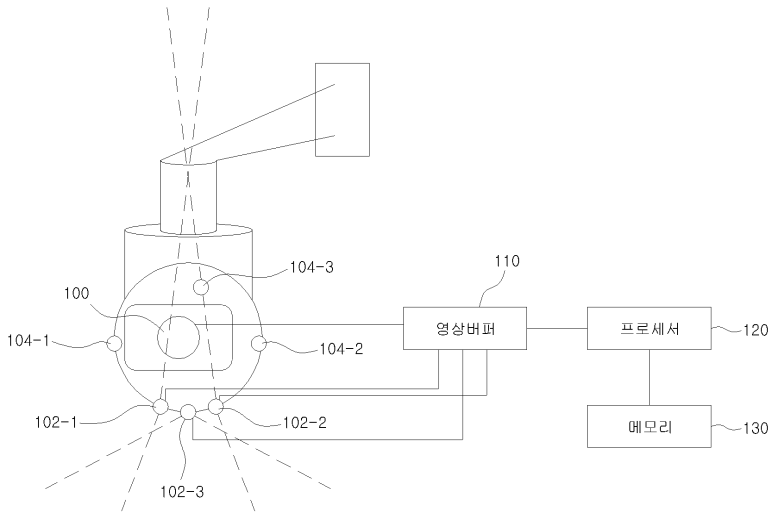
전체 청구항 수 : 총 4 항 심사관 : 이병우

(54) 발명의 명칭 **전 방향 고해상도 추적 녹화 장치 및 방법**

## (57) 요 약

본 발명은 전 방향 고해상도 추적 녹화 장치 및 방법을 개시한다. 본 발명의 일 실시 예에 따르면, 회전 및 줌인이 가능한 팬 틸트 줌 카메라; 어안렌즈를 갖는 복수의 보조 카메라; 빔 조향이 가능한 복수의 마이크로폰; 및 상기 팬 틸트 줌 카메라 및 상기 복수의 보조 카메라에서 촬영된 영상과 상기 복수의 마이크로폰에 녹음된 음원을 분석하여 녹화된 영상에 있는 중요 객체의 위치를 탐색하고, 탐색 된 위치로 상기 팬 틸트 줌 카메라를 회전 및 줌인 하며, 미리 설정된 알고리즘을 통해 음원 방향을 탐색하여 상기 음원 방향에서 발생한 소리가 증폭되도록 제어하는 프로세서를 포함하는 고해상도 추적 녹화 장치가 제공된다.

*대 표 도* - 도1



(52) CPC특허분류
*G08B 13/1963* (2013.01)
*H04N 5/23296* (2013.01)
*H04N 7/18* (2013.01)

공지예외적용 : 있음

# 명 세 서

## 청구범위

### 청구항 1

고해상도 추적 녹화 장치로서,

회전 및 줌인이 가능한 팬 틸트 줌 카메라;

어안렌즈를 갖는 복수의 보조 카메라;

빔 조향이 가능한 복수의 마이크로폰; 및

상기 팬 틸트 줌 카메라 및 상기 복수의 보조 카메라에서 촬영된 영상과 상기 복수의 마이크로폰에 녹음된 음원을 분석하여 녹화된 영상에 있는 중요 객체의 위치를 탐색하고, 탐색 된 위치로 상기 팬 틸트 줌 카메라를 회전 및 줌인 하며, 미리 설정된 알고리즘을 통해 음원 방향을 탐색하여 상기 음원 방향에서 발생한 소리가 증폭되도록 제어한 후, 녹화가 시작되도록 제어하는 프로세서를 포함하되,

상기 프로세서는, 빔포머의 미스얼라인먼트에 따른 위상 미스얼라인먼트가 2π/5보다 작도록 전체 탐색 범위를 일련의 그룹으로 분할하고, 각 그룹의 대표 클러스터에 대해서 최대 파워를 스캔하여 전역 최대 파워 좌표를 포함하는 그룹을 탐색함으로써,

음원 방향 탐색 범위에 있는 좌표들 중 모든 마이크로폰 쌍에 대해 동일한 도착 지연 시간을 갖는 좌표들은 복수의 클러스터로 클러스터링되고,

상기 클러스터 중 상기 모든 마이크로폰 쌍에 대하여, 상기 2π/5의 위상 미스얼라인먼트에 상응하는 지연 시간보다 작은 도착 지연 시간을 갖는 클러스터들은 복수의 그룹으로 클러스터링되는 고해상도 추적 녹화 장치.

### 청구항 2

삭제

### 청구항 3

삭제

### 청구항 4

삭제

### 청구항 5

제1항에 있어서,

상기 프로세서는,

상기 복수의 그룹의 각 대표 클러스터에 대한 출력 파워를 계산하고,

상기 출력 파워가 큰 상위 n개(n은 1이상의 자연수)의 그룹을 탐색하고,

상기 상위 n개의 최대 출력 파워 그룹에 속하는 모든 좌표들 중 최대 출력 파워를 갖는 좌표의 방향을 상기 음원 방향으로 결정하는 고해상도 추적 녹화 장치.

### 청구항 6

고해상도 추적 녹화 방법으로서,

회전 및 줌인이 가능한 팬 틸트 줌 카메라 및 어안렌즈를 갖는 복수의 보조 카메라로부터 촬영 영상을 획득하는 단계;

복수의 마이크로폰을 이용하여 소리를 획득하는 단계;

상기 팬 틸트 줌 카메라 및 상기 복수의 보조 카메라에서 촬영된 영상과 상기 복수의 마이크로폰에 녹음된 음원을 분석하여 녹화된 영상에 있는 중요 객체의 위치를 탐색하는 단계;

상기 탐색 된 위치로 상기 팬 틸트 줌 카메라를 회전 및 줌인 하는 단계;

미리 설정된 알고리즘을 통해 음원 방향을 탐색하는 단계;

상기 탐색 된 음원 방향에서 발생한 소리를 증폭하는 단계; 및

녹화를 시작하고, 상기 발생한 소리를 저장하는 단계를 포함하되,

상기 미리 설정된 알고리즘을 통해 음원 방향을 탐색하는 단계는,

빔포머의 미스얼라인먼트에 따른 위상 미스얼라인먼트가 2π/5보다 작도록 전체 탐색 범위를 일련의 그룹으로 분할하고, 각 그룹의 대표 클러스터에 대해서 최대 파워를 스캔하여 전역 최대 파워 좌표를 포함하는 그룹을 탐색함으로써,

음원 방향 탐색 범위에 있는 좌표들 중 모든 마이크로폰 쌍에 대해 동일한 도착 지연 시간을 갖는 좌표들은 복수의 클러스터로 클러스터링되고,

상기 클러스터 중 상기 모든 마이크로폰 쌍에 대하여, 상기 2π/5의 위상 미스얼라인먼트에 상응하는 지연 시간보다 작은 도착 지연 시간을 갖는 클러스터들은 복수의 그룹으로 클러스터링되는 고해상도 추적 녹화 방법.

**청구항 7**

삭제

**청구항 8**

제6항에 있어서,

상기 음원 방향 탐색 단계는,

상기 복수의 그룹의 각 대표 클러스터에 대한 출력 파워를 계산하는 단계;

상기 출력 파워가 큰 상위 n개(n은 1이상의 자연수)의 그룹을 탐색하는 단계;

상기 상위 n개의 최대 출력 파워 그룹에 속하는 모든 좌표들 중 최대 출력 파워를 갖는 좌표의 방향을 상기 음원 방향으로 결정하는 단계를 포함하는 고해상도 추적 녹화 방법.

## *발명의 설명*

### *기 술 분 야*

[0001] 본 발명은 전 방향 고해상도 추적 녹화 장치 및 방법에 관한 것이다.

### *배 경 기 술*

[0002] 카메라 렌즈의 물리적 한계 때문에 CCTV는 제한된 화각을 가진다.

[0003] 따라서 개방된 공간에서 넓은 각도를 감시하려면 감시하고자 하는 각도를 카메라 화각으로 나눈 수만큼 많은 CCTV를 여러 방향으로 설치해야만 한다.

[0004] 그러나 그 카메라들을 모두 고해상도 카메라로 배치하면 비용이 너무 많이 든다. 고해상도 카메라 그리고 촬영된 고해상도 영상들을 동시에 처리하고 저장하는 장치들이 비싸기 때문이다.

[0005] 그러므로 개방된 넓은 공간을 가격 면에서 가장 효율적으로 감시하기 위한 CCTV 배치에 관한 방안이 필요하다.

[0006] CCTV로 촬영된 장면(영상과 소리) 대부분은 사용되지 않고 버려진다. 그러므로 영상 대부분은 고해상도로 촬영될 필요가 없다.

[0007] 한국 공개특허 제2010-0103109호(발명의 명칭: 전 방위 카메라 및 팬/틸트/줌 카메라를 이용한 무인 감시 장치 및 방법)은 전 방위 팬/틸트/줌 카메라가 감시 대상 객체를 일련의 처리 과정으로 추적하여 촬영하는 방법을 개시하고 있다.

[0008] 그러나 종래기술들은 전 방향 영상을 촬영한다는 점은 개시하고는 있을 뿐, 필요한 영상만 고해상도로 촬영하는 방안은 제시하고 있지 않다.

[0009] 또한, 고해상도로 녹화하고자 하는 물체가 있는 위치에서 수신되는 음원을 증폭하는 방안도 제시하고 있지 않다.

## 발명의 내용

### 해결하려는 과제

[0010] 상기한 종래기술의 문제점을 해결하기 위해, 본 발명은 저비용으로 필요한 영역에 대해서만 고해상도 영상을 획득할 수 있는 전 방향 고해상도 추적 녹화 장치 및 방법을 제안하고자 한다.

### 과제의 해결 수단

[0011] 상기 목적을 달성하기 위해 본 발명의 일 실시 예에 따르면, 고해상도 추적 녹화 장치로서, 회전 및 줌인이 가능한 팬 틸트 줌 카메라; 어안렌즈를 갖는 복수의 보조 카메라; 빔 조향이 가능한 복수의 마이크로폰; 및 상기 팬 틸트 줌 카메라 및 상기 복수의 보조 카메라에서 촬영된 영상과 상기 복수의 마이크로폰에 녹음된 음원을 분석하여 녹화된 영상에 있는 중요 객체의 위치를 탐색하고, 탐색 된 위치로 상기 팬 틸트 줌 카메라를 회전 및 줌인 하며, 미리 설정된 알고리즘을 통해 음원 방향을 탐색하여 상기 음원 방향에서 발생한 소리가 증폭되도록 제어하는 프로세서를 포함하는 고해상도 추적 녹화 장치가 제공된다.

[0012] 상기 프로세서는 상기 팬 틸트 줌 카메라가 상기 객체 위치로 회전 및 줌인 된 이후에 녹화가 시작되도록 제어할 수 있다.

[0013] 음원 방향 탐색 범위에 있는 좌표들 중 모든 마이크로폰 쌍에 대해 동일한 도착 지연 시간을 갖는 좌표들은 복수의 클러스터로 클러스터링되고, 상기 클러스터 중 상기 모든 마이크로폰 쌍에 대해 미리 설정된 크기보다 작은 도착 지연 시간을 갖는 클러스터들은 복수의 그룹으로 클러스터링 된다.

[0014] 상기 미리 설정된 크기는 $2\pi/5$ 미스얼라인먼트로 정해질 수 있다.

[0015] 상기 프로세서는, 상기 복수의 그룹의 각 대표 클러스터에 대한 출력 파워를 계산하고, 상기 출력 파워가 큰 상위 n개(n은 1이상의 자연수)의 그룹을 탐색하고, 상기 상위 n개의 최대 출력 파워 그룹에 속하는 모든 좌표들 중 최대 출력 파워를 갖는 좌표의 방향을 상기 음원 방향으로 결정할 수 있다.

[0016] 본 발명의 다른 측면에 따르면, 고해상도 추적 녹화 방법으로서, 회전 및 줌인이 가능한 팬 틸트 줌 카메라 및 어안렌즈를 갖는 복수의 보조 카메라로부터 촬영 영상을 획득하는 단계; 복수의 마이크로폰을 이용하여 소리를 획득하는 단계; 상기 팬 틸트 줌 카메라 및 상기 복수의 보조 카메라에서 촬영된 영상과 상기 복수의 마이크로폰에 녹음된 음원을 분석하여 녹화된 영상에 있는 중요 객체의 위치를 탐색하는 단계; 상기 탐색 된 위치로 상기 팬 틸트 줌 카메라를 회전 및 줌인 하는 단계; 미리 설정된 알고리즘을 통해 음원 방향을 탐색하는 단계; 및 상기 탐색 된 음원 방향에서 발생한 소리를 증폭하는 단계를 포함하는 고해상도 추적 녹화 방법이 제공된다.

### 발명의 효과

[0017] 본 발명에 따르면, 하나의 팬 틸트 줌 카메라 및 어안렌즈를 갖는 복수의 카메라를 통해 필요한 위치에 대해서만 고해상도로 영상을 녹화할 수 있다.

## 도면의 간단한 설명

[0018] 도 1은 본 발명의 바람직한 일 실시 예에 따른 전 방향 고해상도 추적 녹화 장치의 구성을 도시한 도면이다.

도 2는 본 실시 예에 따른 고해상도 추적 녹화 방법의 순서도이다.

도 3은 출력 파워(SRP-PHAT)를 방위각(azimuth) 및 고도각(elevation)에 따라 출력한 에너지 맵의 일 예를 도시한 것이다.

도 4는 본 발명의 일 실시 예에 따른 음원 방향 탐색 과정을 도시한 순서도이다.

도 5는 본 발명의 일 실시 예에 따른 음원 방향 탐색 과정을 설명하기 위한 도면이다.

## *발명을 실시하기 위한 구체적인 내용*

[0019] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시 예를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세하게 설명하고자 한다.

[0020] 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.

[0021] 이하에서, 본 발명에 따른 실시 예들을 첨부된 도면을 참조하여 상세하게 설명한다.

[0022] 도 1은 본 발명의 바람직한 일 실시 예에 따른 전 방향 고해상도 추적 녹화 장치의 구성을 도시한 도면이다.

[0023] 도 1에 도시된 바와 같이, 본 실시 예에 따른 전 방향 고해상도 추적 녹화 장치는 고해상도 팬 틸트 줌 카메라(100), 어안렌즈가 장착된 복수의 카메라(102-1 내지 102-3, 이하 '102'로 통칭함), 복수의 마이크로폰(104-1 내지 104-3)을 포함할 수 있다.

[0024] 여기서, 팬 틸트 줌 카메라(100)는 정해진 시야각을 가지며 고해상도 영상 촬영이 가능한 카메라이며, 어안렌즈가 장착된 복수의 카메라(102)는 전 방향 촬영이 가능하며 저해상도로 영상 촬영이 가능한 카메라이다.

[0025] 본 실시 예에 따르면, 소리가 발생하는 방향(음원 방향)을 탐색하고, 탐색 된 방향에서 오는 소리를 증폭하기 위해 복수의 마이크로폰이 배치될 수 있다.

[0026] 복수의 마이크로폰을 이용한 음원 방향 탐색 과정에 대해서는 이하에서 상세하게 설명할 것이다.

[0027] 도 1에 도시된 바와 같이, 본 실시 예에 따른 장치는 영상 분석을 통해 이벤트가 발생한 위치(영상 내 객체의 움직임이 발생하여 촬영이 필요한 위치)를 탐색하고, 또한 음원 방향 탐색을 통해 음원으로부터 수신되는 소리를 증폭할 수 있도록 하는 프로세서(120)를 포함하며,프로세서(120)는 메모리(130)에 저장된 프로그램 명령어들을 이용하여 녹화 위치 탐색 및 음원 방향 탐색을 수행한다.

[0028] 또한 프로세서(120)는 탐색 된 결과에 따라 팬 틸트 줌 카메라(100)의 회전 또는 줌인(zoom in)을 제어한다.

[0029] 또한, 프로세서(120)는 움직임이 발생한 위치(녹화 위치)에서 획득된 영상이 저장부(미도시)에 녹화되도록 제어한다.

[0030] 도 2는 본 실시 예에 따른 고해상도 추적 녹화 방법의 순서도이다.

[0031] 도 2를 참조하면, 고해상도 추적 녹화 장치는 복수의 카메라(100,102)를 통해 영상을 획득하고(단계 200), 획득된 영상 분석을 통해 영상의 변화가 있는지 여부를 판단한다(단계 202).

[0032] 영상 버퍼(110)에는 직전에 촬영된 영상이 저장되며, 단계 202에서 현재 시점에 획득된 영상과 이전 영상을 비교하여 영상의 변화가 있는지 여부를 판단한다.

[0033] 예를 들어, 영상 내에서 객체의 큰 움직임이 있는 경우, 직전 영상과 큰 차이를 보일 수 있으며, 영상 분석 알고리즘을 통해 이벤트가 발생(위급 상황 발생)했는지 여부를 판단할 수 있다.

[0034] 보다 상세하게, 고해상도 추적 녹화 장치는 직전 영상과 현재 영상에서 각 픽셀의 색 변화량을 측정하고, 변화량이 임계치를 넘는 픽셀들을 추출한다.

[0035] 고해상도 녹화 추적 장치는 추출된 픽셀들을 군집화하여 군집들 중 가장 규모가 큰 군집의 방향으로 팬 틸트 줌 카메라(100)를 회전시키고, 줌인 되도록 한다(단계 204).

[0036] 상기와 같이 중요 객체의 위치를 탐색하기 위해, 영상 분석뿐만 아니라, 복수의 마이크로폰(104)을 통해 녹음된 음원이 분석이 함께 이루어질 수 있다.

[0037] 한편, 본 발명의 일 실시 예에 따르면, 팬 틸트 줌 카메라(100)가 회전하여 바라보는 방향에서 수신되는 소리를 증폭하기 위해 음원 방향을 탐색한다.

[0038] 고해상도 추적 녹화 장치는 복수의 마이크로폰(104)를 통해 주변 소리를 획득하고(단계 206), 미리 설정된 알고리즘을 이용하여 음원 방향을 탐색한다(단계 208).

[0039] 이후, 탐색 된 음원 방향에서 수신된 소리를 증폭한다(단계 210).

[0040] 본 실시 예에 따르면, 이벤트가 발생한 위치에서의 영상만을 선택적으로 녹화하고, 또한 해당 위치에서의 소리만을 증폭하여 저장하기 때문에 감시 효율을 한층 높일 수 있다.

[0041] 이하에서는 본 발명의 일 실시 예에 따른 음원 방향 탐색 알고리즘에 대해 상세하게 설명한다.

[0042] 본 실시 예에 따른 음원 방향 탐색 알고리즘은 2 단계 탐색 공간 클러스터링을 이용한 고속 음원 방향 측정 방법(twolevel search space clustering: TL-SSC)으로 정의될 수 있다.

[0043] 제안된 방법은 룩업 테이블에서 좌표들을 일련의 그룹으로 나누고, 전역 최대 파워 영역을 포함할 수 있는 적은 수의 그룹들을 탐색한다.

[0044] 이는 방향 탐색의 정확도를 보장하면서 탐색 후보 좌표의 수를 급격히 감소시킬 수 있다.

[0045] 본 실시 예에 따른 TL-SSC를 설명하기에 앞서, SRP-PHAT(steered response power phase transform) 기반 SSL(sound source localization)에 대해 우선 설명한다.

[0046] SRP-PHAT 기반 SSL은 최대 출력 파워 위치를 탐색하기 위해 마이크로폰 배열의 빔 조향(beam steering)을 수행한다.

[0047] 시간 t, 위치 q에서 포커스된 delay-and-sum 빔포머의 출력 y(t,q)는 다음과 같다.

*수학식 1*

[0048] $$y(t,q)=\sum_{m=1}^{M} x_m(t+\tau_{m,q})$$

[0049] 여기서, M은 마이크로폰 수, $x_m(t)$는 m번째 마이크로폰에서 시간 t에 수신된 신호이고, $\tau_{m,q}$는 위치 q에서 마이크로폰 m까지 소리가 이동하는 데 걸린 시간이다.

[0050] 가중 함수를 이용한 filter-and-sum 빔포머는 방의 반향(room reverberation)에 의해 야기되는 복잡한 노이즈를 극복하기 위해 사용될 수 있다.

[0051] 주파수 도메인에서 수학식 1의 filter-and sum 버전은 수학식 2와 같다.

*수학식 2*

[0052] $$Y(\omega,q)=\sum_{m=1}^{M} G_m(\omega)X_m(\omega)e^{j\omega\tau_{m,q}}$$

[0053] 여기서, $X_m(\omega)$와 $G_m(\omega)$는 각각 m번째 마이크로폰 신호의 푸리에 변환 및 임펄스 응답이다.

[0054] 수학식 2에서, 마이크로폰 신호는 필터가 적용된 이후에 이동 지연 및 합산에 의해 위상 정렬된다.

[0055] SRP-PHAT 기반 SSL은 마이크로폰의 빔 조향을 통해 각 공간 지점 q에 초점을 맞추며, q에서의 출력 파워 P(q)는 수학식 3과 같이 계산된다.

*수학식 3*

$$
\begin{aligned}
P(q) &= \int_{-\infty}^{\infty} |Y(\omega, q)|^2 d\omega \\
&= \int_{-\infty}^{\infty} \left( \sum_{l=1}^{M} G_l(\omega) X_l(\omega) e^{j\omega\tau_{l,q}} \right) \\
&\qquad \times \left( \sum_{k=1}^{M} G_k^*(\omega) X_k^*(\omega) e^{-j\omega\tau_{k,q}} \right) d\omega \\
&= \sum_{l=1}^{M} \sum_{k=1}^{M} \int_{-\infty}^{\infty} \Psi_{lk}(\omega) X_l(\omega) X_k^*(\omega) e^{j\omega(\tau_{l,q}-\tau_{k,q})} d\omega
\end{aligned}
$$

[0056]

[0057] 여기서, $\Psi_{lk}(\omega)=G_l(\omega)G_k^*(\omega)$ 이고, 반향 효과를 감소시키기 위해 수학식 4와 같이 정의된다.

*수학식 4*

$$
\Psi_{lk}(\omega) \equiv \frac{1}{\left| X_l(\omega) X_k^*(\omega) \right|}.
$$

[0058]

[0059] 각 후보 위치에 대해 조향된 응답 파워(steered response power) P(q)가 계산된 이후, 최대 출력 파워를 갖는 지점은 수학식 5와 같이 음원 방향으로 결정된다.

*수학식 5*

$$
\hat{q} = \arg\max_{q} P(q).
$$

[0060]

[0061] 그러나 수학식 5에서와 같이 모든 후보 위치에서 탐색 과정이 수행되기 때문에 SRP-PHAT 기반 SSL을 매우 높은 계산 비용을 갖는 문제점이 있다.

[0062] 종래기술에서 설명된 확장성(scalability) 문제를 해결하기 위해, 본 발명은 TL-SSC를 제안한다.

[0063] 이 방법은 출력 파워가 전역 최대 파워 위치에서 높다는 점을 이용하여 착안된 것이다.

[0064] 도 3은 출력 파워가 방위각(azimuth) 및 고도각(elevation)의 함수에 따라 그려진 SRP-PHAT 에너지 맵의 일 예를 도시한 것이다.

[0065] 색이 어두울수록 더 높은 출력 파워를 나타내며, 십자 표시는 실제 음원 방향을 나타낸다.

[0066] 여기서, 음원 방향 주변의 어두운 영역이 상당히 넓다는 것을 확인할 수 있다.

[0067] 이는 빔포머가 약간 잘못 정렬되더라도 빔포머의 출력 파워가 여전히 높게 유지됨을 보여준다. 따라서 $2\pi/5$보다 작은 위상 미스얼라인먼트를 가지는 좌표들도 충분히 높은 출력 파워를 갖는 정답 좌표로 고려될 수 있다.

[0068] 따라서 미스얼라인먼트가 $2\pi/5$보다 작도록 전체 탐색 범위를 일련의 그룹으로 분할하고, 각 그룹의 대표 클러스터에 대해서만 최대 파워를 스캔하여도 전역 최대 파워 좌표를 포함하는 그룹을 탐색할 수 있다.

[0069] 보다 상세하게, 본 실시 예에 따르면, 음원 방향 탐색 범위에 있는 좌표들 중 모든 마이크로폰 쌍에 대해 동일한 도착 지연 시간을 갖는 좌표들은 복수의 클러스터로 클러스터링되고, 상기 클러스터 중 상기 모든 마이크로폰 쌍에 대해 미리 설정된 크기보다 작은 도착 지연 시간을 갖는 클러스터들은 복수의 그룹으로 클러스터링 된다.

[0070] 여기서, 미리 설정된 크기는 $2\pi/5$ 미스얼라인먼트로 정해진다.

[0071] 본 실시 예에 따르면, 룩업 테이블의 좌표들은 수학식 6과 같이 2 단계에서 다시 그룹화된다.

*수학식 6*

$$\sum_{i=1}^{I}\sum_{j=1}^{J} p(q_i \in z_j)\, e(q_i, z_j) = 0$$

$$e(q_i, z_j) = \sum_{k=1}^{M}\sum_{l=1}^{M} d(s_{k,l,q_i}, s_{k,l,z_j})$$

$$d(s_{k,l,q_i}, s_{k,l,z_j}) = \begin{cases} 0, & \text{if } |s_{k,l,q_i} - s_{k,l,z_j}| \le \theta \\ 1, & \text{otherwise} \end{cases}$$

[0072]

[0073] 여기서, I는 룩업 테이블에서 좌표의 수, J는 그룹의 수이다.

[0074] $p(q_i \in z_j)$는 룩업 테이블 내에서 만일 i번째 좌표 $q_i$가 j번째 그룹에 속하면 1이고, 그렇지 않으면 0이다.

[0075] $e(q_i, z_j)$는 클러스터링 에러이고, 이는 모든 가능한 마이크로폰의 쌍 $d(s_{k,l,q}, s_{k,l,z})$의 합으로 정의된다.

[0076] $s_{k,l,q}$는 위치 q로부터 마이크로폰 쌍 k 와 l까지의 TDOA(time difference of arrival)이다. 만일 $s_{k,l,q}$와 $s_{k,l,z}$의 차이가 미리 설정된 임계치 $\Theta$보다 작다면, $d(s_{k,l,q}, s_{k,l,z})$는 0이며, 그렇지 않은 경우는 1로 가정된다.

[0077] J의 초기값은 1이고, 수학식 6을 만족할 때까지 1씩 증가한다. J가 증가함에 따라 모든 i 및 j에 대해 $p(q_i \in z_j)$, $e(q_i, z_j)$, $z_j$는 k-means 알고리즘에서와 유사하게 갱신된다.

[0078] $2\pi/5$ 미스얼라인먼트 제한을 충족시키기 위해, 임계값은 수학식 7과 같이 정의된다.

*수학식 7*

$$\theta = \text{round}\left(\frac{1}{5} \times \frac{\lambda}{c} \times r\right)$$

[0079]

[0080] 여기서, $\lambda$는 파장, c는 음속, r은 샘플링 율이다.

[0081] 임계값은 소리의 최대 주파수에 의존하며, 파장은 인간 음성의 경우 5.5cm고, 샘플링 율는 SSL을 위해 6kHz가 사용된다.

[0082] 미리 계산된 룩업 테이블 및 클러스터링된 그룹의 대표 클러스터들은 2 단계 탐색에 이용된다.

[0083] 1단계에서 상위 n개의 최대 출력 파워 그룹이 선택되고, 2단계에서 상위 n개의 그룹에 대한 상세 탐색이 수행된다.

[0084] 도 4는 본 발명의 일 실시 예에 따른 음원 방향 탐색 과정을 도시한 순서도이다.

[0085] 도 4를 설명함에 있어 도 5와 같이 그룹으로 분할한 경우를 예로 들어 설명한다.

[0086] 도 4 내지 도 5를 참조하면, 상기한 수학식 3을 이용하여(SRP-PHAT 알고리즘을 이용하여) 각 그룹(Group 1 내지 Group 5)의 대표 클러스터(500-1 내지 500-5)에서의 출력 파워를 계산할 수 있다(단계 400).

[0087] 대표 클러스터에 대한 출력 파워 계산을 통해, 출력 파워가 큰 상위 n개(여기서 n은 1이상의 자연수이다)의 그룹을 탐색한다(단계 402).

[0088] 도 5에서는 단계 402를 통해 Group 3과 5가 선택된 경우를 나타낸 것이다.

[0089] 이후, 수학식 3을 이용하여 상위 n개의 그룹에 속하는 모든 좌표들(510)의 출력 파워를 계산한다(단계 404).

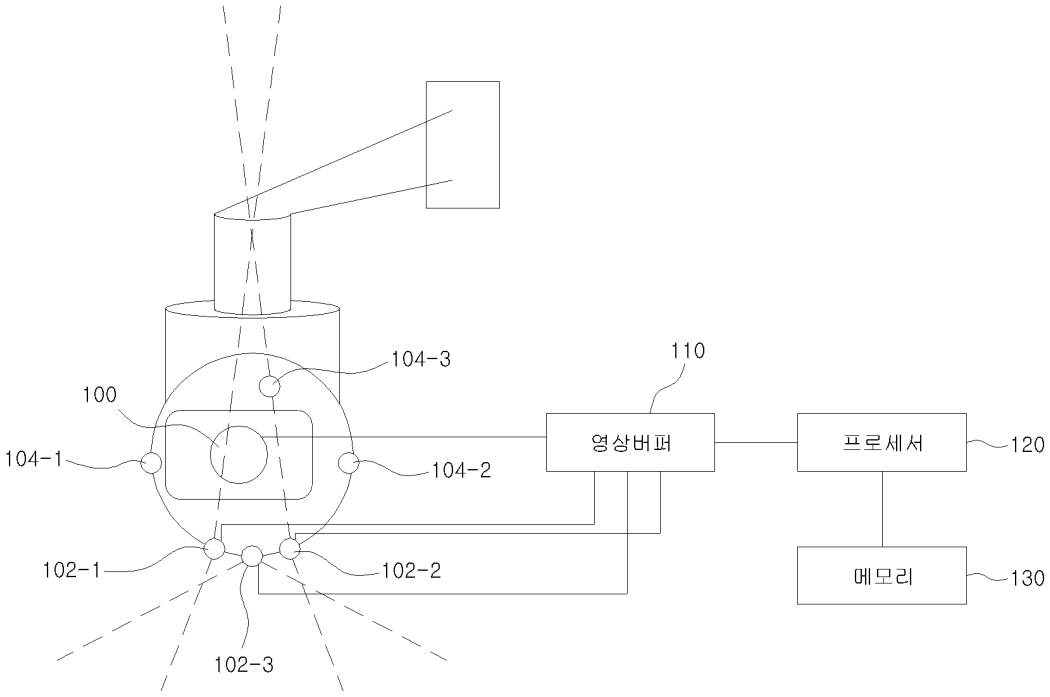
[0090] 마지막으로, 상위 n개의 그룹에 속하는 좌표들 중 최대 출력 파워를 갖는 좌표의 방향을 탐색한다(단계 406).

[0091] 단계 406을 통해 복수의 좌표들(510) 중 하나가 최대 출력 파워로 결정되고, 그 좌표가 있는 방향이 음원 방향으로 결정될 수 있다.

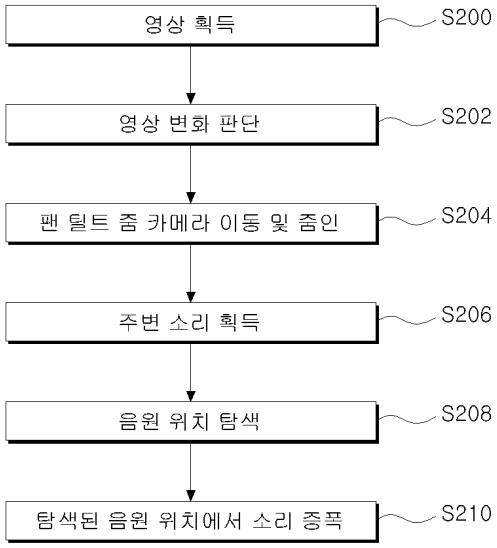
[0092] 이상과 같이 본 발명에서는 구체적인 구성 요소 등과 같은 특정 사항들과 한정된 실시 예 및 도면에 의해 설명되었으나 이는 본 발명의 보다 전반적인 이해를 돕기 위해서 제공된 것일 뿐, 본 발명은 상기의 실시 예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서 본 발명의 사상은 설명된 실시 예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등하거나 등가적 변형이 있는 모든 것들은 본 발명 사상의 범주에 속한다고 할 것이다.

## 도면

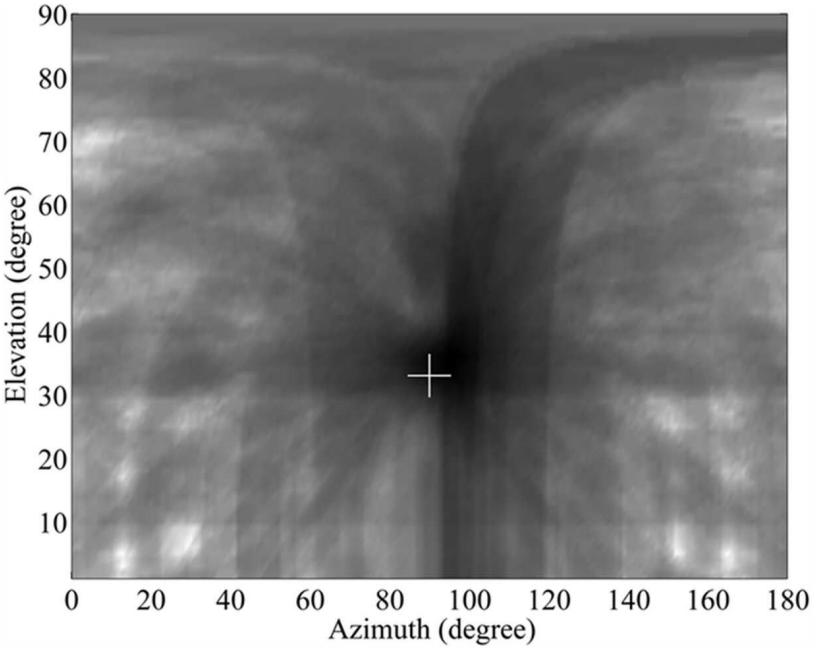
### 도면1


104-3
110
100
영상버퍼
프로세서
120
104-1
104-2
102-1
102-2
메모리
130
102-3

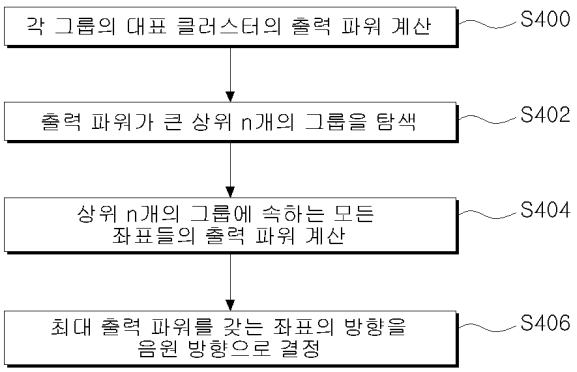
### 도면2


영상 획득
S200
영상 변화 판단
S202
팬 틸트 줌 카메라 이동 및 줌인
S204
주변 소리 획득
S206
음원 위치 탐색
S208
탐색된 음원 위치에서 소리 증폭
S210

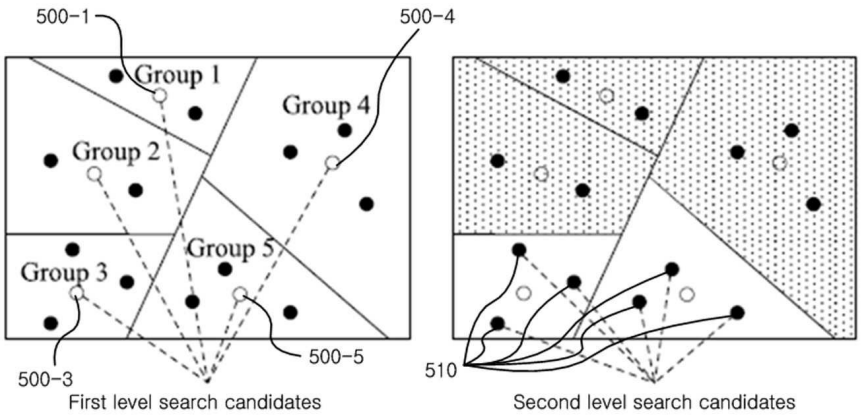
도면3




도면4




도면5




【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 6

【변경전】

전역 최대 파원 좌표

【변경후】

전역 최대 파워 좌표