(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0137491
(43) 공개일자 2015년12월09일

(51) 국제특허분류(Int. Cl.)
G01S 11/14 (2006.01) *G06T 7/00* (2006.01)
(21) 출원번호 10-2014-0065446
(22) 출원일자 2014년05월29일
심사청구일자 2014년05월29일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
정연모
경기도 성남시 분당구 미금로 251, 708동 1201호 (금곡동)
조재식
서울특별시 동대문구 겸재로 24, 705호 (휘경동, 가든아파트)
(74) 대리인
특허법인 신지

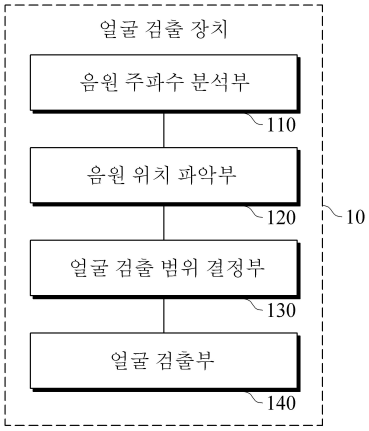
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 **다중 음원 국지화 기법을 이용한 얼굴 검출 장치 및 방법**

(57) 요 약

다중 음원 환경에서 얼굴을 검출하는 얼굴 검출 장치 및 방법이 개시된다. 일 실시예에 따른 얼굴 검출 장치는 음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 음원 주파수 분석부; 상기 음원 주파수 분석부에 의하여 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 상기 다중 음원의 위치를 파악하는 음원 위치 파악부; 및 상기 음원 위치 파악부에 의하여 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 상기 다중 음원에 대한 얼굴 검출을 수행하는 얼굴 검출부를 포함한다.

대 표 도 - 도1



이 발명을 지원한 국가연구개발사업
과제고유번호 C0142628
부처명 경기지방 중소기업청
연구관리전문기관 한국산학연협회
연구사업명 산학연협력기술개발
연구과제명 전장로봇에 적용하기 위한 음원국지화와 얼굴검출의 최적화 기술 개발
기 여 율 1/1
주관기관 경희대학교 산학협력단
연구기간 2013.11.01 ~ 2014.10.31

명 세 서

청구범위

청구항 1

음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 음원 주파수 분석부;

상기 음원 주파수 분석부에 의하여 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 상기 다중 음원의 위치를 파악하는 음원 위치 파악부; 및

상기 음원 위치 파악부에 의하여 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 상기 다중 음원에 대한 얼굴 검출을 수행하는 얼굴 검출부

를 포함하는 얼굴 검출 장치.

청구항 2

제1항에 있어서,

상기 음원 위치 파악부에 의하여 파악된 다중 음원의 위치에 기초하여 상기 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 얼굴 검출 범위 결정부를 더 포함하고,

상기 얼굴 검출부는 상기 얼굴 검출 범위 결정부에 의하여 결정된 얼굴 검출 범위에서 얼굴 검출을 수행하는 것을 특징으로 하는 얼굴 검출 장치.

청구항 3

제1항에 있어서,

상기 음원 획득 장치는 전방 마이크로폰, 후방 마이크로폰, 좌측 마이크로폰 및 우측 마이크로폰이 십자 형태로 배열된 것을 특징으로 하는 얼굴 검출 장치.

청구항 4

제1항에 있어서,

상기 음원 주파수 분석부는 상기 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고 상기 디지털 신호를 주파수 영역의 신호로 변환하여 상기 다중 음원의 주파수를 분석하는 것을 특징으로 하는 얼굴 검출 장치.

청구항 5

제4항에 있어서,

상기 음원 위치 파악부는 상기 주파수 영역에서 주파수 대역별 시간차를 산출하고, 상기 산출된 주파수 대역별 시간차로부터 산출된 최종 시간차를 이용하여 상기 다중 음원의 위치를 파악하는 것을 특징으로 하는 얼굴 검출 장치.

청구항 6

제1항에 있어서,

상기 얼굴 검출부는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행하는 것을 특징으로 하는 얼굴 검출 장치.

청구항 7

음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 단계;

상기 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 상기 다중 음원의 위치를 파악하는

단계; 및

상기 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 상기 다중 음원에 대한 얼굴 검출을 수행하는 단계

를 포함하는 얼굴 검출 방법.

청구항 8

제7항에 있어서,

상기 파악된 다중 음원의 위치에 기초하여 상기 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 단계를 더 포함하고,

상기 얼굴 검출 수행 단계에서는 상기 얼굴 검출 범위 결정 단계에서 결정된 얼굴 검출 범위에서 얼굴 검출을 수행하는 것을 특징으로 하는 얼굴 검출 방법.

청구항 9

제7항에 있어서,

상기 음원 획득 장치는 전방 마이크로폰, 후방 마이크로폰, 좌측 마이크로폰 및 우측 마이크로폰이 십자 형태로 배열된 것을 특징으로 하는 얼굴 검출 방법.

청구항 10

제7항에 있어서,

상기 음원 주파수 분석 단계에서는 상기 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고 상기 디지털 신호를 주파수 영역의 신호로 변환하여 상기 다중 음원의 주파수를 분석하고,

상기 음원 위치 파악 단계에서는 상기 주파수 영역에서 주파수 대역별 시간차를 산출하고, 상기 산출된 주파수 대역별 시간차로부터 산출된 최종 시간차를 이용하여 상기 다중 음원의 위치를 파악하는 것을 특징으로 하는 얼굴 검출 방법.

청구항 11

제7항에 있어서,

상기 얼굴 검출 수행 단계에서는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행하는 것을 특징으로 하는 얼굴 검출 방법.

청구항 12

제7항 내지 제11항 중 어느 한 항에 따른 방법을 컴퓨터로 실행하기 위한 프로그램이 기록된 컴퓨터 판독가능 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 얼굴 검출 기술에 관한 것으로서, 보다 구체적으로 다중 음원 환경에서 얼굴을 검출하는 얼굴 검출 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 얼굴 검출 기술은 최근 IPTV, 스마트폰, 디지털 카메라, 출입통제장치 등의 다양한 디지털 기기 및 휴먼 컴퓨터 인터페이스, 비디오 감시 시스템, 로봇의 시각 시스템, 출입통제 시스템, 모바일 콘텐츠 등의 다양한 응용분야가 생겨나면서 그 중요성이 점점 커지고 있다. 얼굴 인식을 위해서는 촬영된 디지털 영상에서 먼저 얼굴 영역을 검출해야 하며, 얼굴 영역의 검출 속도는 전체 시스템 성능을 좌우하는 중요한 요소이다.

[0003] 얼굴 검출 기술에 대한 많은 연구가 이루어지고 있으나, 실제 생활에 적용되기에는 아직 알고리즘의 처리속도,

신뢰성 등이 만족스럽지 못하다. 종래 Paul Viola와 Michael J. Jones에 의하여 개발된 얼굴 검출 방법은 얼굴 특성을 기반으로 프레임들을 다양한 크기로 리사이즈하여, 리사이즈된 프레임들의 모든 이미지에 대하여 얼굴을 검출한다. 이는 소프트웨어를 기반으로 순차적으로 수행되므로, 다양한 크기의 얼굴 검출을 위하여 각 영상을 순차적으로 리사이즈하여, 각 리사이즈된 영상 각각에 대하여 얼굴 검출을 수행해야 하기 때문에, 계산량이 많아 얼굴 검출을 하는데 시간이 오래 걸리는 문제가 있다.

[0004] 즉, 기존의 얼굴 검출 방법들은 얼굴 검출을 위하여 방대한 계산량을 필요로 하기 때문에, 최근에는 입력된 이미지의 크기를 줄이거나 얼굴이 있을 후보 영역을 전처리 단계에서 빠르게 계산하여 탐색 범위를 좁히는 방법 등이 개발된 바 있다.

[0005] 그러나, 이러한 얼굴 검출 방법은 얼굴 검출 정확도의 손해를 감수해야 하거나 속도 향상이 월등하지 않다. 구체적으로 말하면, 현재까지 개발된 얼굴 검출 기법들은 하드웨어 또는 소프트웨어 방식과 무관하게, VGA급의 낮은 해상도에서 실시간으로 동작을 하거나, FHD급의 높은 해상도에서는 매우 낮은 프레임으로 동작을 하고 있다.

[0006] 실시간 보안 분야에서는 얼굴 검출의 범위와 정확도를 높이기 위하여 높은 해상도가 필수적이지만, 높은 해상도로 검출하게 되면 실시간으로 검출하지 못하는 문제점이 있고, 실시간 얼굴 검출을 위하여 해상도를 낮추면 얼굴 검출의 범위와 정확도가 낮아지는 문제점이 있다.

[0007] 한편, 음원 국지화 기술은 센서로부터 음원이 위치한 장소를 찾아내는 기술을 의미한다. 소리에 의한 방향 감각은 양쪽 귀에 음파가 도달하는 차이에 의하여 판별되는 것으로 사람의 경우에 16방향까지 인지가 가능하다. 오래전부터 이러한 음원의 방향감지에 대하여 연구가 진행되었고, 음원의 방향 감지 기술은 음원 국지화 기술의 바탕이 되었다. 음원 국지화 기술은 레이더, SONAR 시스템, 인간과 로봇간의 상호작용, 스마트 감시 카메라 등 다양한 분야에서 적용 및 연구되고 있다. 음원 국지화를 위한 음원 방향 지각 방식에는 마이크로폰 어레이를 이용하는 방법 등이 있다.

발명의 내용

해결하려는 과제

[0008] 본 발명이 해결하고자 하는 하나의 과제는 다중 음원 환경에서 얼굴을 검출하는 얼굴 검출 장치 및 방법을 제공하는 것이다.

[0009] 본 발명이 해결하고자 하는 다른 하나의 과제는 보다 신속하게 얼굴 검출이 가능한 얼굴 검출 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0010] 상기한 과제를 해결하기 위한 본 발명의 일 실시예에 따른 얼굴 검출 장치는 음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 음원 주파수 분석부; 상기 음원 주파수 분석부에 의하여 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 상기 다중 음원의 위치를 파악하는 음원 위치 파악부; 및 상기 음원 위치 파악부에 의하여 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 상기 다중 음원에 대한 얼굴 검출을 수행하는 얼굴 검출부를 포함한다.

[0011] 상기 실시예의 일 측면에 의하면, 상기 음원 위치 파악부에 의하여 파악된 다중 음원의 위치에 기초하여 상기 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 얼굴 검출 범위 결정부를 더 포함하고, 상기 얼굴 검출부는 상기 얼굴 검출 범위 결정부에 의하여 결정된 얼굴 검출 범위에서 얼굴 검출을 수행할 수 있다.

[0012] 상기 실시예의 다른 측면에 의하면, 상기 음원 획득 장치는 전방 마이크로폰, 후방 마이크로폰, 좌측 마이크로폰 및 우측 마이크로폰이 십자 형태로 배열될 수 있다.

[0013] 상기 실시예의 또 다른 측면에 의하면, 상기 음원 주파수 분석부는 상기 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고 상기 디지털 신호를 주파수 영역의 신호로 변환하여 상기 다중 음원의 주파수를 분석할 수 있다. 이 경우에, 상기 음원 위치 파악부는 상기 주파수 영역에서 주파수 대역별 시간차를 산출하고, 상기 산출된 주파수 대역별 시간차로부터 산출된 최종 시간차를 이용하여 상기 다중 음원의 위치를 파악할 수 있다.

[0014] 상기 실시예의 또 다른 측면에 의하면, 상기 얼굴 검출부는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local

Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행할 수 있다.

[0015] 상기한 과제를 해결하기 위한 본 발명의 일 실시예에 따른 얼굴 검출 방법은 음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 단계; 상기 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 상기 다중 음원의 위치를 파악하는 단계; 및 상기 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 상기 다중 음원에 대한 얼굴 검출을 수행하는 단계를 포함한다.

[0016] 상기 실시예의 일 측면에 의하면, 상기 파악된 다중 음원의 위치에 기초하여 상기 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 단계를 더 포함하고, 상기 얼굴 검출 수행 단계에서는 상기 얼굴 검출 범위 결정 단계에서 결정된 얼굴 검출 범위에서 얼굴 검출을 수행할 수 있다.

[0017] 상기 실시예의 다른 측면에 의하면, 상기 음원 획득 장치는 전방 마이크로폰, 후방 마이크로폰, 좌측 마이크로폰 및 우측 마이크로폰이 십자 형태로 배열될 수 있다.

[0018] 상기 실시예의 또 다른 측면에 의하면, 상기 음원 주파수 분석 단계에서는 상기 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고 상기 디지털 신호를 주파수 영역의 신호로 변환하여 상기 다중 음원의 주파수를 분석하고, 상기 음원 위치 파악 단계에서는 상기 주파수 영역에서 주파수 대역별 시간차를 산출하고, 상기 산출된 주파수 대역별 시간차로부터 산출된 최종 시간차를 이용하여 상기 다중 음원의 위치를 파악할 수 있다.

[0019] 상기 실시예의 또 다른 측면에 의하면, 상기 얼굴 검출 수행 단계에서는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행할 수 있다.

[0020] 상기한 과제를 해결하기 위한 본 발명의 일 실시예에 따른 컴퓨터 판독가능한 기록매체는 본 발명의 일 실시예에 따른 얼굴 검출 방법을 컴퓨터로 실행하기 위한 프로그램이 기록된다.

발명의 효과

[0021] 이상과 같이 본 발명의 일 실시 예에 따르면, 다중 음원 환경에서 얼굴을 검출함으로써 영상의 이미지 중심이 아닌 음원 중심으로 얼굴을 검출할 수 있는 효과가 있다.

[0022] 그리고 본 발명의 일 실시 예에 따르면, 다중 음원 환경에서 음원을 중심으로 얼굴 검출시 다중 음원 위치의 밀도가 가장 큰 영역을 선정하고 얼굴 검출 범위를 축소하여 얼굴 검출을 수행함으로써 얼굴 검출의 성능을 향상시킬 수 있는 효과가 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 일 실시예에 따른 얼굴 검출 장치의 구성을 보여주는 블록도이다.

도 2 및 도 3은 본 발명의 일실시예에 따른 얼굴 검출 범위 결정을 보여주는 도면으로, 도 2는 4개의 마이크로폰의 배치와 음원 검출 영역 분할을 보여주는 도면이고, 도 3은 음원 위치의 밀도가 가장 큰 범위를 선정하여 결정된 얼굴 검출 범위를 보여주는 도면이다.

도 4는 본 발명의 일 실시예에 따른 얼굴 검출 방법을 보여주는 흐름도이다.

도 5는 본 발명의 다른 실시예에 따른 LBP 알고리즘을 이용한 얼굴 검출 방법을 보여주는 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0024] 이하 첨부된 도면을 참조하여 본 발명을 보다 상세히 설명한다. 그러나 이러한 도면은 본 발명의 기술적 사상의 내용과 범위를 쉽게 설명하기 위한 예시일 뿐, 이에 의해 본 발명의 기술적 범위가 한정되거나 변경되는 것은 아니다. 또한 이러한 예시에 기초하여 본 발명의 기술적 사상의 범위 안에서 다양한 변형과 변경이 가능함은 통상의 기술자에게는 당연할 것이다.

[0025] 또한, 본 명세서에서 사용되는 용어 및 단어들은 실시예에서의 기능을 고려하여 선택된 용어들로서, 그 용어의 의미는 발명의 의도 또는 관례 등에 따라 달라질 수 있다. 따라서 후술하는 실시예에서 사용된 용어는, 본 명세서에 구체적으로 정의된 경우에는 그 정의에 따르며, 구체적인 정의가 없는 경우는 통상의 기술자들이 일반적으로 인식하는 의미로 해석되어야 할 것이다.

[0026] '얼굴 검출(Face Detection)' 은 컴퓨터 비전의 한 분야로 영상(Image)에서 얼굴이 존재하는 위치를 알려주는 기술이다. 얼굴 검출의 알고리즘적인 기본 구조는 Rowley, Baluja, Kanade의 논문에 의해 정의되었다. 일반적으로 다양한 크기의 얼굴을 검출하기 위하여 피라미드 영상을 생성한 후, 한 픽셀씩 이동하며 특정 크기(예를 들어, 20 X 20 픽셀)의 해당 영역이 얼굴인지 아닌지를 분류기(예를 들어, 신경망(Neural Network), 아다부스트(Adaboost), 서포트 벡터 머신(Support Vector Machine) 등)로 얼굴인지 아닌지를 결정한다.

[0027] 초기 얼굴 검출에 사용된 특징은 영상에서 얼굴의 강도(intensity)였다. 하지만 인종, 조명 등에 따라 성능이 달라짐에 따라 이에 무관한 특징이 필요하게 되었다. 이에 하-라이크 특징(Haar-like feature), 국부 이진 패턴(Local Binary Pattern, LBP), Modified Census Transform(MCT) 등의 특징이 등장하였다.

[0028] 제안하는 본 발명의 얼굴 검출 장치 및 방법은 얼굴 검출시 입력된 영상 전체를 대상으로 얼굴 검출을 수행하지 않고, 다중 음원을 기초로 하여 얼굴이 있을 수 있는 일부 영역을 먼저 검출한 다음, 해당 영역을 중심으로 얼굴 검출을 수행함으로써, 보다 신속하게 얼굴 검출이 가능하다. 특히, 일반적인 경우보다는 영상에서 얼굴들이 특정부분에 모여있고 음성으로 파악이 가능한 경우 얼굴 검출을 효율적으로 할 수 있다. 이하에서, 본 발명의 실시예에 따른 얼굴 검출 장치 및 방법에 대하여 하기 도면들을 참조하여 자세히 살펴본다.

[0029] 먼저, 도 1은 본 발명의 일 실시예에 따른 얼굴 검출 장치의 구성을 보여주는 블록도이다.

[0030] 도 1을 참조하면, 얼굴 검출 장치(10)는 음원 주파수 분석부(110), 음원 위치 파악부(120), 얼굴 검출 범위 결정부(130) 및 얼굴 검출부(140)를 포함한다. 그리고, 도 1에 도시되어 있는 얼굴 검출 장치(10)의 구성은 예시적인 것으로서, 얼굴 검출 장치(10)는 도 1에 개시되어 있는 모듈들의 일부만을 구비하거나 및/또는 그 동작을 위하여 필수적인 다른 모듈들을 추가로 구비할 수도 있다. 예를 들어, 얼굴 검출 장치(10)는 다른 장치와 통신을 위하여 통신 기능을 가지는 통신부 등을 추가로 구비할 수 있다.

[0031] 음원 주파수 분석부(110)는 음원 획득 장치로부터 획득된 다중 음원의 주파수를 분석하는 기능을 제공한다. 여기서, '음원 획득 장치' 는 가청음을 전기적인 에너지 변환기나 센서로 전달하여 소리를 전기 신호로 변환해주는 장치로서, 마이크로폰(Microphone) 또는 마이크로폰 어레이(Microhphone Array)일 수 있으나, 여기에만 한정되는 것은 아니며, 음원을 획득할 수 있는 장치라면 해당될 수 있다. 통상적으로 마이크로폰 어레이는 다수의 마이크로폰들을 조합하여 사운드 자체뿐만 아니라 취득하려는 사운드의 방향이나 위치와 같은 지향성에 관한 부가적인 성질을 얻을 수 있다. 여기서, '지향성(Directivity)' 이라 함은 소리, 전파, 빛 등의 송수신에 있어, 그것들의 장치 특성이 특정 방향에서 강한 성질을 나타내는 현상을 의미하는 것으로, 음원 신호가 어레이를 구성하는 다수의 마이크로폰들 각각에 도달하는 시간 차이를 이용하여 특정 방향에 위치한 음원으로부터 방사되는 음원 신호에 대한 감도를 크게 하는 것을 말한다. 따라서, 이러한 마이크로폰 어레이를 이용하여 음원 신호들을 취득함으로써 특정 방향으로부터 입력되는 음원 신호를 강조하거나 억제할 수도 있다.

[0032] 또한, 본 발명의 실시예에서 음원 획득 장치는 동영상 획득 대상인 영역을 복수의 음원 검출 영역으로 각각 나누고, 각각의 영역을 인식할 수 있도록 마이크로폰 또는 마이크로폰 어레이를 설치할 수 있다. 예를 들어, 4개의 마이크로폰을 서로 대칭되게 십자 형상으로 전방, 후방, 좌측 및 우측에 배치함으로써, 각각의 영역을 4개의 마이크로폰을 이용하여 인식할 수 있도록 구현할 수 있다. 본 발명의 실시예에서는 4개의 마이크로폰을 음원 획득 장치로 이용하고, 음원 검출 영역을 16 영역으로 분할하였다. 음원 획득 장치의 배치와 복수의 얼굴 검출 영역 분할에 대하여는 뒤에서 도 2 및 도 3을 참고하여 상세히 설명한다.

[0033] 한편, 음원 주파수 분석부(110)는 획득된 다중 음원 신호를 분석하여 다중 음원의 주파수를 분석할 수 있다. 또는, 주파수 영역의 신호에 기초하여 음원 위치를 파악하기 위하여, 음원 주파수 분석부(110)는 획득된 다중 음원 신호를 디지털 신호로 변환하고 변환된 디지털 신호를 주파수 영역의 신호로 변환하여 다중 음원의 주파수를 분석할 수 있다.

[0034] 음원 위치 파악부(120)는 음원 국지화(Sound Source Localization)를 실행하여 다중 음원의 위치를 파악하는 기능을 제공한다. 이 경우에, 음원 국지화는 음원 주파수 분석부(110)에 의하여 분석된 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하면 실행될 수 있다. 사람이 들을 수 있는 가청 주파수 영역은 평균적으로 20~20,000Hz인 반면에, 사람이 낼 수 있는 주파수 영역은 평균적으로 300~3,400Hz의 범위를 가진다. 또한, 평균적으로 남자에 비해서 여자의 발성 주파수가 더 높다. 일반적으로 남성의 목소리 주파수는 100~150Hz, 여성은 200~250Hz로 나타난다. 여기서, '음원 국지화' 라 함은 음원 감지 센서로부터 음원이 위치하는 장소를 찾아내는 기술을 의미한다. 음원 국지화 기술은 군사용 무기나 로봇 등 다양한 분야에서 활용되고 있다. 이러한 음원

국지화 기술은 다수의 마이크로폰들을 이용하여 만들어진 마이크로폰 어레이를 사용하여 이루어진다. 마이크로폰 어레이를 사용하여 음원의 방향을 추정하기 위하여 음성 신호의 도달 지연 시간을 측정할 필요가 있고, 음성 신호의 도달 지연 시간을 측정하기 위하여 GCC(Generalized Cross Correlation) 알고리즘이 일반적으로 사용된다.

[0035] 여기서, '음원의 위치'라 함은 기준점(마이크로폰 또는 마이크로폰 어레이가 될 수 있다.)을 중심으로 음원이 위치한 방향 또는 거리를 의미한다. 즉, 마이크로폰 또는 마이크로폰 어레이를 구성하는 개별 마이크로폰들을 각각 다르게 지연시키면 특정 방향에 위치한 음원 신호에 대하여 지향성을 가지게 되고, 이러한 과정을 전 방향에 대하여 수행한다. 만약, 특정 방향으로부터 획득된 음원 신호의 음압이 최대값을 갖는다면 해당 방향에 음원이 존재한다고 파악할 수 있다.

[0036] 또는, 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고, 변환된 디지털 신호를 다시 주파수 영역의 신호로 변환하여, 다중 음원의 주파수 영역에서 주파수 대역별 시간차를 산출하고, 산출된 주파수 대역별 시간차로부터 최종 시간차를 산출하여, 최종 시간차를 이용하여 다중 음원의 위치를 파악할 수 있다.

[0037] 얼굴 검출 범위 결정부(130)는 동영상 획득 대상인 영역에 대하여 복수의 음원 검출 영역으로 분할하고, 음원 위치 파악부(120)에 의하여 파악된 다중 음원의 위치에 기초하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 하는 기능을 제공한다. 여기서, '얼굴 검출을 수행할 범위를 결정'한다는 것은 동영상 내에서 얼굴 검출을 수행할 때 전체 영상에 대하여 얼굴 검출을 수행하거나 또는 이미지 상에서 얼굴 가능 영역을 중심으로 얼굴 검출을 수행하는 것이 아니라, 다중 음원 위치의 밀도를 중심으로 하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 것을 의미한다.

[0038] 그리고, '다중 음원 위치의 밀도가 가장 큰 범위'라 함은 동영상 획득 대상인 영역에 대하여 복수의 음원 검출 영역으로 분할한 후 음원 국지화를 수행한 결과, 복수의 음원 검출 영역 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 의미한다. 또는, 동영상 획득 대상인 영역에 대하여 음원 국지화를 수행한 후 복수의 음원 검출 영역으로 분할한 결과, 복수의 음원 검출 영역 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 의미할 수도 있다.

[0039] 또한, '복수의 음원 검출 영역으로 분할'한다는 것은 동영상 획득 대상인 영역에 대하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하기 위하여, 동영상 획득 대상인 영역을 소정의 기준에 따라 복수 개의 영역으로 분할하는 것을 의미한다. 여기서, 동영상 획득 대상인 영역을 복수의 음원 검출 영역으로 분할하는 방법은 동영상 획득 대상인 영역에 배치된 음원 획득 장치의 개수에 따라 달라질 수 있으나, 여기에만 한정되는 것은 아니다. 예를 들어, 동영상 획득 대상인 영역의 범위가 넓은 경우라면 음원 검출 영역을 더욱 세분화할 수 있다. 본 발명의 실시예에서는 4개의 마이크로폰을 음원 획득 장치로 이용하고, 음원 검출 영역을 16 영역으로 분할하였다. 그러나, 음원 검출 영역은 16 영역이 아니라 9 영역 또는 25 영역 또는 36 영역 등으로 분할할 수도 있다.

[0040] 한편, 얼굴 검출 범위를 결정하기 위하여 소정의 시간 동안 음원 획득 장치로부터 다중 음원을 획득하여 음원 주파수 분석 및 음원 위치를 파악할 수 있다. 소정의 시간 동안 다중 음원을 획득하여 음원 주파수 분석 및 음원 위치를 파악하는 것은 동영상 획득 대상인 영역에서 사람의 목소리가 가장 많이 파악되는 영역, 즉 사람이 가장 많은 곳을 찾기 위함이다.

[0041] 얼굴 검출부(140)는 음원 위치 파악부(120)에 의하여 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 다중 음원에 대한 얼굴 검출을 수행하는 기능을 제공한다. 또는, 얼굴 검출 범위 결정부(130)에 의하여 결정된 얼굴 검출 범위에서 얼굴 검출을 수행할 수 있다. 기존의 얼굴 검출 기술에는 스킨 컬러 기반 기법(Skin Color based approach), 서포트 벡터 머신 기법(Support Vector Machine approach: SVM), 가우시안 혼합 기법(Gaussian Mixture approach), 최대 유사 기법(Maximum likelihood approach), 신경망 기법(Neural Network approach) 등이 있다. 이러한 기술들을 하드웨어로 구현하기 위해서는, 기본적으로 얼굴 패턴에 대한 정보와 비 얼굴 패턴에 대한 정보가 등록된 데이터베이스의 구출과 얼굴의 특징에 대한 코스트 값(Cost Value)이 저장된 룩-업 테이블(Look-up Table)이 요구된다. 여기서, '코스트 값'이란 내부적으로 수집된 통계 정보에 기초하여 계산된 예측값이다.

[0042] 한편, 본 발명의 실시예에 의하면, 얼굴 검출 방법에는 특별한 제한이 없는데, 얼굴 검출을 위하여 얼굴 검출부(140)는 종래 주지된 얼굴 검출 인식 알고리즘 등을 이용할 수 있다. 예를 들어, 얼굴 검출에서 다양한 얼굴 크기를 검출하기 위해서는 피라미드 방식을 이용하여 입력영상의 크기를 점점 변경하고 고정된 크기의 얼굴 검출기로 검출을 수행하는 방법 또는 입력영상의 크기는 고정하고 얼굴 검출기의 특징추출 필터의 크기를 점점 변경

하여 검출하는 방법 등을 이용할 수 있다. 이때, 일반적으로 매번 얼굴 검출기의 필터 크기를 변경하는 것보다 입력영상을 피라미드 방식으로 변경함으로써 보다 높은 검출 속도를 얻을 수 있다. 여기서, 얼굴 검출을 수행하는데 이용될 수 있는 하나의 방법은 이미지를 스캐닝하는 슬라이딩 윈도우를 이용하는 것이다. 슬라이딩 윈도우는 얼굴 샘플과 동일한 크기일 수 있는 20 X 20과 같은 미리 정의된 크기를 가질 수 있다. 얼굴 검출을 수행하기 위하여 이용되는 스캐닝 및 매칭 처리는 최대 얼굴 크기가 도달될 때까지 입력 이미지를 다운 샘플링함으로써 수회 반복될 수 있다.

[0043] 또한, 얼굴 검출 동작은 예를 들면, 아다부스트 학습 알고리즘에 의해서 학습되어 선택된 특정 패턴들을 이용하여 수행될 수 있지만, 반드시 여기에만 한정되는 것은 아니며 다양한 변형 예를 상정할 수 있다. 이때, 얼굴 검출부(140)는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행할 수 있다. LBP(Local Binary Pattern) 알고리즘은 3 X 3 마스크내의 중심 픽셀에 대해 주변 픽셀과의 크기 비교를 통해 선택된 이진 패턴을 의미한다. 주변의 8개 픽셀에 대해 일정한 방향이 정해지면 각 픽셀이 중심보다 크면 1, 작거나 같으면 0을 할당하여 8비트의 값이 결정되고, 이를 중심픽셀의 값으로 할당한다. LBP는 8개의 인접한 픽셀들과 비교하여 구성되기 때문에 8비트로 표현되며, 0부터 255 사이의 값을 갖는다.

[0044] 이러한 방식을 통하여 영상의 해상도를 축소시키지 않은 채 얼굴 검출 범위를 축소시켰기 때문에 사람의 얼굴 크기는 그대로 이므로 얼굴 검출률을 유지할 수 있고, 얼굴 검출 범위가 축소되었기 때문에 실시간 얼굴 검출이 가능하다.

[0045] 도 2 및 도 3은 본 발명의 일실시예에 따른 얼굴 검출 범위 결정을 보여주는 도면으로, 도 2는 4개의 마이크로폰의 배치와 음원 검출 영역 분할을 보여주는 도면이고, 도 3은 음원 위치의 밀도가 가장 큰 범위를 선정하여 결정된 얼굴 검출 범위를 보여주는 도면이다.

[0046] 도 2를 참조하면, 4개의 마이크로폰을 음원 획득 장치(21, 22, 23, 24)로 이용하고, 동영상 획득 대상인 영역에 대하여 음원 검출 영역(200)을 16 영역으로 분할하였다. 구체적으로, 4개의 마이크로폰을 서로 대칭되게 십자 형상으로 전방, 후방, 좌측 및 우측에 배치함으로써, 각각의 영역을 4개의 마이크로폰을 이용하여 인식할 수 있도록 구현할 수 있다. 여기서, 2차원 영역에서 음원 검출을 하는 경우 마이크로폰이 전방, 후방, 좌측 및 우측에 배치되도록 해야 음원의 방향성을 정확히 알 수 있다. 만일, 3차원 영역에서 음원 검출을 하는 경우라면 전방, 후방, 좌측 및 우측에 추가하여 상방 및 하방에 마이크로폰이 배치되도록 해야 음원의 방향성을 정확히 알 수 있다. 만일, 마이크로폰을 전방 및 후방에 2개만 배치한다면 음원의 전후 각도만 알 수 있고, 좌측 및 우측에 2개만 배치한다면 음원의 좌우 각도만 알 수 있다. 따라서, 2차원 영역에서 음원의 위치 좌표를 알기 위해서는 최소한 4방향에 마이크로폰을 배치해야 음원 국지화를 할 수 있게 된다. 또한, 도 2와 같이 마이크로폰을 동영상 획득 대상인 영역을 구성하는 변의 중심부가 아닌 꼭지점 부분에 배치할 수도 있으나, 배치 방법은 다양할 수 있고 여기에만 한정되는 것은 아니다.

[0047] 또한, 본 발명의 실시예에서는 음원 검출 영역(200)을 16 영역으로 분할하였으나, 분할되는 영역의 개수는 다양할 수 있고, 여기에만 한정되는 것은 아니다. 예를 들면, 음원 검출 영역(200)은 4 영역 또는 9 영역 또는 25 영역 또는 36 영역 등으로 분할할 수도 있다. 그러나, 음원 검출 영역(200)을 더 많은 영역으로 분할하면 더 정확하게 영상을 분할하여 처리함으로써 얼굴 검출 장치(10)의 효율성을 높일 수는 있으나, 구현상의 복잡성이 있을 수도 있다.

[0048] 한편, 얼굴 검출 성능을 보다 높이기 위해서는 각 분할 영역마다 음원 획득 장치, 즉 마이크로폰을 배치할 수도 있으나, 얼굴 검출 장치의 간소화와 효율성을 고려하면 도 2와 같이 마이크로폰을 전방, 후방, 좌측 및 우측에 배치하는 것이 바람직할 수 있다.

[0049] 도 3을 참조하면, 도 2에 따라 음원 획득 장치(21, 22, 23, 24)를 배치하고, 동영상 획득 대상인 영역에 대하여 복수의 음원 검출 영역(200)으로 분할하여, 다중 음원을 획득한 결과를 보여준다. 도 3에서 동영상 획득 대상 영역에서 점으로 표현된 부분은 다중 음원 각각의 위치를 나타낸다. 본 발명의 실시예에서는 동영상 획득 영역에 대하여 16 영역으로 음원 검출 영역(200)을 분할하여 다중 음원을 획득한 결과, 중심으로부터 위쪽의 4 영역의 음원 위치의 밀도가 가장 크다. 따라서 해당 4 영역이 얼굴 검출을 수행할 범위(300)에 해당하게 된다. 이때, 도 1에서 설명한 바와 같이, 동영상 획득 대상인 영역에 대하여 복수의 음원 검출 영역(200)으로 분할한

후 음원 국지화를 수행한 결과, 복수의 음원 검출 영역(200) 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 얼굴 검출 범위(300)로 결정할 수 있다. 또는, 동영상 획득 대상인 영역에 대하여 음원 국지화를 수행한 후 복수의 음원 검출 영역(200)으로 분할한 결과, 복수의 음원 검출 영역(200) 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 얼굴 검출 범위(300)로 결정할 수도 있다.

[0050] 도 4는 본 발명의 일 실시예에 따른 얼굴 검출 방법을 보여주는 흐름도이다. 본 발명의 실시예를 포함하여 이하에서 설명되는 절차는 다양한 형태로 구현하는 것이 가능하다. 도 4에 도시된 얼굴 검출 방법은 도 1의 얼굴 검출 장치(200) 또는 이를 구비하는 전자 기기를 이용하여 얼굴을 검출하는 방법일 수 있다. 따라서, 불필요한 반복을 피하기 위하여 얼굴 검출 방법에 관하여 간략히 설명하며, 여기에서 상세히 설명되지 않은 사항은 도 1 내지 도 3을 참조하여 설명한 사항이 동일하게 적용될 수 있다.

[0051] 도 1 내지 도 4를 참조하면, 먼저 음원 획득 장치로부터 다중 음원을 획득한다(S401). 여기서, '음원 획득 장치'는 가청음을 전기적인 에너지 변환기나 센서로 전달하여 소리를 전기 신호로 변환해주는 장치로서, 마이크로폰(Microphone) 또는 마이크로폰 어레이(Microhphone Array)일 수 있으나, 여기에만 한정되는 것은 아니며, 음원을 획득할 수 있는 장치라면 여기에 해당될 수 있다. 또한, 본 발명에서 음원 획득 장치는 동영상 획득 대상인 영역을 복수의 음원 검출 영역으로 각각 나누고, 각각의 영역을 인식할 수 있도록 마이크로폰 또는 마이크로폰 어레이를 설치할 수 있다. 예를 들어, 4개의 마이크로폰을 서로 대칭되게 십자 형상으로 전방, 후방, 좌측 및 우측에 배치함으로써, 각각의 영역을 4개의 마이크로폰을 이용하여 인식할 수 있도록 구현할 수 있다. 본 발명의 실시예에서는 4개의 마이크로폰을 음원 획득 장치로 이용하고, 음원 검출 영역을 16 영역으로 분할하였다.

[0052] 다음으로, 다중 음원의 주파수를 분석한다(S402). 이 경우에, 음원 획득 장치로부터 획득된 다중 음원 신호를 분석하여 다중 음원의 주파수를 분석할 수 있다. 또는, 주파수 영역의 신호에 기초하여 음원 위치를 파악하기 위하여, 음원 획득 장치로부터 획득된 다중 음원 신호를 디지털 신호로 변환하고 변환된 디지털 신호를 주파수 영역의 신호로 변환하여 다중 음원의 주파수를 분석할 수 있다.

[0053] 다음으로, 다중 음원 주파수 분석 결과 다중 음원의 주파수가 사람의 목소리 주파수 범위에 해당하는지 여부를 판단한다(S403). 판단 결과, 단계 S402에서 분석된 주파수가 사람의 목소리 주파수 범위에 해당하지 않는 것으로 판단되면, 다시 다중 음원을 획득한다(S401).

[0054] 반면에, 단계 S402에서 분석된 주파수가 사람의 목소리 주파수 범위에 해당하는 것으로 판단되면, 다중 음원의 위치를 파악하기 위하여 음원 국지화를 실행한다(S404). 여기서, '음원 국지화'라 함은 음원 감지 센서로부터 음원이 위치하는 장소를 찾아내는 기술을 의미한다. 이러한 음원 국지화 기술은 다수의 마이크로폰들을 이용하여 만들어진 마이크로폰 어레이를 사용한다. 마이크로폰 어레이를 사용하여 음원의 방향을 추정하기 위하여 음성 신호의 도달 지연 시간을 측정할 필요가 있고, 음성 신호의 도달 지연 시간을 측정하기 위하여 GCC(Generalized Cross Correlation) 알고리즘이 일반적으로 사용된다.

[0055] 다음으로, 실행된 음원 국지화 결과에 따라 다중 음원의 위치를 파악한다(S405). 여기서, '음원의 위치'라 함은 기준점(마이크로폰 또는 마이크로폰 어레이가 될 수 있다.)을 중심으로 음원이 위치한 방향 또는 거리를 의미한다. 즉, 마이크로폰 또는 마이크로폰 어레이를 구성하는 개별 마이크로폰들을 각각 다르게 지연시키면 특정 방향에 위치한 음원 신호에 대하여 지향성을 가지게 되고, 이러한 과정을 전 방향에 대하여 수행한다. 만약, 특정 방향으로부터 획득된 음원 신호의 음압이 최대값을 갖는다면 해당 방향에 음원이 존재한다고 파악할 수 있다. 또는, 음원 획득 장치에 의하여 획득된 다중 음원 신호를 디지털 신호로 변환하고, 변환된 디지털 신호를 다시 주파수 영역의 신호로 변환하여, 다중 음원의 주파수 영역에서 주파수 대역별 시간차를 산출하고, 산출된 주파수 대역별 시간차로부터 최종 시간차를 산출하여, 최종 시간차를 이용하여 다중 음원의 위치를 파악할 수 있다.

[0056] 다음으로, 다중 음원의 측정 시간이 소정의 시간을 초과하는지 여부를 판단한다(S406). 이는 소정의 시간 동안 다중 음원을 획득하여 음원 주파수 분석 및 음원 위치를 파악하는 것은 동영상 획득 대상인 영역에서 사람의 목소리가 가장 많이 파악되는 영역, 즉 사람이 가장 많은 곳을 찾기 위함이다. 판단 결과 소정의 측정 시간을 초과하지 않는 것으로 판단되면, 다시 다중 음원을 획득한다(S401).

[0057] 반면에, 판단 결과 소정의 측정 시간을 초과하는 것으로 판단되면, 단계 S405에서 파악된 다중 음원의 위치에 기초하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정한다(S407). 여기서, '얼굴 검출을 수행할 범위를 결정'한다는 것은 동영상 내에서 얼굴 검출을 수행할 때 전체 영상에 대하여

얼굴 검출을 수행하거나 또는 이미지 상에서 얼굴 가능 영역을 중심으로 얼굴 검출을 수행하는 것이 아니라, 다중 음원 위치의 밀도를 중심으로 하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하여 얼굴 검출을 수행할 범위를 결정하는 것을 의미한다.

[0058] 그리고, ‘다중 음원 위치의 밀도가 가장 큰 범위’라 함은 동영상 획득 대상인 영역에 대하여 복수의 음원 검출 영역으로 분할한 후 음원 국지화를 수행한 결과, 복수의 음원 검출 영역 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 의미한다. 또는, 동영상 획득 대상인 영역에 대하여 음원 국지화를 수행한 후 복수의 음원 검출 영역으로 분할한 결과, 복수의 음원 검출 영역 중에서 다중 음원 위치의 밀도가 가장 높은 음원 검출 영역을 의미할 수도 있다.

[0059] 또한, ‘복수의 음원 검출 영역으로 분할’한다는 것은 동영상 획득 대상인 영역에 대하여 다중 음원 위치의 밀도가 가장 큰 범위를 선정하기 위하여, 동영상 획득 대상인 영역을 소정의 기준에 따라 복수 개의 영역으로 분할하는 것을 의미한다. 여기서, 동영상 획득 대상인 영역을 복수의 음원 검출 영역으로 분할하는 방법은 동영상 획득 대상인 영역에 배치된 음원 획득 장치의 개수에 따라 달라질 수 있으나, 여기에만 한정되는 것은 아니다. 예를 들어, 동영상 획득 대상인 영역의 범위가 넓은 경우라면 음원 검출 영역을 더욱 세분화할 수 있다. 본 발명의 실시예에서는 4개의 마이크로폰을 음원 획득 장치로 이용하고, 음원 검출 영역을 16 영역으로 분할하였다. 그러나, 음원 검출 영역은 16 영역이 아니라 9 영역 또는 25 영역 또는 36 영역 등으로 분할할 수도 있다.

[0060] 다음으로, 파악된 다중 음원의 위치에 기초하고 소정의 알고리즘을 이용하여 다중 음원에 대한 얼굴 검출을 수행한다(S408). 또는, 단계 S407에서 결정된 얼굴 검출 범위에서 얼굴 검출을 수행할 수 있다. 본 발명의 실시예에 의하면, 얼굴 검출 방법에는 특별한 제한이 없는데, 얼굴 검출을 위하여 종래 주지된 얼굴 검출 인식 알고리즘 등을 이용할 수 있다. 예를 들어, 얼굴 검출에서 다양한 얼굴 크기를 검출하기 위해서는 피라미드 방식을 이용하여 입력영상의 크기를 점점 변경하고 고정된 크기의 얼굴 검출기로 검출을 수행하는 방법 또는 입력영상의 크기는 고정하고 얼굴 검출기의 특징추출 필터의 크기를 점점 변경하여 검출하는 방법 등을 이용할 수 있다. 이때, 일반적으로 매번 얼굴 검출기의 필터 크기를 변경하는 것보다 입력영상을 피라미드 방식으로 변경함으로써 보다 높은 검출 속도를 얻을 수 있다. 여기서, 얼굴 검출을 수행하는데 이용될 수 있는 하나의 방법은 이미지를 스캐닝하는 슬라이딩 윈도우를 이용하는 것이다. 슬라이딩 윈도우는 얼굴 샘플과 동일한 크기일 수 있는 20 X 20과 같은 미리 정의된 크기를 가질 수 있다. 얼굴 검출을 수행하기 위하여 이용되는 스캐닝 및 매칭 처리는 최대 얼굴 크기가 도달될 때까지 입력 이미지를 다운 샘플링함으로써 수회 반복될 수 있다.

[0061] 또한, 얼굴 검출 동작은 예를 들면, 아다부스트 학습 알고리즘에 의해서 학습되어 선택된 특정 패턴들을 이용하여 수행될 수 있지만, 반드시 여기에만 한정되는 것은 아니며 다양한 변형예를 상정할 수 있다. 이때, 얼굴 검출부(140)는 주변 픽셀들 사이의 상관 관계에 기반한 LBP(Local Binary Pattern) 알고리즘을 이용하여 얼굴 검출을 수행할 수 있다. LBP(Local Binary Pattern) 알고리즘은 중앙 픽셀값을 기준으로 8개의 인접한 픽셀값들과 비교하여 구성되며, 시계방향으로 중앙 픽셀값과 비교하여 크면 1, 그렇지 않으면 0의 값으로 변환한다. LBP는 8개의 인접한 픽셀들과 비교하여 구성되기 때문에 8비트로 표현되며, 0부터 255 사이의 값을 갖는다.

[0062] 도 5는 본 발명의 다른 실시예에 따른 LBP 알고리즘을 이용한 얼굴 검출 방법을 보여주는 흐름도이다.

[0063] 본 발명의 실시예에 의하면, 얼굴 검출 방법에는 특별한 제한이 없는데, 얼굴 검출을 위하여 얼굴 검출부(140)는 종래 주지된 얼굴 검출 인식 알고리즘 등을 이용할 수 있다. 예를 들어, 얼굴 검출에서 다양한 얼굴 크기를 검출하기 위해서는 피라미드 방식을 이용하여 입력영상의 크기를 점점 변경하고 고정된 크기의 얼굴 검출기로 검출을 수행하는 방법 또는 입력영상의 크기는 고정하고 얼굴 검출기의 특징추출 필터의 크기를 점점 변경하여 검출하는 방법 등을 이용할 수 있다. 이때 일반적으로 매번 얼굴 검출기의 필터 크기를 변경하는 것보다 입력영상을 피라미드 방식으로 변경함으로써 보다 높은 검출 속도를 얻을 수 있다. 여기서, 얼굴 검출을 수행하는데 이용될 수 있는 하나의 방법은 이미지를 스캐닝하는 슬라이딩 윈도우를 이용하는 것이다. 슬라이딩 윈도우는 얼굴 샘플과 동일한 크기일 수 있는 20 X 20과 같은 미리 정의된 크기를 가질 수 있다. 얼굴 검출을 수행하기 위하여 이용되는 스캐닝 및 매칭 처리는 최대 얼굴 크기가 도달될 때까지 입력 이미지를 다운 샘플링함으로써 수회 반복될 수 있다.

[0064] 도 5를 참조하면, 먼저 얼굴 검출이 시작되면 LBP(Local Binary Pattern) 변환을 수행하고(S501), 변환된 값에 가중치 값을 중첩한다(S502). 여기서, LBP(Local Binary Pattern) 알고리즘은 3 X 3 마스크내의 중심 픽셀에 대해 주변 픽셀과의 크기 비교를 통해 선택된 이진 패턴을 의미한다. 주변의 8개 픽셀에 대해 일정한 방향이 정

해지면 각 픽셀이 중심보다 크면 1, 작거나 같으면 0을 할당하여 8비트의 값이 결정되고, 이를 중심픽셀의 값으로 할당한다. LBP는 8개의 인접한 픽셀들과 비교하여 구성되기 때문에 8비트로 표현되며, 0부터 255 사이의 값을 갖는다.

[0065] 다름으로, LBP 변환값에 가중치값이 중첩된 결과값이 얼굴 후보에 해당하면 얼굴 후보를 등록하고(S503), 영상 크기가 슬라이딩 윈도우 기준 크기보다 큰지 여부를 판단한다(S504). 판단 결과, 영상 크기가 슬라이딩 윈도우 기준 크기보다 큰 것으로 판단되면, 영상을 축소하고(S505), 다시 LBP 변환을 수행한다(S501).

[0066] 반면에, 영상 크기가 슬라이딩 윈도운 기준 크기보다 작은 것으로 판단되면, 얼굴 검출 인증된 것으로 파악한다(S506).

[0067] 또한, 이러한 얼굴 검출 방법은 컴퓨터로 실행하기 위한 프로그램이 기록된 컴퓨터 판독가능 기록매체에 저장될 수 있다. 이때, 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록 장치를 포함한다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 제안된 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 통상의 기술자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 상기 매체는 프로그램 명령, 데이터 구조 등을 지정하는 신호를 전송하는 반송파를 포함하는 광 또는 금속선, 도파관 등의 전송 매체일 수도 있다. 프로그램명령의 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 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다. 또한, 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 장치에 분산되어 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

[0068] 이상의 설명은 본 발명의 기술 사상을 예시적으로 설명한 것에 불과한 것으로서, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자라면 본 발명의 본질적인 특성에서 벗어나지 않는 범위 내에서 다양한 수정, 변경 및 치환이 가능할 것이다. 따라서, 본 발명에 개시된 실시예 및 첨부된 도면들은 본 발명의 기술 사상을 한정하기 위한 것이 아니라 설명하기 위한 것이고, 이러한 실시예 및 첨부된 도면에 의하여 본 발명의 기술 사상의 범위가 한정되는 것은 아니다. 본 발명의 보호 범위는 아래의 청구범위에 의하여 해석되어야 하며, 그와 동등한 범위 내에 있는 모든 기술 사상은 본 발명의 권리범위에 포함되는 것으로 해석되어야 할 것이다.

부호의 설명

[0069] 10 : 얼굴 검출 장치

110 : 음원 주파수 분석부

120 : 음원 위치 파악부

130 : 얼굴 검출 범위 결정부

140 : 얼굴 검출부

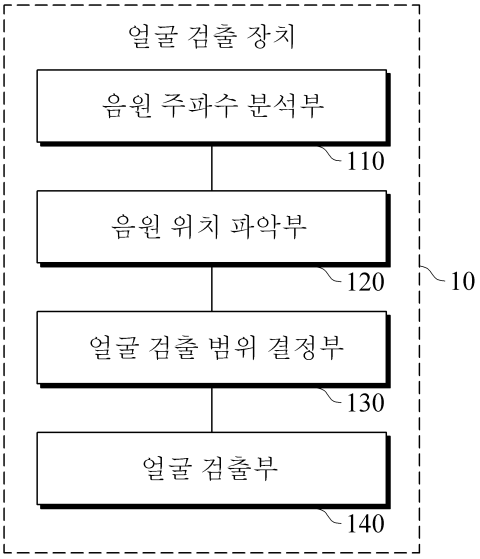
21, 22, 23, 24 : 음원 획득 장치

200 : 음원 검출 영역

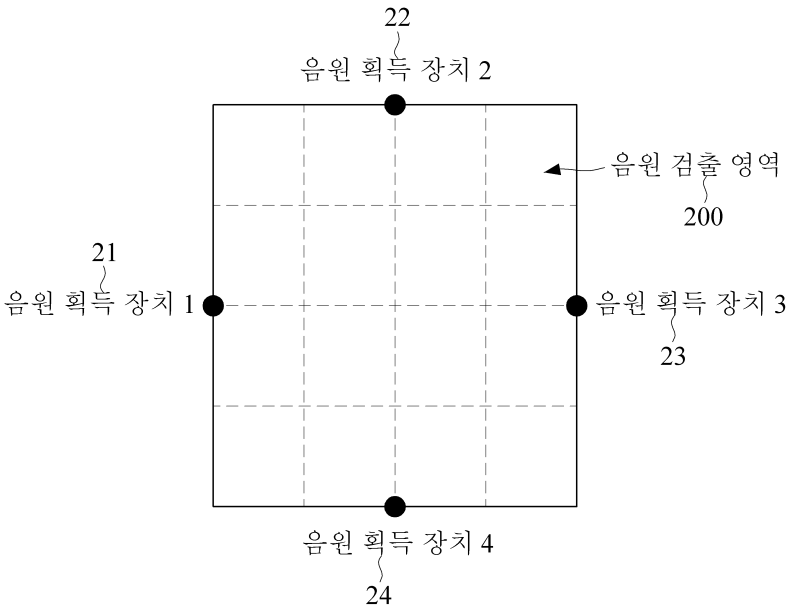
300 : 얼굴 검출 범위

도면

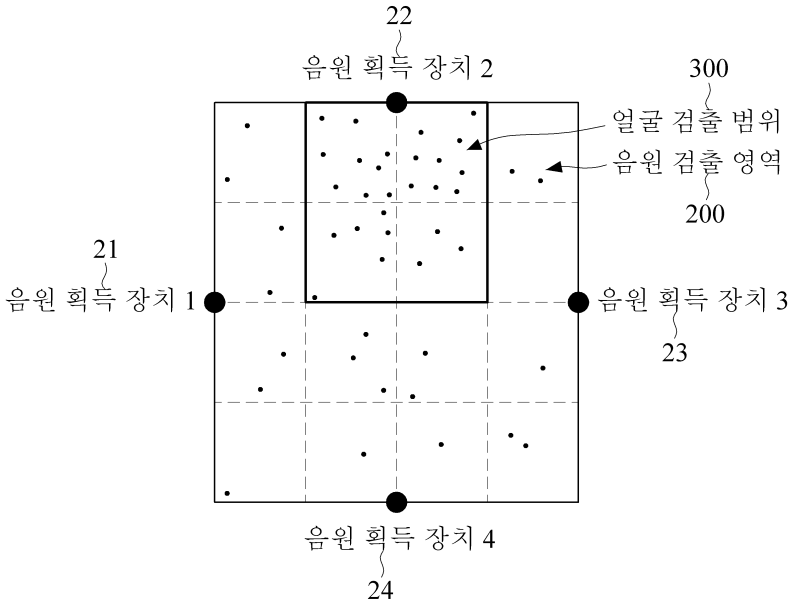
도면1


얼굴 검출 장치
음원 주파수 분석부
110
음원 위치 파악부
120
얼굴 검출 범위 결정부
130
얼굴 검출부
140
10

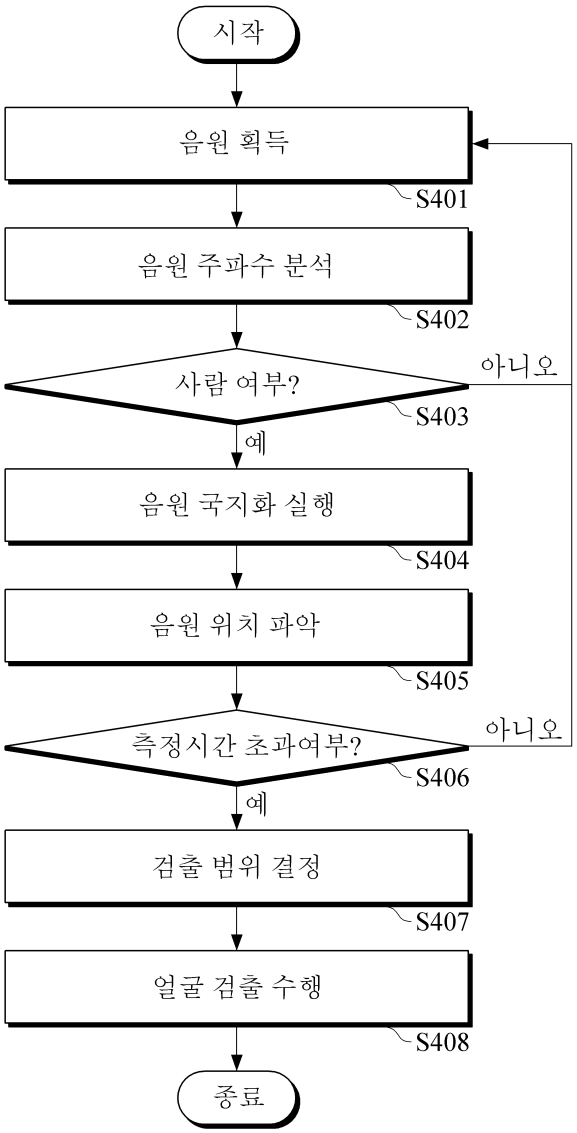
도면2


22
음원 획득 장치 2
음원 검출 영역
200
21
음원 획득 장치 1
음원 획득 장치 3
23
음원 획득 장치 4
24

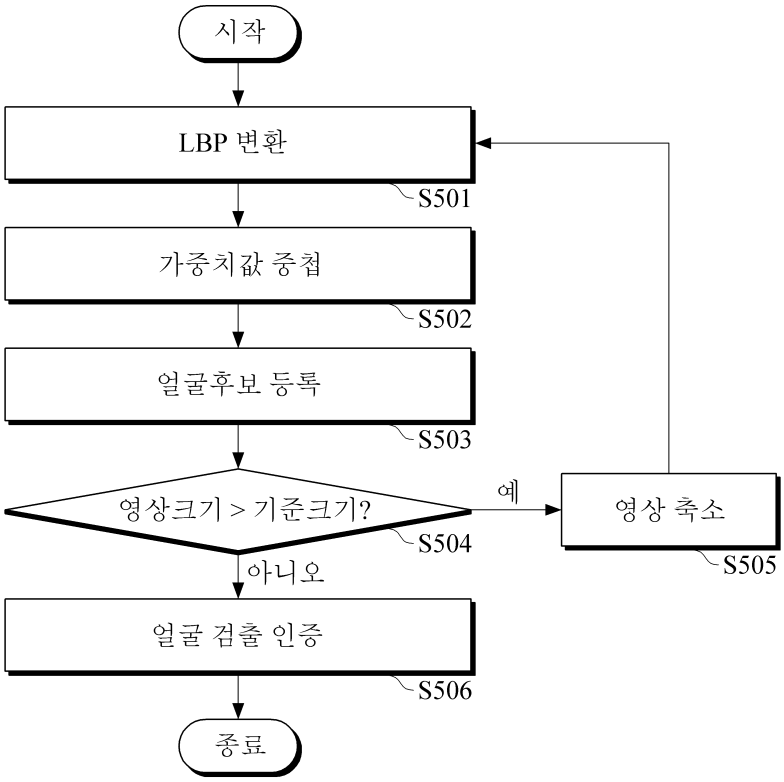
도면3


22
음원 획득 장치 2
300
얼굴 검출 범위
음원 검출 영역
200
21
음원 획득 장치 1
음원 획득 장치 3
23
음원 획득 장치 4
24

도면4


시작
음원 획득
S401
음원 주파수 분석
S402
사람 여부?
아니오
S403
예
음원 국지화 실행
S404
음원 위치 파악
S405
측정시간 초과여부?
아니오
S406
예
검출 범위 결정
S407
얼굴 검출 수행
S408
종료

도면5


시작
LBP 변환
S501
가중치값 중첩
S502
얼굴후보 등록
S503
영상크기 > 기준크기?
S504
예
영상 축소
S505
아니오
얼굴 검출 인증
S506
종료