

명세서

발명의 명칭: 오디오 신호에 방향성을 적용하는 전자 장치 및 그 방법

기술분야

- [1] 본 문서에 개시된 다양한 실시예는 전자 장치가 획득한 오디오 신호에 방향성을 적용하는 방법에 관한 것이다.

배경기술

- [2] 스테레오 오디오는, 복수의 음성 출력 구성들을 이용하여 둘 이상의 독립적인 오디오 채널을 사용하는 음향 제공 방법이다. 동일한 오디오 데이터에 스테레오 오디오의 정보를 포함할 수 있고, 각각의 독립적 오디오 채널을 이용하여 복수의 음성 출력 구성들로 하여금 각기 상이한 오디오를 출력하도록 하여 청취자에게 현장감을 제공할 수 있다.
- [3] 바이노럴(binaural) 오디오는, 복수의 음성 출력 구성들을 이용하여 바이노럴 효과(쌍이효과, 雙耳效果)를 제공하는 음향 제공 방법이다. 바이노럴 효과란, 사람의 양쪽 귀로 듣는 음향의 강도차, 시간차, 및/또는 위상차 등과 같은 차이를 이용하여 청취자로 하여금 원근감, 실제감, 방향감, 공간감, 음장(acoustic field)감을 갖도록 하는 효과를 의미한다.

발명의 상세한 설명

기술적 과제

- [4] 동영상 촬영과 함께 촬영된 영상에 대응하는 현장감 있는 오디오 신호의 녹음이 필요할 수 있다. 전자 장치를 이용하여 동영상을 촬영하는 경우, 녹음되는 배경 오디오 신호에 대하여 현장감, 방향감 등을 부여하는 것이 필요할 수 있다. 특히 동영상 촬영 시 무선 마이크로폰 등 외부 전자 장치를 이용하여 오디오 신호를 수집하는 경우, 단일한 외부 전자 장치를 이용하는 경우가 대부분일 수 있다. 이러한 단일 외부 전자 장치로 객체의 소리를 수음하여 동영상을 촬영하는 경우, 객체 소리의 방향성에 대한 정보가 없어, 촬영 시 객체가 촬영된 화면을 기준으로 좌측이나 우측 또는 어느 방향에 있든 관계 없이, 녹음된 객체의 오디오 신호는 고정된 방향에서 소리가 들려오는 것과 같은 효과를 줄 수 있다. 이 때 오디오 신호는 모노(mono)로 녹음될 수 있고, 모노로 녹음된 오디오 신호는 단조롭거나 현장감, 공간감을 제공하기 어려울 수 있다. 모노 신호로 녹음된 음성을 이용하여, 현장감, 공간감 등을 제공할 수 있는 방법이 필요할 수 있다.

과제 해결 수단

- [5] 본 문서에 개시된 다양한 실시예에 따른 전자 장치는, 근거리 무선 통신을 지원하는 통신 모듈, 동영상 이미지를 촬영하는 카메라 모듈, 상기 촬영하는 동영상 이미지를 표시하는 디스플레이, 및 상기 통신 모듈, 카메라 모듈 및

디스플레이와 작동적으로(operatively) 연결되는 프로세서를 포함하고, 상기 프로세서는, 상기 통신 모듈을 이용하여 외부 전자 장치와 연결을 수립하고, 상기 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오 신호를 수신하고, 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체 가운데 타겟이 되는 타겟 객체를 확인하고, 상기 타겟 객체가 상기 디스플레이 상에 표시되는 위치에 관한 제1위치 정보를 확인하고, 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여, 실제 위치에 관한 제2위치 정보를 생성하고, 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하도록 설정될 수 있다.

- [6] 본 문서에 개시된 다양한 실시예에 따른 전자 장치가 오디오 신호를 처리하는 방법에 있어서, 외부 전자 장치와 연결을 수립하는 동작, 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오 신호를 수신하는 동작, 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체 가운데 타겟이 되는 타겟 객체를 확인하는 동작, 상기 타겟 객체가 상기 전자 장치의 디스플레이 상에 표시되는 위치에 관한 제1위치 정보를 확인하는 동작, 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여, 실제 위치에 관한 제2위치 정보를 생성하는 동작, 및 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하는 동작을 포함할 수 있다.

발명의 효과

- [7] 다양한 실시예에 따르면, 촬영된 동영상에 대응하는, 현장감과 공간감을 가진 오디오 신호를 생성할 수 있다. 사용자에게 동영상의 이미지와 부합하는 공간감 있는 음향을 제공함으로써 사용자 경험을 제고할 수 있다.
- [8] 이 외에, 본 문서를 통해 직접적 또는 간접적으로 파악되는 다양한 효과들이 제공될 수 있다.

도면의 간단한 설명

- [9] 도면의 설명과 관련하여, 동일 또는 유사한 구성 요소에 대하여는 동일 또는 유사한 참조 부호가 사용될 수 있다.
- [10] 도 1은 다양한 실시예들에 따른, 네트워크 환경 내의 전자 장치의 블록도이다.
- [11] 도 2는 다양한 실시예에 따른 전자 장치가 동영상을 촬영하는 예시도이다.
- [12] 도 3은 다양한 실시예에 따른 전자 장치의 블록도이다.
- [13] 도 4는 다양한 실시예에 따른 전자 장치가 오디오 신호에 방향성을 적용하는 동작 흐름도이다.
- [14] 도 5, 도 6 및 도 7은 다양한 실시예에 따른 전자 장치가 타겟 객체를 확인하는 예시도이다.
- [15] 도 8은 다양한 실시예에 따른 부가 정보를 설명하기 위한 예시도이다.
- [16] 도 9는 다양한 실시예에 따른 스테레오 음향을 설명하기 위한 예시도이다.
- [17] 도 10 및 도 11은 다양한 실시예에 따른 공간감이 부여된 오디오 신호를

설명하기 위한 예시도이다.

발명의 실시를 위한 형태

- [18] 도 1은, 다양한 실시예들에 따른, 네트워크 환경(100) 내의 전자 장치(101)의 블록도이다. 도 1을 참조하면, 네트워크 환경(100)에서 전자 장치(101)는 제 1 네트워크(198)(예: 근거리 무선 통신 네트워크)를 통하여 전자 장치(102)와 통신하거나, 또는 제 2 네트워크(199)(예: 원거리 무선 통신 네트워크)를 통하여 전자 장치(104) 또는 서버(108) 중 적어도 하나와 통신할 수 있다. 일실시예에 따르면, 전자 장치(101)는 서버(108)를 통하여 전자 장치(104)와 통신할 수 있다. 일실시예에 따르면, 전자 장치(101)는 프로세서(120), 메모리(130), 입력 모듈(150), 음향 출력 모듈(155), 디스플레이 모듈(160), 오디오 모듈(170), 센서 모듈(176), 인터페이스(177), 연결 단자(178), 햅틱 모듈(179), 카메라 모듈(180), 전력 관리 모듈(188), 배터리(189), 통신 모듈(190), 가입자 식별 모듈(196), 또는 안테나 모듈(197)을 포함할 수 있다. 어떤 실시예에서는, 전자 장치(101)에는, 이 구성요소들 중 적어도 하나(예: 연결 단자(178))가 생략되거나, 하나 이상의 다른 구성요소가 추가될 수 있다. 어떤 실시예에서는, 이 구성요소들 중 일부들(예: 센서 모듈(176), 카메라 모듈(180), 또는 안테나 모듈(197))은 하나의 구성요소(예: 디스플레이 모듈(160))로 통합될 수 있다.
- [19] 프로세서(120)는, 예를 들면, 소프트웨어(예: 프로그램(140))를 실행하여 프로세서(120)에 연결된 전자 장치(101)의 적어도 하나의 다른 구성요소(예: 하드웨어 또는 소프트웨어 구성요소)를 제어할 수 있고, 다양한 데이터 처리 또는 연산을 수행할 수 있다. 일실시예에 따르면, 데이터 처리 또는 연산의 적어도 일부로서, 프로세서(120)는 다른 구성요소(예: 센서 모듈(176) 또는 통신 모듈(190))로부터 수신된 명령 또는 데이터를 휘발성 메모리(132)에 저장하고, 휘발성 메모리(132)에 저장된 명령 또는 데이터를 처리하고, 결과 데이터를 비휘발성 메모리(134)에 저장할 수 있다. 일실시예에 따르면, 프로세서(120)는 메인 프로세서(121)(예: 중앙 처리 장치 또는 어플리케이션 프로세서) 또는 이와는 독립적으로 또는 함께 운영 가능한 보조 프로세서(123)(예: 그래픽 처리 장치, 신경망 처리 장치(NPU: neural processing unit), 이미지 시그널 프로세서, 센서 허브 프로세서, 또는 커뮤니케이션 프로세서)를 포함할 수 있다. 예를 들어, 전자 장치(101)가 메인 프로세서(121) 및 보조 프로세서(123)를 포함하는 경우, 보조 프로세서(123)는 메인 프로세서(121)보다 저전력을 사용하거나, 지정된 기능에 특화되도록 설정될 수 있다. 보조 프로세서(123)는 메인 프로세서(121)와 별개로, 또는 그 일부로서 구현될 수 있다.
- [20] 보조 프로세서(123)는, 예를 들면, 메인 프로세서(121)가 인액티브(예: 슬립) 상태에 있는 동안 메인 프로세서(121)를 대신하여, 또는 메인 프로세서(121)가 액티브(예: 어플리케이션 실행) 상태에 있는 동안 메인 프로세서(121)와 함께, 전자 장치(101)의 구성요소들 중 적어도 하나의 구성요소(예: 디스플레이

모듈(160), 센서 모듈(176), 또는 통신 모듈(190))와 관련된 기능 또는 상태들의 적어도 일부를 제어할 수 있다. 일실시예에 따르면, 보조 프로세서(123)(예: 이미지 시그널 프로세서 또는 커뮤니케이션 프로세서)는 기능적으로 관련 있는 다른 구성요소(예: 카메라 모듈(180) 또는 통신 모듈(190))의 일부로서 구현될 수 있다. 일실시예에 따르면, 보조 프로세서(123)(예: 신경망 처리 장치)는 인공지능 모델의 처리에 특화된 하드웨어 구조를 포함할 수 있다. 인공지능 모델은 기계 학습을 통해 생성될 수 있다. 이러한 학습은, 예를 들어, 인공지능 모델이 수행되는 전자 장치(101) 자체에서 수행될 수 있고, 별도의 서버(예: 서버(108))를 통해 수행될 수도 있다. 학습 알고리즘은, 예를 들어, 지도형 학습(supervised learning), 비지도형 학습(unsupervised learning), 준지도형 학습(semi-supervised learning) 또는 강화 학습(reinforcement learning)을 포함할 수 있으나, 전술한 예에 한정되지 않는다. 인공지능 모델은, 복수의 인공 신경망 레이어들을 포함할 수 있다. 인공 신경망은 심층 신경망(DNN: deep neural network), CNN(convolutional neural network), RNN(recurrent neural network), RBM(restricted boltzmann machine), DBN(deep belief network), BRDNN(bidirectional recurrent deep neural network), 심층 Q-네트워크(deep Q-networks) 또는 상기 중 둘 이상의 조합 중 하나일 수 있으나, 전술한 예에 한정되지 않는다. 인공지능 모델은 하드웨어 구조 이외에, 추가적으로 또는 대체적으로, 소프트웨어 구조를 포함할 수 있다.

- [21] 메모리(130)는, 전자 장치(101)의 적어도 하나의 구성요소(예: 프로세서(120) 또는 센서 모듈(176))에 의해 사용되는 다양한 데이터를 저장할 수 있다. 데이터는, 예를 들어, 소프트웨어(예: 프로그램(140)) 및, 이와 관련된 명령에 대한 입력 데이터 또는 출력 데이터를 포함할 수 있다. 메모리(130)는, 휘발성 메모리(132) 또는 비휘발성 메모리(134)를 포함할 수 있다.
- [22] 프로그램(140)은 메모리(130)에 소프트웨어로서 저장될 수 있으며, 예를 들면, 운영 체제(142), 미들 웨어(144) 또는 어플리케이션(146)을 포함할 수 있다.
- [23] 입력 모듈(150)은, 전자 장치(101)의 구성요소(예: 프로세서(120))에 사용될 명령 또는 데이터를 전자 장치(101)의 외부(예: 사용자)로부터 수신할 수 있다. 입력 모듈(150)은, 예를 들면, 마이크, 마우스, 키보드, 키(예: 버튼), 또는 디지털 펜(예: 스타일러스 펜)을 포함할 수 있다.
- [24] 음향 출력 모듈(155)은 음향 신호를 전자 장치(101)의 외부로 출력할 수 있다. 음향 출력 모듈(155)은, 예를 들면, 스피커 또는 리시버를 포함할 수 있다. 스피커는 멀티미디어 재생 또는 녹음 재생과 같이 일반적인 용도로 사용될 수 있다. 리시버는 착신 전화를 수신하기 위해 사용될 수 있다. 일실시예에 따르면, 리시버는 스피커와 별개로, 또는 그 일부로서 구현될 수 있다.
- [25] 디스플레이 모듈(160)은 전자 장치(101)의 외부(예: 사용자)로 정보를 시각적으로 제공할 수 있다. 디스플레이 모듈(160)은, 예를 들면, 디스플레이, 홀로그램 장치, 또는 프로젝터 및 해당 장치를 제어하기 위한 제어 회로를 포함할 수 있다. 일실시예에 따르면, 디스플레이 모듈(160)은 터치를 감지하도록

설정된 터치 센서, 또는 상기 터치에 의해 발생하는 힘의 세기를 측정하도록 설정된 압력 센서를 포함할 수 있다.

- [26] 오디오 모듈(170)은 소리를 전기 신호로 변환시키거나, 반대로 전기 신호를 소리로 변환시킬 수 있다. 일실시예에 따르면, 오디오 모듈(170)은, 입력 모듈(150)을 통해 소리를 획득하거나, 음향 출력 모듈(155), 또는 전자 장치(101)와 직접 또는 무선으로 연결된 외부 전자 장치(예: 전자 장치(102))(예: 스피커 또는 헤드폰)를 통해 소리를 출력할 수 있다.
- [27] 센서 모듈(176)은 전자 장치(101)의 작동 상태(예: 전력 또는 온도), 또는 외부의 환경 상태(예: 사용자 상태)를 감지하고, 감지된 상태에 대응하는 전기 신호 또는 데이터 값을 생성할 수 있다. 일실시예에 따르면, 센서 모듈(176)은, 예를 들면, 제스처 센서, 자이로 센서, 기압 센서, 마그네틱 센서, 가속도 센서, 그립 센서, 근접 센서, 컬러 센서, IR(infrared) 센서, 생체 센서, 온도 센서, 습도 센서, 또는 조도 센서를 포함할 수 있다.
- [28] 인터페이스(177)는 전자 장치(101)가 외부 전자 장치(예: 전자 장치(102))와 직접 또는 무선으로 연결되기 위해 사용될 수 있는 하나 이상의 지정된 프로토콜들을 지원할 수 있다. 일실시예에 따르면, 인터페이스(177)는, 예를 들면, HDMI(high definition multimedia interface), USB(universal serial bus) 인터페이스, SD카드 인터페이스, 또는 오디오 인터페이스를 포함할 수 있다.
- [29] 연결 단자(178)는, 그를 통해서 전자 장치(101)가 외부 전자 장치(예: 전자 장치(102))와 물리적으로 연결될 수 있는 커넥터를 포함할 수 있다. 일실시예에 따르면, 연결 단자(178)는, 예를 들면, HDMI 커넥터, USB 커넥터, SD 카드 커넥터, 또는 오디오 커넥터(예: 헤드폰 커넥터)를 포함할 수 있다.
- [30] 햅틱 모듈(179)은 전기적 신호를 사용자가 촉각 또는 운동 감각을 통해서 인지할 수 있는 기계적인 자극(예: 진동 또는 움직임) 또는 전기적인 자극으로 변환할 수 있다. 일실시예에 따르면, 햅틱 모듈(179)은, 예를 들면, 모터, 압전 소자, 또는 전기 자극 장치를 포함할 수 있다.
- [31] 카메라 모듈(180)은 정지 영상 및 동영상을 촬영할 수 있다. 일실시예에 따르면, 카메라 모듈(180)은 하나 이상의 렌즈들, 이미지 센서들, 이미지 시그널 프로세서들, 또는 플래시들을 포함할 수 있다.
- [32] 전력 관리 모듈(188)은 전자 장치(101)에 공급되는 전력을 관리할 수 있다. 일실시예에 따르면, 전력 관리 모듈(188)은, 예를 들면, PMIC(power management integrated circuit)의 적어도 일부로서 구현될 수 있다.
- [33] 배터리(189)는 전자 장치(101)의 적어도 하나의 구성요소에 전력을 공급할 수 있다. 일실시예에 따르면, 배터리(189)는, 예를 들면, 재충전 불가능한 1차 전지, 재충전 가능한 2차 전지 또는 연료 전지를 포함할 수 있다.
- [34] 통신 모듈(190)은 전자 장치(101)와 외부 전자 장치(예: 전자 장치(102), 전자 장치(104), 또는 서버(108)) 간의 직접(예: 유선) 통신 채널 또는 무선 통신 채널의 수립, 및 수립된 통신 채널을 통한 통신 수행을 지원할 수 있다. 통신 모듈(190)은

프로세서(120)(예: 어플리케이션 프로세서)와 독립적으로 운영되고, 직접(예: 유선) 통신 또는 무선 통신을 지원하는 하나 이상의 커뮤니케이션 프로세서를 포함할 수 있다. 일실시예에 따르면, 통신 모듈(190)은 무선 통신 모듈(192)(예: 셀룰러 통신 모듈, 근거리 무선 통신 모듈, 또는 GNSS(global navigation satellite system) 통신 모듈) 또는 유선 통신 모듈(194)(예: LAN(local area network) 통신 모듈, 또는 전력선 통신 모듈)을 포함할 수 있다. 이들 통신 모듈 중 해당하는 통신 모듈은 제 1 네트워크(198)(예: 블루투스, WiFi(wireless fidelity) direct 또는 IrDA(infrared data association)와 같은 근거리 통신 네트워크) 또는 제 2 네트워크(199)(예: 레거시 셀룰러 네트워크, 5G 네트워크, 차세대 통신 네트워크, 인터넷, 또는 컴퓨터 네트워크(예: LAN 또는 WAN)와 같은 원거리 통신 네트워크)를 통하여 외부의 전자 장치(104)와 통신할 수 있다. 이런 여러 종류의 통신 모듈들은 하나의 구성요소(예: 단일 칩)로 통합되거나, 또는 서로 별도의 복수의 구성요소들(예: 복수 칩들)로 구현될 수 있다. 무선 통신 모듈(192)은 가입자 식별 모듈(196)에 저장된 가입자 정보(예: 국제 모바일 가입자 식별자(IMSII))를 이용하여 제 1 네트워크(198) 또는 제 2 네트워크(199)와 같은 통신 네트워크 내에서 전자 장치(101)를 확인 또는 인증할 수 있다.

- [35] 무선 통신 모듈(192)은 4G 네트워크 이후의 5G 네트워크 및 차세대 통신 기술, 예를 들어, NR 접속 기술(new radio access technology)을 지원할 수 있다. NR 접속 기술은 고용량 데이터의 고속 전송(eMBB(enhanced mobile broadband)), 단말 전력 최소화 및 다수 단말의 접속(mMTC(massive machine type communications)), 또는 고신뢰도와 저지연(URLLC(ultra-reliable and low-latency communications))을 지원할 수 있다. 무선 통신 모듈(192)은, 예를 들어, 높은 데이터 전송률 달성을 위해, 고주파 대역(예: mmWave 대역)을 지원할 수 있다. 무선 통신 모듈(192)은 고주파 대역에서의 성능 확보를 위한 다양한 기술들, 예를 들어, 빔포밍(beamforming), 거대 배열 다중 입출력(massive MIMO(multiple-input and multiple-output)), 전차원 다중입출력(FD-MIMO: full dimensional MIMO), 어레이 안테나(array antenna), 아날로그 빔형성(analog beam-forming), 또는 대규모 안테나(large scale antenna)와 같은 기술들을 지원할 수 있다. 무선 통신 모듈(192)은 전자 장치(101), 외부 전자 장치(예: 전자 장치(104)) 또는 네트워크 시스템(예: 제 2 네트워크(199))에 규정되는 다양한 요구사항을 지원할 수 있다. 일실시예에 따르면, 무선 통신 모듈(192)은 eMBB 실현을 위한 Peak data rate(예: 20Gbps 이상), mMTC 실현을 위한 손실 Coverage(예: 164dB 이하), 또는 URLLC 실현을 위한 U-plane latency(예: 다운링크(DL) 및 업링크(UL) 각각 0.5ms 이하, 또는 라운드 트립 1ms 이하)를 지원할 수 있다.

- [36] 안테나 모듈(197)은 신호 또는 전력을 외부(예: 외부의 전자 장치)로 송신하거나 외부로부터 수신할 수 있다. 일실시예에 따르면, 안테나 모듈(197)은 서브스트레이트(예: PCB) 위에 형성된 도전체 또는 도전성 패턴으로 이루어진 방사체를 포함하는 안테나를 포함할 수 있다. 일실시예에 따르면, 안테나

모듈(197)은 복수의 안테나들(예: 어레이 안테나)을 포함할 수 있다. 이런 경우, 제 1 네트워크(198) 또는 제 2 네트워크(199)와 같은 통신 네트워크에서 사용되는 통신 방식에 적합한 적어도 하나의 안테나가, 예를 들면, 통신 모듈(190)에 의하여 상기 복수의 안테나들로부터 선택될 수 있다. 신호 또는 전력은 상기 선택된 적어도 하나의 안테나를 통하여 통신 모듈(190)과 외부의 전자 장치 간에 송신되거나 수신될 수 있다. 어떤 실시예에 따르면, 방사체 이외에 다른 부품(예: RFIC(radio frequency integrated circuit))이 추가로 안테나 모듈(197)의 일부로 형성될 수 있다.

- [37] 다양한 실시예에 따르면, 안테나 모듈(197)은 mmWave 안테나 모듈을 형성할 수 있다. 일 실시예에 따르면, mmWave 안테나 모듈은 인쇄 회로 기판, 상기 인쇄 회로 기판의 제 1 면(예: 아래 면)에 또는 그에 인접하여 배치되고 지정된 고주파 대역(예: mmWave 대역)을 지원할 수 있는 RFIC, 및 상기 인쇄 회로 기판의 제 2 면(예: 윗 면 또는 측 면)에 또는 그에 인접하여 배치되고 상기 지정된 고주파 대역의 신호를 송신 또는 수신할 수 있는 복수의 안테나들(예: 어레이 안테나)을 포함할 수 있다.
- [38] 상기 구성요소들 중 적어도 일부는 주변 기기들간 통신 방식(예: 버스, GPIO(general purpose input and output), SPI(serial peripheral interface), 또는 MIPI(mobile industry processor interface))를 통해 서로 연결되고 신호(예: 명령 또는 데이터)를 상호간에 교환할 수 있다.
- [39] 일 실시예에 따르면, 명령 또는 데이터는 제 2 네트워크(199)에 연결된 서버(108)를 통해서 전자 장치(101)와 외부의 전자 장치(104)간에 송신 또는 수신될 수 있다. 외부의 전자 장치(102, 또는 104) 각각은 전자 장치(101)와 동일한 또는 다른 종류의 장치일 수 있다. 일 실시예에 따르면, 전자 장치(101)에서 실행되는 동작들의 전부 또는 일부는 외부의 전자 장치들(102, 104, 또는 108) 중 하나 이상의 외부의 전자 장치들에서 실행될 수 있다. 예를 들면, 전자 장치(101)가 어떤 기능이나 서비스를 자동으로, 또는 사용자 또는 다른 장치로부터의 요청에 반응하여 수행해야 할 경우에, 전자 장치(101)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 하나 이상의 외부의 전자 장치들에게 그 기능 또는 그 서비스의 적어도 일부를 수행하라고 요청할 수 있다. 상기 요청을 수신한 하나 이상의 외부의 전자 장치들은 요청된 기능 또는 서비스의 적어도 일부, 또는 상기 요청과 관련된 추가 기능 또는 서비스를 실행하고, 그 실행의 결과를 전자 장치(101)로 전달할 수 있다. 전자 장치(101)는 상기 결과를, 그대로 또는 추가적으로 처리하여, 상기 요청에 대한 응답의 적어도 일부로서 제공할 수 있다. 이를 위하여, 예를 들면, 클라우드 컴퓨팅, 분산 컴퓨팅, 모바일 에지 컴퓨팅(MEC: mobile edge computing), 또는 클라이언트-서버 컴퓨팅 기술이 이용될 수 있다. 전자 장치(101)는, 예를 들어, 분산 컴퓨팅 또는 모바일 에지 컴퓨팅을 이용하여 초저지연 서비스를 제공할 수 있다. 다른 실시예에 있어서, 외부의 전자 장치(104)는 IoT(internet of things)

기기를 포함할 수 있다. 서버(108)는 기계 학습 및/또는 신경망을 이용한 지능형 서버일 수 있다. 일 실시예에 따르면, 외부의 전자 장치(104) 또는 서버(108)는 제 2 네트워크(199) 내에 포함될 수 있다. 전자 장치(101)는 5G 통신 기술 및 IoT 관련 기술을 기반으로 지능형 서비스(예: 스마트 홈, 스마트 시티, 스마트 카, 또는 헬스케어)에 적용될 수 있다.

- [40] 도 2는 다양한 실시예에 따른 전자 장치(200)가 동영상을 촬영하는 예시도이다.
- [41] 도 2를 참조하면 전자 장치(200)(예: 도 1의 전자 장치(101))는 다양한 피사체(예: 외부 전자 장치(220) 및/또는 인물(230))의 영상을 촬영할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 전자 장치(200)에 포함된 카메라(예: 도 3의 카메라 모듈(320))를 이용하여 적어도 하나의 피사체(예: 외부 전자 장치(220) 및/또는 인물(230))를 촬영할 수 있다. 피사체가 될 수 있는 것은 인물(230), 장치(예: 외부 전자 장치(220))와 같이 다양한 인물 또는 물체가 될 수 있고, 피사체가 될 수 있는 객체에 제한이 없으나, 본 문서에서는 편의상 적어도 하나의 인물 및/또는 적어도 하나의 외부 전자 장치를 피사체로 하는 경우를 중심으로 설명하도록 한다.
- [42] 다양한 실시예에 따르면, 전자 장치(200)는 피사체(예: 인물(230) 및/또는 외부 전자 장치(220))를 촬영하여 촬영한 피사체의 이미지를 생성할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영한 이미지를 디스플레이(210) 상에 표시할 수 있다. 일 실시예에 따르면, 전자 장치(200)가 촬영한 이미지는 동영상 이미지일 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영 중인 동영상 이미지를 디스플레이(210)에 표시할 수 있다.
- [43] 다양한 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와 연결을 구성할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와 통신적으로(communicatively) 연결을 수립할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와 유선(예: 직접 통신) 및/또는 무선 통신 네트워크(예: 도 1의 제 1 네트워크(198))를 이용하여 연결을 수립할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와 근거리 무선 통신(예: 블루투스)을 이용하여 연결될 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)로 통신 연결 수립 및/또는 기능 수행에 필요한 데이터를 전송하거나, 외부 전자 장치(220)로부터 수신할 수 있다.
- [44] 다양한 실시예에 따르면, 전자 장치(200)는 오디오 신호를 획득할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 동영상을 촬영하는 경우 영상의 배경음에 해당하는 오디오 신호를 획득할 수 있다. 전자 장치(200)는 전자 장치(200)에 포함된 마이크로폰(예: 도 1의 입력 모듈(150))를 이용하여 외부 음성을 입력 받고, 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 연결된 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 예를 들면, 외부 전자 장치(220)는 수집한 음성을 이용하여 오디오 신호를 생성할 수 있고, 전자 장치(200)로 생성한 오디오 신호를 전송할 수 있다. 전자

장치(200)는 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 영상을 촬영함과 동시에 외부 전자 장치(220)로부터 영상에 대응하는 음성의 오디오 신호를 수신할 수 있다. 도 2를 참조하면, 전자 장치(200)가 표시한 동영상 이미지는 적어도 하나의 객체를 포함할 수 있다. 동영상 이미지에 포함된 적어도 하나의 객체는, 예를 들면, 인물(230)이 촬영된 이미지 객체인 제1객체(211) 및/또는 외부 전자 장치(220)가 촬영된 이미지 객체인 제2객체(221)를 포함할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영된 동영상 이미지 또는 촬영 중인 동영상 이미지를 분석하여 동영상 이미지에 포함된 적어도 하나의 이미지 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 메모리(예: 도 3의 메모리(340))에 기 저장된 알고리즘을 이용하여 이미지를 분석할 수 있고, 이미지 분석을 통해 동영상 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 표시된 이미지를 분석하여 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))가 디스플레이(210) 상에 표시되는 위치에 관한 정보(예: 좌표)(예: 제1위치 정보)를 확인할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))의 디스플레이(210) 상 표시되는 좌표(예: 제1위치 정보)를 지속적으로 확인할 수 있다. 전자 장치(200)는 촬영 중 디스플레이(210) 상에서 실시간으로 이동하는 객체(예: 제1객체(211) 및/또는 제2객체(221))의 좌표(예: 제1위치 정보)를 확인할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영한 동영상 이미지를 분석하여 타겟이 되는 객체(예: 타겟 객체)를 확인할 수 있다. 타겟 객체는, 예를 들면, 전자 장치(200)가 실제 위치를 추정하고자 하는 객체를 의미할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 이미지 분석을 이용하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 동영상 이미지를 분석하여 안면 인식을 수행할 수 있고, 안면 인식 수행 결과에 기초하여 인물 객체(예: 제1객체(211))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 인물 객체(예: 제1객체(211))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 촬영된 이미지를 분석하여 시각적 신호(예: 점멸 중인 LED 신호)를 식별하여 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로 확인할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 타겟 객체를 확인하기 위한 조건을 메모리(예: 도 3의 메모리(340))에 미리 저장할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 사용자(미도시)의 터치 입력을 수신할 수 있고, 수신한 터치 입력에 기초하여 타겟 객체를 확인할 수

수 있다. 예를 들면, 인식된 적어도 하나의 객체 중 터치 입력의 위치에 대응하는 객체를 타겟 객체로 인식할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 센서(예: 도 1의 센서 모듈(176))를 이용하여 센서 정보를 생성하고, 센서 정보에 기초하여 외부 전자 장치(220) 또는 인물(230)을 인식하고, 인식 결과에 기초하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와의 통신을 이용하여 외부 전자 장치(220)의 위치에 관한 정보를 수신하고, 수신한 위치 정보를 저장할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영된 동영상 이미지 분석, 수신된 터치 입력 분석, 센서 정보, 및 수신한 위치 정보 중 적어도 하나에 기초하여 타겟 객체를 식별하고 확인할 수 있다.

- [45] 다양한 실시예에 따르면, 전자 장치(200)는 타겟 객체가 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 존재하는 모든 객체(예: 제1객체(211) 및 제2객체(221))의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 표시된 적어도 하나의 이미지 객체(예: 제1객체(211) 및/또는 제2객체(221)) 가운데, 타겟 객체(예: 제1객체(211) 또는 제2객체(221))의 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다. 제1위치 정보는, 타겟 객체가 디스플레이(210) 상에 표시되는 위치에 관한 정보일 수 있다. 일 실시예에 따르면, 제1위치 정보는 디스플레이(210) 상의 일정한 좌표로 표현되는 정보일 수 있다. 일 실시예에 따르면, 제1위치 정보는 전자 장치(200)의 동영상 촬영 중에 실시간으로 변경되는 정보일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 동영상 촬영 중에 제1위치 정보를 지속적 및/또는 즉각적으로 확인할 수 있다.
- [46] 다양한 실시예에 따르면, 전자 장치(200)는 부가 정보를 확인할 수 있다. 부가 정보는 전자 장치(200)가 타겟 객체에 대응되는 피사체(예: 외부 전자 장치(220) 또는 인물(230))의 실제 위치에 관한 정보(예: 제2위치 정보)를 추정하기 위하여 제1위치 정보 이외에 이용되는 정보일 수 있다. 부가 정보는, 전자 장치(200)에 포함된 카메라(예: 도 3의 카메라 모듈(320))에 설정된 정보를 포함할 수 있다. 다양한 실시예에 따르면, 부가 정보는 촬영 중인 전자 장치(200)의 화각(field of view, FOV) 및/또는 배율과 같이, 카메라(예: 카메라 모듈(320))의 상태 및/또는 설정에 관한 정보를 포함할 수 있다. 일 실시예에 따르면, 카메라(예: 도 3의 카메라 모듈(320))는 거리를 측정할 수 있는 깊이 카메라(depth camera)를 포함할 수 있고, 전자 장치(200) 및 타겟 객체의 거리를 측정할 수 있다. 일 실시예에 따르면, 부가 정보는 전자 장치(200)와 타겟 객체 간의 거리 정보를 포함할 수 있다. 일 실시예에 따르면, 부가 정보는 타겟 객체의 크기를 포함할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 타겟 객체의 이미지에 대응하는 디스플레이의 영역 범위를 확인하고, 타겟 객체의 크기(예: 길이 및/또는 면적)를 확인할 수 있다.

- [47] 다양한 실시예에 따르면, 전자 장치(200)는 피사체(예: 인물(230) 및/또는 외부 전자 장치(220))의 위치를 추정할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 촬영한 동영상 이미지에 포함된 모든 객체(예: 제1객체(211) 및 제2객체(221))에 대응하는 피사체(예: 인물(230) 및 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 타겟 객체에 대응하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치만을 추정할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 타겟 객체에 대응하는 피사체의 실제 위치를 추정하고, 추정한 위치에 관한 제2위치 정보를 생성할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 제1위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 촬영된 이미지 객체(예: 타겟 객체)의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인하고, 디스플레이(210) 상 위치에 기초하여 실제 위치(예: 제2위치 정보)를 추정할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 부가 정보에 기초하여 제2위치 정보를 추정할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 타겟 객체의 제1위치 정보 및 부가 정보를 이용하여, 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 전자 장치(200)가 추정하는 실제 위치는, 전자 장치(200)를 기준으로 하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 상대적인 위치를 의미할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 구비한 센서를 통해 생성한 센서 정보에 기초하여 제2위치 정보를 생성할 수 있다. 예를 들면, 전자 장치(200)는 적외선 센서와 같은 거리 측정 센서를 이용하여 타겟 객체와의 거리를 확인하고, 확인한 거리를 기초로 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)로부터 외부 전자 장치(220)의 위치 정보를 수신할 수 있고, 수신한 위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 제2위치 정보는, 좌우 편향된 위치만을 표현하는 1차원적 위치 정보일 수 있고, 또는, 전자 장치(200)를 기준으로 앞뒤(예: 거리)를 포함할 수 있고, 또는, 상하로 편향된 위치를 표현하여 3차원적 위치 정보일 수 있다. 다양한 실시예에 따르면, 제2위치 정보는 피사체의 전후 위치, 좌우 위치, 및 상하 위치 중 적어도 하나 또는 그 조합(예: 1차원적 위치, 2차원적 위치, 또는 3차원적 위치)을 포함할 수 있다.
- [48] 다양한 실시예에 따르면, 전자 장치(200)는 제2위치 정보에 기초하여 오디오 신호를 처리할 수 있다. 오디오 신호의 처리는, 획득한 오디오 신호에 방향감을 부여하는 동작일 수 있다. 오디오 신호의 처리는, 예를 들면, 오디오 신호의 변경, 및/또는 변환을 포함할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 획득한 오디오 신호를 패닝(panning)하여 스테레오 오디오 신호로 변환할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 획득한 오디오 신호를 렌더링(rendering) 하여 공간감, 위치감 및/또는 방향감을 제공하는 3차원 음향(예: 바이노럴 음향)으로 변환할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 획득한 오디오 신호의

음량을 조절하여 거리감을 제공하도록 오디오 신호를 처리할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 하나의 오디오 신호를 처리하여 청취자 기준 좌측 귀로 듣는 신호(좌측 오디오) 및 우측 귀로 듣는 신호(우측 오디오)를 각각 생성할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 좌측 오디오 및 우측 오디오에 각각 음향의 강도 차이, 시간 차이 및 위상 차이 중 적어도 하나를 생성하여 오디오 신호를 처리할 수 있다.

- [49] 도 2를 참조하면, 외부 전자 장치(220)는 전자 장치(200)와 통신적으로 연결을 수립할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 전자 장치(200)로 오디오 신호를 전송할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 음성을 입력 받고 입력 받은 음성을 이용하여 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 센서를 구비하여 센서 정보를 생성하고, 생성한 센서 정보를 전자 장치(200)로 전송할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 전자 장치(200)로 위치를 알리는 신호(예: UWB(ultra wide band) 신호)를 전송할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 외부 전자 장치(220)의 위치를 확인하여 위치 정보를 생성하고, 생성된 위치 정보를 전자 장치(200)로 전송할 수 있다.
- [50] 도 3은 다양한 실시예에 따른 전자 장치의 블록도이다.
- [51] 도 3을 참조하면 전자 장치(300)(예: 도 1의 전자 장치(101) 및/또는 도 2의 전자 장치(200))는 통신 모듈(310), 카메라 모듈(320), 디스플레이(330), 메모리(340) 및 프로세서(350)를 포함할 수 있다. 전자 장치(300)는 도 1의 전자 장치(101)의 구성 및/또는 기능 가운데 적어도 일부를 포함할 수 있다.
- [52] 다양한 실시예들에 따르면, 통신 모듈(310)은 유선 및/또는 무선 네트워크 통신(예: 도 1의 제 1 네트워크(198) 또는 제 2 네트워크(199))를 이용해 외부 전자 장치(예: 도 2의 외부 전자 장치(220))와 통신할 수 있다. 통신 모듈(310)에 의해 지원되는 원거리 통신에는 제한이 없고, 다양한 통신 방식(예: 블루투스, UWB)을 지원할 수 있다. 다양한 실시예에 따르면, 통신 모듈(310)은 근거리 무선 통신(예: 블루투스(Bluetooth), BLE(Bluetooth low energy), WiFi(wireless fidelity) direct, 및/또는 UWB(ultra wide band))를 지원할 수 있고, 근거리 무선 통신을 이용하여 외부 전자 장치(220)로 정보를 전송할 수 있다. 다양한 실시예에 따르면, 통신 모듈(310)은 외부 전자 장치(220)와 단방향성 또는 양방향성 통신을 수행할 수 있다. 단방향성 통신은, 예를 들면, 다른 전자 장치에게 정보를 송신하는 것에 한정될 수 있고, 정보의 송신은 단순히 일정 신호를 외부에 출력하는 것으로 이루어질 수도 있다.
- [53] 다양한 실시예에 따르면, 다양한 실시예에 따르면, 카메라 모듈(320)은 전자 장치(300)의 외부 환경의 이미지 및/또는 영상을 촬영할 수 있다. 카메라 모듈(320)은 도 1의 카메라 모듈(180)의 구성 및/또는 기능의 적어도 일부를 포함할 수 있다. 다양한 실시예에 따르면, 카메라 모듈(320)은 외부로부터 입사된 광을 전기적 신호로 변환하여 이미지 정보를 생성할 수 있다. 다양한 실시예에

따르면, 카메라 모듈(320)은 전자 장치(300)의 외부 환경을 촬영하고, 주변 환경을 촬영한 동영상 이미지를 생성할 수 있다. 일 실시예에 따르면, 카메라 모듈(320)은 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))를 촬영하고, 촬영한 피사체의 디지털 이미지를 생성할 수 있다. 일 실시예에 따르면, 카메라 모듈(320)은 거리를 측정할 수 있는 깊이 카메라(depth camera)를 포함할 수 있다.

- [54] 다양한 실시예에 따르면, 디스플레이(330)는 전자 장치(300) 외부로 정보를 표시할 수 있다. 디스플레이(330)은 도 1의 디스플레이 모듈(160)의 구성 및/또는 기능 가운데 적어도 일부를 포함할 수 있다. 다양한 실시예에 따르면, 디스플레이(330)은 디스플레이 패널을 포함하고, 프로세서(350)로부터 수신한 정보를 시각적으로 표시할 수 있다. 다양한 실시예에 따르면, 디스플레이(330)은 입력 모듈(331)을 포함할 수 있다. 일 실시예에 따르면, 디스플레이(330)은 터치 센서 및/또는 압력 센서를 포함하고, 사용자의 터치 입력을 수신할 수 있다.
- [55] 다양한 실시예에 따르면, 메모리(340)는 디지털 데이터들을 일시적 또는 영구적으로 저장하기 위한 것으로서, 도 1의 메모리(130)의 구성 및/또는 기능 가운데 적어도 일부를 포함할 수 있다. 또한, 메모리(340)는 도 1의 프로그램(140) 가운데 적어도 일부를 저장할 수 있다. 메모리(340)는 프로세서(350)에 의하여 수행될 수 있는 다양한 인스트럭션(instruction)을 저장할 수 있다. 이와 같은 인스트럭션들은 프로세서(350)에 의하여 인식되고 실행될 수 있는 논리 연산, 데이터 입출력 등의 제어 명령을 포함할 수 있다. 메모리(340)가 저장할 수 있는 데이터의 종류 및/또는 양에 대하여는 한정됨이 없을 것이나, 본 문서에서는 다양한 실시예들에 따른 오디오 신호를 처리하는 방법 및 그 방법을 수행하는 프로세서(350)의 동작과 관련된 메모리의 구성 및 기능에 대하여만 설명하기로 한다.
- [56] 다양한 실시예에 따르면, 프로세서(350)는 전자 장치(300)의 각 구성요소들의 제어 및/또는 통신에 관한 연산이나 데이터를 처리할 수 있다. 프로세서(350)는 도 1의 프로세서(120)의 구성 및/또는 기능 가운데 적어도 일부를 포함할 수 있다. 프로세서는 통신 모듈(310), 카메라 모듈(320), 디스플레이(330) 및 메모리(340)와 같은 전자 장치(300)의 구성요소와 작동적(operatively), 전기적(electrically) 및/또는 기능적(functionally)으로 연결될 수 있다. 다양한 실시예에 따르면 프로세서(350)의 각 동작은 실시간으로 이루어질 수 있다. 예를 들면, 프로세서(350)가 오디오 신호를 처리하기 위하여 수행하는 일련의 연산 및/또는 동작들은 순차적 또는 동시에, 매우 작은 시간 범위 내에서 이루어질 수 있다. 프로세서(350)가 수행할 수 있는 동작, 연산 및 데이터 처리의 종류 및/또는 양에 대하여는 한정됨이 없을 것이나, 본 문서에서는 다양한 실시예들에 따른 오디오 신호를 처리하는 방법 및 그 방법을 수행하는 동작과 관련된 프로세서(350)의 구성 및 기능에 대하여만 설명하기로 한다.
- [57] 다양한 실시예에 따르면, 프로세서(350)는 동영상을 촬영하고 오디오 신호를

수신할 수 있다. 프로세서(350)는 다양한 피사체(예: 도 2의 외부 전자 장치(220) 및/또는 인물(230))의 영상을 촬영할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 카메라 모듈(320)을 이용하여 적어도 하나의 피사체(예: 외부 전자 장치(220) 및/또는 인물(230))를 촬영할 수 있다. 피사체가 될 수 있는 것은 인물(230), 장치(예: 외부 전자 장치(220))와 같이 다양한 객체가 될 수 있고, 피사체가 될 수 있는 객체에 제한이 없으나, 본 문서에서는 편의상 적어도 하나의 인물 및/또는 적어도 하나의 외부 전자 장치를 피사체로 하는 경우를 중심으로 설명하도록 한다. 다양한 실시예에 따르면, 프로세서(350)는 피사체(예: 인물(230) 및/또는 외부 전자 장치(220))를 촬영하여 촬영한 피사체의 이미지를 생성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영한 이미지를 디스플레이(330)(예: 도 2의 디스플레이(210)) 상에 표시할 수 있다. 일 실시예에 따르면, 프로세서(350)가 촬영한 이미지는 동영상 이미지일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영 중인 동영상 이미지를 디스플레이(330)에 표시할 수 있다.

[58] 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(예: 도 2의 외부 전자 장치(220))와 연결을 구성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 통신적으로(communicatively) 연결을 수립할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 유선(예: 직접 통신) 및/또는 무선 통신 네트워크(예: 도 1의 제 1 네트워크(198))를 이용하여 연결을 수립할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 근거리 무선 통신(예: 블루투스)을 이용하여 연결될 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)로 통신 연결 수립 및/또는 기능 수행에 필요한 데이터를 전송하거나, 외부 전자 장치(220)로부터 수신할 수 있다.

[59] 다양한 실시예에 따르면, 프로세서(350)는 오디오 신호를 획득할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 동영상을 촬영하는 경우, 영상의 배경음에 해당하는 오디오 신호를 획득할 수 있다. 프로세서(350)는 프로세서(350)에 포함된 마이크로폰(예: 도 1의 입력 모듈(150))를 이용하여 외부 음성을 입력 받고, 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 연결된 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 예를 들면, 외부 전자 장치(220)는 수집한 음성을 이용하여 오디오 신호를 생성할 수 있고, 프로세서(350)로 생성한 오디오 신호를 전송할 수 있다. 프로세서(350)는 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 일 실시예에 따르면, 프로세서(350)는 영상을 촬영함과 동시에 외부 전자 장치(220)로부터 영상에 대응하는 음성의 오디오 신호를 수신할 수 있다. 일 실시예에 따르면, 프로세서(350)가 외부 전자 장치(220)로부터 수신한 오디오 신호는 모노(mono) 음향일 수 있다.

[60] 다양한 실시예에 따르면, 프로세서(350)는 타겟 객체를 확인할 수 있다. 다양한

실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지를 분석하여 타겟이 되는 객체(예: 타겟 객체)를 확인할 수 있다. 타겟 객체는, 예를 들면, 프로세서(350)가 실제 위치를 추정하고자 하는 객체를 의미할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지를 디스플레이(330)에 표시할 수 있다. 프로세서(350)가 표시한 동영상 이미지는 적어도 하나의 객체를 포함할 수 있다. 동영상 이미지에 포함된 적어도 하나의 객체는, 예를 들면, 인물(예: 도 2의 인물(230))이 촬영된 이미지 객체인 제1객체(예: 도 2의 제1객체(211)) 및/또는 외부 전자 장치(220)가 촬영된 이미지 객체인 제2객체(예: 도 2의 제2객체(221))를 포함할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영된 동영상 이미지 또는 촬영 중인 동영상 이미지를 분석하여 동영상 이미지에 포함된 적어도 하나의 이미지 객체(예: 도 2의 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 메모리(340)에 기 저장된 알고리즘을 이용하여 이미지를 분석할 수 있고, 이미지 분석을 통해 동영상 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 표시된 이미지를 분석하여 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다.

- [61] 다양한 실시예에 따르면, 프로세서(350)는 이미지 분석을 이용하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 동영상 이미지를 분석하여 안면 인식을 수행할 수 있고, 안면 인식 수행 결과에 기초하여 인물 객체(예: 제1객체(211))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 식별된 인물 객체(예: 제1객체(211))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 식별된 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영된 이미지를 분석하여 시각적 신호(예: 점멸 중인 LED 신호)를 식별하여 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 타겟 객체를 확인하기 위한 조건을 메모리(예: 도 3의 메모리(340))에 미리 저장할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 사용자(미도시)의 터치 입력을 수신할 수 있고, 수신한 터치 입력에 기초하여 타겟 객체를 확인할 수 있다. 예를 들면, 인식된 적어도 하나의 객체 중 터치 입력의 위치에 대응하는 객체를 타겟 객체로 인식할 수 있다. 일 실시예에 따르면, 프로세서(350)는 센서(예: 도 1의 센서 모듈(176))를 이용하여 센서 정보를 생성하고, 센서 정보에 기초하여 외부 전자 장치(220) 또는 인물(230)을 인식하고, 인식 결과에 기초하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와의 통신을 이용하여 외부 전자 장치(220)의 위치에 관한 정보를 수신하고, 수신한 위치 정보를 저장할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영된 동영상 이미지 분석, 수신된

터치 입력 분석, 센서 정보, 및 수신한 위치 정보 중 적어도 하나에 기초하여 타겟 객체를 식별하고 확인할 수 있다.

- [62] 다양한 실시예에 따르면, 프로세서(350)는 제1위치 정보 및 부가 정보를 확인할 수 있다. 제1위치 정보는, 타겟 객체가 디스플레이(210) 상에 표시되는 위치에 관한 정보일 수 있다. 일 실시예에 따르면, 제1위치 정보는 디스플레이(210) 상의 일정한 좌표로 표현되는 정보일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))가 디스플레이(210) 상에 표시되는 위치에 관한 정보(예: 좌표)(예: 제1위치 정보)를 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))의 디스플레이(210) 상 표시되는 좌표(예: 제1위치 정보)를 지속적으로 확인할 수 있다. 프로세서(350)는 촬영 중 디스플레이(210) 상에서 실시간으로 이동하는 객체(예: 제1객체(211) 및/또는 제2객체(221))의 좌표(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 제1위치 정보는 프로세서(350)의 동영상 촬영 중에 실시간으로 변경되는 정보일 수 있다. 일 실시예에 따르면, 프로세서(350)는 동영상 촬영 중에 제1위치 정보를 지속적 및 즉각적으로 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 타겟 객체가 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 존재하는 모든 객체(예: 제1객체(211) 및 제2객체(221))의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 표시된 적어도 하나의 이미지 객체(예: 제1객체(211) 및/또는 제2객체(221)) 가운데 타겟 객체(예: 제1객체(211) 또는 제2객체(221))의 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다.

- [63] 다양한 실시예에 따르면, 프로세서(350)는 부가 정보를 확인할 수 있다. 부가 정보는 프로세서(350)가 타겟 객체에 대응되는 피사체(예: 외부 전자 장치(220) 또는 인물(230))의 실제 위치에 관한 정보(예: 제2위치 정보)를 추정하기 위하여 제1위치 정보 이외에 이용되는 정보일 수 있다. 부가 정보는, 전자 장치(300)에 포함된 카메라 모듈(320)에 설정된 정보를 포함할 수 있다. 다양한 실시예에 따르면, 부가 정보는 촬영 중인 프로세서(350)의 화각(field of view, FOV) 및/또는 배율과 같이, 카메라 모듈(320)의 상태 및/또는 설정에 관한 정보를 포함할 수 있다. 일 실시예에 따르면, 프로세서(350) 및 타겟 객체의 거리를 측정할 수 있다. 일 실시예에 따르면, 부가 정보는 전자 장치(300)와 타겟 객체 간의 거리정보를 포함할 수 있다. 일 실시예에 따르면, 부가 정보는 타겟 객체의 크기를 포함할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체의 이미지에 대응하는 디스플레이의 영역 범위를 확인하고, 타겟 객체의 크기(예: 길이 및/또는 면적)를 확인할 수 있다.

- [64] 다양한 실시예에 따르면, 프로세서(350)는 제2위치 정보를 생성할 수 있다. 제2위치 정보는, 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))의

실제 위치에 관한 정보일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))의 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지에 포함된 모든 객체(예: 제1객체(211) 및 제2객체(221))에 대응하는 피사체(예: 인물(230) 및 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체에 대응하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치만을 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체에 대응하는 피사체의 실제 위치를 추정하고, 추정된 위치에 관한 제2위치 정보를 생성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 제1위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영된 이미지 객체(예: 타겟 객체)의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인하고, 디스플레이(210) 상 위치에 기초하여 실제 위치(예: 제2위치 정보)를 추정할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 부가 정보에 기초하여 제2위치 정보를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체의 제1위치 정보 및 부가 정보를 이용하여 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)가 추정하는 실제 위치는, 프로세서(350)를 기준으로 하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 상대적인 위치를 의미할 수 있다. 일 실시예에 따르면, 프로세서(350)는 전자 장치(300)에 포함된 센서(예: 도 1의 센서 모듈(176))를 통해 생성한 센서 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)로부터 외부 전자 장치(220)의 위치 정보를 수신할 수 있고, 수신한 위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 제2위치 정보는, 좌우 편향된 위치만을 표현하는 1차원적 위치 정보일 수 있고, 또는, 프로세서(350)를 기준으로 앞뒤(예: 거리)를 포함할 수 있고, 또는, 상하로 편향된 위치를 표현하여 3차원적 위치 정보일 수 있다. 다양한 실시예에 따르면, 제2위치 정보는 피사체의 전후 위치, 좌우 위치 및 상하 위치 중 적어도 하나 또는 그 조합을 포함할 수 있다.

- [65] 다양한 실시예에 따르면, 프로세서(350)는 제2위치 정보에 기초하여 오디오 신호를 처리할 수 있다. 오디오 신호의 처리는, 획득한 오디오 신호에 방향감을 부여하는 동작일 수 있다. 오디오 신호의 처리는, 예를 들면, 오디오 신호의 변경 및/또는 변환을 포함할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호를 패닝(panning)하여 스테레오 오디오 신호로 변환할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호를 렌더링(rendering) 하여 공간감, 위치감 및/또는 방향감을 제공하는 3차원 음향(예: 바이노럴 음향)으로 변환할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호의 음량을 조절하여 거리감을 제공하도록 오디오 신호를 처리할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 하나의 오디오 신호를 처리하여 청취자 기준

좌측 귀로 듣는 신호(좌측 오디오) 및 우측 귀로 듣는 신호(우측 오디오)를 각각 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 좌측 오디오 및 우측 오디오에 각각 음향의 강도 차이, 시간 차이, 및 위상 차이 중 적어도 하나를 생성하여 오디오 신호를 처리할 수 있다.

- [66] 도 4는 다양한 실시예에 따른 전자 장치가 오디오 신호에 방향성을 적용하는 동작 흐름도이다.
- [67] 도 4를 참조하면, 전자 장치(예: 도 3의 전자 장치(300))가 오디오 신호를 처리하는 일련의 각 동작은, 전자 장치(300)에 포함된 프로세서(예: 도 3의 프로세서(350))가 수행하는 동작으로 표현될 수 있다.
- [68] 동작 410을 참조하면, 프로세서(350)는 동영상을 촬영하고 오디오 신호를 수신할 수 있다. 프로세서(350)는 다양한 피사체(예: 도 2의 외부 전자 장치(220) 및/또는 인물(230))의 영상을 촬영할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 카메라 모듈(예: 도 3의 카메라 모듈(320))을 이용하여 적어도 하나의 피사체(예: 외부 전자 장치(220) 및/또는 인물(230))를 촬영할 수 있다. 피사체가 될 수 있는 것은 인물(230), 장치(예: 외부 전자 장치(220))와 같이 다양한 객체가 될 수 있고, 피사체가 될 수 있는 객체에 제한이 없으나, 본 문서에서는 편의상 적어도 하나의 인물 및/또는 적어도 하나의 외부 전자 장치를 피사체로 하는 경우를 중심으로 설명하도록 한다. 다양한 실시예에 따르면, 프로세서(350)는 피사체(예: 인물(230) 및/또는 외부 전자 장치(220))를 촬영하여 촬영한 피사체의 이미지를 생성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영한 이미지를 디스플레이(예: 도 3의 디스플레이(330)) 상에 표시할 수 있다. 일 실시예에 따르면, 프로세서(350)가 촬영한 이미지는 동영상 이미지일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영 중인 동영상 이미지를 디스플레이(330)에 표시할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(예: 도 2의 외부 전자 장치(220))와 연결을 구성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 통신적으로(communicatively) 연결을 수립할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 유선(예: 직접 통신) 및/또는 무선 통신 네트워크(예: 도 1의 제 1 네트워크(198))를 이용하여 연결을 수립할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와 근거리 무선 통신(예: 블루투스)을 이용하여 연결될 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)로 통신 연결 수립 및/또는 기능 수행에 필요한 데이터를 전송하거나, 외부 전자 장치(220)로부터 수신할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 오디오 신호를 획득할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 동영상을 촬영하는 경우, 영상의 배경음에 해당하는 오디오 신호를 획득할 수 있다. 프로세서(350)는 프로세서(350)에 포함된 마이크로폰(예: 도 1의 입력 모듈(150))를 이용하여 외부 음성을 입력 받고, 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는

연결된 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 예를 들면, 외부 전자 장치(220)는 수집한 음성을 이용하여 오디오 신호를 생성할 수 있고, 프로세서(350)로 생성한 오디오 신호를 전송할 수 있다. 프로세서(350)는 외부 전자 장치(220)로부터 오디오 신호를 수신할 수 있다. 일 실시예에 따르면, 프로세서(350)는 영상을 촬영함과 동시에 외부 전자 장치(220)로부터 영상에 대응하는 음성의 오디오 신호를 수신할 수 있다. 일 실시예에 따르면, 프로세서(350)가 외부 전자 장치(220)로부터 수신한 오디오 신호는 모노(mono) 음향일 수 있다.

- [69] 동작 420을 참조하면, 프로세서(350)는 타겟 객체를 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지를 분석하여 타겟이 되는 객체(예: 타겟 객체)를 확인할 수 있다. 타겟 객체는, 예를 들면, 프로세서(350)가 실제 위치를 추정하고자 하는 객체를 의미할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지를 디스플레이(330)에 표시할 수 있다. 프로세서(350)가 표시한 동영상 이미지는 적어도 하나의 객체를 포함할 수 있다. 동영상 이미지에 포함된 적어도 하나의 객체는, 예를 들면, 인물(예: 도 2의 인물(230))이 촬영된 이미지 객체인 제1객체(예: 도 2의 제1객체(211)) 및/또는 외부 전자 장치(220)가 촬영된 이미지 객체인 제2객체(예: 도 2의 제2객체(221))를 포함할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영된 동영상 이미지 또는 촬영 중인 동영상 이미지를 분석하여 동영상 이미지에 포함된 적어도 하나의 이미지 객체(예: 도 2의 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 메모리(예: 도 3의 메모리(340))에 기 저장된 알고리즘을 이용하여 이미지를 분석할 수 있고, 이미지 분석을 통해 동영상 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 표시된 이미지를 분석하여 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다.
- [70] 다양한 실시예에 따르면, 프로세서(350)는 이미지 분석을 이용하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 동영상 이미지를 분석하여 안면 인식을 수행할 수 있고, 안면 인식 수행 결과에 기초하여 인물 객체(예: 제1객체(211))를 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 식별된 인물 객체(예: 제1객체(211))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 식별된 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로서 식별할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영된 이미지를 분석하여 시각적 신호(예: 점멸 중인 LED 신호)를 식별하여 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 타겟 객체를 확인하기 위한 조건을 메모리(예: 도 3의 메모리(340))에 미리 저장할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에

사용자(미도시)의 터치 입력을 수신할 수 있고, 수신한 터치 입력에 기초하여 타겟 객체를 확인할 수 있다. 예를 들면, 인식된 적어도 하나의 객체 중 터치 입력의 위치에 대응하는 객체를 타겟 객체로 인식할 수 있다. 일 실시예에 따르면, 프로세서(350)는 센서(예: 도 1의 센서 모듈(176))를 이용하여 센서 정보를 생성하고, 센서 정보에 기초하여 외부 전자 장치(220) 또는 인물(230)을 인식하고, 인식 결과에 기초하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)와의 통신을 이용하여 외부 전자 장치(220)의 위치에 관한 정보를 수신하고, 수신한 위치 정보를 저장할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 촬영된 동영상 이미지 분석, 수신된 터치 입력 분석, 센서 정보, 및 수신한 위치 정보 중 적어도 하나에 기초하여 타겟 객체를 식별하고 확인할 수 있다.

- [71] 동작 430을 참조하면, 프로세서(350)는 제1위치 정보 및 부가 정보를 확인할 수 있다. 제1위치 정보는, 타겟 객체가 디스플레이(210) 상에 표시되는 위치에 관한 정보일 수 있다. 일 실시예에 따르면, 제1위치 정보는 디스플레이(210) 상의 일정한 좌표로 표현되는 정보일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))가 디스플레이(210) 상에 표시되는 위치에 관한 정보(예: 좌표)(예: 제1위치 정보)를 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 식별한 각 객체(예: 제1객체(211) 및/또는 제2객체(221))의 디스플레이(210) 상 표시되는 좌표(예: 제1위치 정보)를 지속적으로 확인할 수 있다. 프로세서(350)는 촬영 중 디스플레이(210) 상에서 실시간으로 이동하는 객체(예: 제1객체(211) 및/또는 제2객체(221))의 좌표(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 제1위치 정보는 프로세서(350)의 동영상 촬영 중에 실시간으로 변경되는 정보일 수 있다. 일 실시예에 따르면, 프로세서(350)는 동영상 촬영 중에 제1위치 정보를 지속적 및 즉각적으로 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 타겟 객체가 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 존재하는 모든 객체(예: 제1객체(211) 및 제2객체(221))의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인할 수 있다. 일 실시예에 따르면, 프로세서(350)는 디스플레이(210) 상에 표시된 적어도 하나의 이미지 객체(예: 제1객체(211) 및/또는 제2객체(221)) 가운데, 타겟 객체(예: 제1객체(211) 또는 제2객체(221))의 디스플레이(210) 상에 표시되는 위치(예: 제1위치 정보)를 확인할 수 있다.

- [72] 다양한 실시예에 따르면, 프로세서(350)는 부가 정보를 확인할 있다. 부가 정보는 프로세서(350)가 타겟 객체에 대응되는 피사체(예: 외부 전자 장치(220) 또는 인물(230))의 실제 위치에 관한 정보(예: 제2위치 정보)를 추정하기 위하여 제1위치 정보 이외에 이용되는 정보일 수 있다. 부가 정보는, 전자 장치(300)에 포함된 카메라 모듈(예: 도 3의 카메라 모듈(320))에 설정된 정보를 포함할 수 있다. 다양한 실시예에 따르면, 부가 정보는 촬영 중인 프로세서(350)의

화각(field of view, FOV) 및/또는 배율과 같이, 카메라 모듈(320)의 상태 및/또는 설정에 관한 정보를 포함할 수 있다.

- [73] 동작 440을 참조하면, 프로세서(350)는 제2위치 정보를 생성할 수 있다. 제2위치 정보는, 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))의 실제 위치에 관한 정보일 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))의 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영한 동영상 이미지에 포함된 모든 객체(예: 제1객체(211) 및 제2객체(221))에 대응하는 피사체(예: 인물(230) 및 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체에 대응하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치만을 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체에 대응하는 피사체의 실제 위치를 추정하고, 추정된 위치에 관한 제2위치 정보를 생성할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 제1위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영된 이미지 객체(예: 타겟 객체)의 디스플레이(210) 상 위치(예: 제1위치 정보)를 확인하고, 디스플레이(210) 상 위치에 기초하여 실제 위치(예: 제2위치 정보)를 추정할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 부가 정보에 기초하여 제2위치 정보를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)는 타겟 객체의 제1위치 정보 및 부가 정보를 이용하여 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 실제 위치를 추정할 수 있다. 일 실시예에 따르면, 프로세서(350)가 추정하는 실제 위치는, 프로세서(350)를 기준으로 하는 피사체(예: 인물(230) 또는 외부 전자 장치(220))의 상대적인 위치를 의미할 수 있다. 일 실시예에 따르면, 프로세서(350)는 전자 장치(300)에 포함된 센서(예: 도 1의 센서 모듈(176))를 통해 생성한 센서 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 외부 전자 장치(220)로부터 외부 전자 장치(220)의 위치 정보를 수신할 수 있고, 수신한 위치 정보에 기초하여 제2위치 정보를 생성할 수 있다. 일 실시예에 따르면, 제2위치 정보는, 좌우 편향된 위치만을 표현하는 1차원적 위치 정보일 수 있고, 또는, 프로세서(350)를 기준으로 앞뒤(예: 거리)를 포함할 수 있고, 또는, 상하로 편향된 위치를 표현하여 3차원적 위치 정보일 수 있다. 다양한 실시예에 따르면, 제2위치 정보는 피사체의 전후 위치, 좌우 위치 및 상하 위치 중 적어도 하나 또는 그 조합을 포함할 수 있다.
- [74] 동작 450을 참조하면, 프로세서(350)는 제2위치 정보에 기초하여 오디오 신호를 처리할 수 있다. 오디오 신호의 처리는, 획득한 오디오 신호에 방향감을 부여하는 동작일 수 있다. 오디오 신호의 처리는, 예를 들면, 오디오 신호의 변경 및/또는 신호의 변환을 포함할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호를 패닝(panning)하여 스테레오 오디오 신호로 변환할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호를

렌더링(rendering) 하여 공간감, 위치감 및/또는 방향감을 제공하는 3차원 음향(예: 바이노럴 음향)으로 변환할 수 있다. 일 실시예에 따르면, 프로세서(350)는 획득한 오디오 신호의 음량을 조절하여 거리감을 제공하도록 오디오 신호를 처리할 수 있다. 다양한 실시예에 따르면, 프로세서(350)는 하나의 오디오 신호를 처리하여 청취자 기준 좌측 귀로 듣는 신호(좌측 오디오) 및 우측 귀로 듣는 신호(우측 오디오)를 각각 생성할 수 있다. 일 실시예에 따르면, 프로세서(350)는 좌측 오디오 및 우측 오디오에 각각 음향의 강도 차이, 시간 차이 및 위상 차이 중 적어도 하나를 생성하여 오디오 신호를 처리할 수 있다.

- [75] 일 실시예에 따르면, 프로세서(350)는 처리된 오디오 신호를 메모리(예: 도 3의 메모리(340))에 저장할 수 있다. 일 실시예에 따르면, 프로세서(350)는 촬영된 동영상 이미지를 동영상 데이터로서 저장하고, 처리된 오디오 신호와 인코딩할 수 있다. 일 실시예에 따르면, 프로세서(350)는 처리된 오디오 신호를 동영상에 대응하는 배경 오디오로써 인코딩하여 저장할 수 있다. 일 실시예에 따르면, 프로세서(350)는 생성한 제2위치 정보를 오디오 신호와 별도로 인코딩할 수 있다. 예를 들면, 프로세서(350)는 제2위치 정보와 오디오 신호를 별도로 인코딩한 뒤, 추후 재생 시에, 다시 디코딩하여 제2위치 정보에 기초한 오디오 신호 처리를 수행할 수 있다.
- [76] 도 5, 도 6 및 도 7은 다양한 실시예에 따른 전자 장치가 타겟 객체를 확인하는 예시도이다.
- [77] 도 5, 도 6 및 도 7을 참조하면, 전자 장치(200)(예: 도 1의 전자 장치(101) 및/또는 도 3의 전자 장치(300))가 적어도 하나의 피사체(예: 도 2의 인물(230) 및/또는 외부 전자 장치(220))를 촬영하여 디스플레이(210)(예: 도 3의 디스플레이(330))에 표시한 동영상 이미지의 예시일 수 있다.
- [78] 도 5를 참조하면, 전자 장치(200)는 디스플레이(210) 상에 표시된 동영상 이미지를 분석하여 타겟 객체를 확인할 수 있다.
- [79] 다양한 실시예에 따르면, 전자 장치(200)가 표시한 동영상 이미지는 적어도 하나의 객체를 포함할 수 있다. 동영상 이미지에 포함된 적어도 하나의 객체는, 예를 들면, 인물 피사체(예: 도 2의 인물(230))가 촬영된 이미지 객체인 제1객체(211) 및/또는 외부 전자 장치(예: 도 2의 외부 전자 장치(220))가 촬영된 이미지 객체인 제2객체(221)를 포함할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영된 동영상 이미지 또는 촬영 중인 동영상 이미지를 분석하여 동영상 이미지에 포함된 적어도 하나의 이미지 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 메모리(예: 도 3의 메모리(340))에 기 저장된 알고리즘을 이용하여 이미지를 분석할 수 있고, 이미지 분석을 통해 동영상 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 표시된 이미지를 분석하여 이미지에 포함된 객체(예: 제1객체(211) 및/또는 제2객체(221))를 식별할 수 있다. 다양한 실시예에

따르면, 전자 장치(200)는 촬영한 동영상 이미지를 분석하여 타겟이 되는 객체(예: 타겟 객체)를 확인할 수 있다. 타겟 객체는, 예를 들면, 전자 장치(200)가 실제 위치를 추정하고자 하는 객체를 의미할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 이미지 분석을 이용하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 동영상 이미지를 분석하여 안면 인식을 수행할 수 있고, 안면 인식 수행 결과에 기초하여 인물 객체(예: 제1객체(211))를 식별할 수 있다.

[80] 도 5의 [a]를 참조하면, 전자 장치(200)가 식별한 객체(500)는 제1객체(211)일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 동영상 이미지를 분석하여 인식한 안면을 기초로 인물 객체(예: 제1객체(211))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 인물 객체(예: 제1객체(211))를 타겟 객체로서 확인할 수 있다.

[81] 도 5의 [b]를 참조하면, 전자 장치(200)가 식별한 객체(500)는 제2객체(221)일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 동영상 이미지를 분석하여 인식한 안면을 기초로 인물 객체(예: 제2객체(221))를 식별할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 촬영된 이미지를 분석하여 시각적 신호(예: 점멸 중인 LED 신호)를 식별하여 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로 확인할 수 있다. 일 실시예에 따르면, 외부 전자 장치(220)는 시각적 신호(예: LED 점멸 신호)를 출력할 수 있고, 전자 장치(200)는 외부 전자 장치(220)의 신호를 식별하여 외부 전자 장치(220)의 이미지에 대응되는 객체를 식별할 수 있다. 실시예에 따르면, 전자 장치(200)는 식별된 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 타겟 객체로서 식별할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 타겟 객체를 확인하기 위한 조건을 메모리(예: 도 3의 메모리(340))에 미리 저장할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 센서(예: 도 1의 센서 모듈(176))를 이용하여 센서 정보를 생성하고, 센서 정보에 기초하여 외부 전자 장치(220)에 대응되는 객체(예: 제2객체(221))를 인식하고, 인식 결과에 기초하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 외부 전자 장치(220)와의 통신을 이용하여 외부 전자 장치(220)의 위치에 관한 정보를 수신하고, 수신한 위치 정보를 저장할 수 있다. 도 5의 [b]를 참조하면, 외부 전자 장치(220)는 특정한 주파수를 가진 신호(예: UWB 신호)를 지속적으로 출력하는 상태일 수 있다. 이 경우 전자 장치(200)는 외부 전자 장치(220)가 출력하는 신호(예: UWB 신호)를 센서(예: 도 1의 센서 모듈(176))를 이용하여 수신하고, 센서 정보를 생성하여 제2객체(221)를 인식할 수 있다.

[82] 도 5의 [c]를 참조하면, 전자 장치(200)는 복수의 이미지 객체들(예: 제1객체(211) 및 제2객체(221))가 인식되는 경우, 인식된 객체들 가운데 미리 설정된 조건에 따라 타겟 객체를 선택할 수 있다. 도 5의 [c]를 참조하면, 식별된 객체(예: 제1식별된 객체(501) 및 제2식별된 객체(502))가 복수개인 경우일 수

- 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 복수의 객체(501 및 502) 가운데 적어도 하나를 선택하여 타겟 객체로 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 복수의 객체들을 타겟 객체로 확인할 수 있다.
- [83] 도 6 및 도 7을 참조하면, 전자 장치(200)는 터치 입력에 기초하여 타겟 객체를 확인할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 사용자(미도시)의 터치 입력을 수신할 수 있고, 수신한 터치 입력에 기초하여 타겟 객체를 확인할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)의 디스플레이(210)는 터치 패널(예: 도 3의 입력 모듈(331))을 포함할 수 있고, 터치 패널(예: 입력 모듈(331))을 통해 사용자의 터치 입력을 수신할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 터치 입력을 수신하고, 터치 입력이 발생한 영역(예: 터치 영역(212))을 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)의 디스플레이(210) 상 좌표를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)의 넓이를 다양하게 설정할 수 있다. 도 6 및 도 7을 참조하면, 전자 장치(200)는 비교적 좁은 범위의 터치 영역(예: 도 6의 터치 영역(212)) 또는 비교적 넓은 범위의 터치 영역(예: 도 7의 터치 영역(212))을 설정할 수 있다. 전자 장치(200)는 터치 입력이 수신된 위치로부터 미리 정해진 일정 반경 이내의 영역을 터치 영역(212)으로 설정할 수 있다.
- [84] 도 6을 참조하면, 전자 장치(200)는 예를 들면, 인식된 적어도 하나의 객체 중 터치 입력의 위치에 대응하는 객체를 타겟 객체로 인식할 수 있다. 도 6의 [a] 및 [b]를 참조하면, 전자 장치(200)는 터치 입력을 수신할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 디스플레이(210) 상에 사용자의 터치 입력을 수신할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 수신한 터치 입력이 수신된 터치 영역(212)을 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)을 확인하고, 터치 영역(212)의 디스플레이(210) 상의 위치에 기초하여 타겟 객체를 확인할 수 있다. 도 6의 [a]를 참조하면, 전자 장치(200)는 디스플레이(210) 상에 촬영 중인 이미지를 표시할 수 있고, 촬영 이미지는 이미지 객체(211)를 포함할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 입력을 수신하고, 입력을 수신한 터치 영역(212)의 위치를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)의 위치에 기초하여 터치 영역(212)과 실질적으로 동일한 위치에 존재하는 이미지 객체(211)를 타겟 객체로 확인할 수 있다.
- [85] 도 6의 [b]를 참조하면, 전자 장치(200)는 복수의 이미지 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))를 포함하는 촬영 영상을 디스플레이(210) 상에 표시할 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 촬영 영상에 포함된 복수의 이미지 객체들 중 타겟이 되는 타겟 객체(예: 제1객체(211a))를 타겟 객체로 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및

제3객체(211c)) 가운데, 터치 영역(212)의 위치에 기초하여 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 이미지 분석을 통해 디스플레이(210) 상에 표시된 영상 이미지에 포함된 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))를 인식할 수 있다. 일 실시예에 따르면, 전자 장치(200)가 인식하는 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))는 인물 이미지 객체일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 영상 이미지를 분석하여 인물 이미지 객체(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))를 인식할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 인물 이미지에 포함된 얼굴 형태를 식별할 수 있고, 영상 이미지 내에 포함된 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c)) 가운데, 복수의 식별된 객체들(예: 제1식별된 객체(501), 제2식별된 객체(502) 및 제3식별된 객체(503))를 확인할 수 있다. 식별된 객체들(예: 제1식별된 객체(501), 제2식별된 객체(502) 및 제3식별된 객체(503))은, 영상 이미지 내의 복수의 객체들 가운데 전자 장치(200)가 식별한 객체일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 식별된 객체들(예: 제1식별된 객체(501), 제2식별된 객체(502) 및 제3식별된 객체(503))의 위치, 및 터치 영역(212)의 위치를 기초로 타겟 객체를 확인할 수 있다. 도 6의 [b]를 참조하면, 전자 장치(200)는 터치 영역(212)과 가장 가까운 제1식별된 객체(501)를 타겟 객체로 확인할 수 있다.

- [86] 도 7을 참조하면, 전자 장치(200)는 터치 위치로부터 일정 반경 범위의 영역을 터치 영역(212)으로 확인할 수 있다. 도 7의 [a] 및 [b]를 참조하면, 터치 영역(212)은 터치 입력이 발생한 위치를 기준으로 하여 일정한 면적을 가지는 디스플레이(210) 상의 영역일 수 있다. 다양한 실시예에 따르면, 전자 장치(200)는 터치 영역(212)을 기초로 타겟 객체를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)과 대응하는 위치에 있는 객체를 타겟 객체로 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 터치 영역(212)에 포함되거나, 중첩 또는 가장 근사한 위치에 존재하는 객체를 타겟 객체로 확인할 수 있다. 도 7의 [b]를 참조하면, 전자 장치(200)가 디스플레이(210)상에 표시하는 이미지는 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))를 포함할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 이미지 분석을 통해 디스플레이(210) 상에 복수의 객체들(예: 제1객체(211a), 제2객체(211b) 및 제3객체(211c))를 각각 식별할 수 있다. 전자 장치(200)는 식별된 객체들(예: 제1식별된 객체(501), 제2식별된 객체(502) 및 제3식별된 객체(503)) 가운데, 터치 영역(212)과 대응되는 객체(예: 제1객체(211a))를 타겟 객체로 확인할 수 있다. 도 7의 [b]를 참조하면, 전자 장치(200)는 터치 영역(212)에 제1식별된 객체(501)가 포함되는 것을 확인할 수 있고, 제1식별된 객체(501)에 대응하는 제1객체(211a)를 타겟 객체로 확인할 수 있다.

- [87] 도 8은 다양한 실시예에 따른 부가 정보를 설명하기 위한 예시도이다.

- [88] 다양한 실시예에 따르면, 전자 장치(200)는 부가 정보를 확인할 수 있다. 다양한 실시예에 따르면, 프로세서(예: 도 3의 프로세서(350))는 부가 정보를 확인할 수 있다. 부가 정보는 프로세서(350)가 타겟 객체에 대응되는 피사체(예: 외부 전자 장치(220) 또는 인물(230))의 실제 위치에 관한 정보(예: 제2위치 정보)를 추정하기 위하여 제1위치 정보 이외에 이용되는 정보일 수 있다. 부가 정보는, 전자 장치(300)에 포함된 카메라 모듈(예: 도 3의 카메라 모듈(320))에 설정된 정보를 포함할 수 있다. 다양한 실시예에 따르면, 부가 정보는 촬영 중인 카메라 모듈(320)의 화각(예: 제1각도(01))(field of view, FOV) 및/또는 배율(m)과 같이, 카메라 모듈(320)의 상태 및/또는 설정에 관한 정보를 포함할 수 있다.
- [89] 도 8을 참조하면, 전자 장치(200)는 피사체(230)를 동영상 촬영하는 중일 수 있다. 일 실시예에 따르면, 전자 장치(200)는 카메라 모듈(예: 도 3의 카메라 모듈(320))을 이용하여 동영상을 촬영할 수 있다.
- [90] 도 8의 [a]를 참조하면, 전자 장치(200)의 카메라 모듈(320)은 일정한 각도(예: 제1각도(01))의 화각을 형성할 수 있다. 일 실시예에 따르면, 카메라 모듈(320)은 적어도 하나의 렌즈(미도시)를 포함할 수 있고, 적어도 하나의 렌즈가 포함하는 구경 및/또는 배율로 인하여 촬영할 수 있는 화각(예: 제1각도(01))을 형성할 수 있다. 전자 장치(200)는 메모리(예: 도 3의 메모리(340))에 화각에 대한 정보를 미리 저장할 수 있다.
- [91] 도 8의 [b]를 참조하면, 전자 장치(200)는 현재 촬영 중인 카메라 모듈(320)에 설정된 배율(m)에 대한 정보 및 식별된 객체(500)(예: 타겟 객체)의 디스플레이 상의 위치에 대한 좌표 정보를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 디스플레이(210)의 x축 상의 길이 dx 및 y축 상의 길이 dy에서 식별된 객체(500)가 표시되는 위치를 확인할 수 있다. 일 실시예에 따르면, 전자 장치(200)는 촬영 중인 영상 이미지에 적용된 배율(m) 정보를 확인할 수 있다. 배율(m)에 관한 정보는, 전자 장치(200)가 동영상을 촬영할 때 확대한 비율에 관한 정보일 수 있다. 다양한 실시예에 따르면, 부가 정보는 화각 및/또는 배율 정보를 포함할 수 있고, 전자 장치(200)는 화각 및/또는 배율 정보를 이용하여 객체(예: 타겟 객체)가 디스플레이(210)의 중심으로부터 벗어난 각도(예: 상하 방향의 고도각 및/또는 좌우 방향의 방위각)를 계산할 수 있고, 실제 위치 정보(예: 제2위치 정보)를 생성할 수 있다.
- [92] 도 9는 다양한 실시예에 따른 스테레오 음향을 설명하기 위한 예시도이다.
- [93] 도 10 및 도 11은 다양한 실시예에 따른 공간감이 부여된 오디오 신호를 설명하기 위한 예시도이다.
- [94] 도 9의 [a] 및 [b]를 참조하면, 음향을 청취하는 청취자(90)는 음향을 청취할 수 있다. 일 실시예에 따르면, 청취자(90)는 이어폰과 같은 음원 출력 장치를 이용하여 음원을 청취하는 중일 수 있다. 다양한 실시예에 따르면, 청취자(90)는 음향을 청취하면서, 특정한 거리 및/또는 방향에 음원이 존재하는 것으로 인식할 수 있고, 해당하는 거리 및/또는 방향에 대응하는 가상의 위치에 음상(900)(sound

image 또는 acoustic image)이 형성될 수 있다. 일 실시예에 따르면, 음상(900)이 형성됨으로써 청취자(90)는 음원이 존재하는 공간(예: 음장(acoustic field))이 있는 것처럼 느끼는 효과를 제공 받을 수 있다. 예를 들면, 청취자(90)는 음상(900)이 형성됨으로써 음장이 있는 느낌, 즉, 음장감을 느낄 수 있다.

- [95] 도 9의 [a]를 참조하면, 청취자(90)는 모노(mono) 음향에 해당하는 오디오 신호를 청취하는 중일 수 있다. 모노 오디오 신호는, 좌이(左耳) 및 우이(右耳)에 동일한 음성을 출력하도록 하는 신호일 수 있다. 동일한 음성은, 동일한 위상의 음성, 동일한 음량의 음성 및/또는 동일한 시점의 음성으로 이해될 수 있다. 도 9의 [a]를 참조하면, 모노 오디오 신호의 경우 하나의 제1음상(910)만을 형성할 수 있다. 제1음상(910)은 사용자의 정면에, 사용자로부터 일정한 거리만큼 이격되어 형성될 수 있다.
- [96] 도 9의 [b], 도 10 및 도 11을 참조하면, 청취자(90)는 스테레오(stereo) 음향에 해당하는 오디오 신호를 청취하는 중일 수 있다. 스테레오 오디오 신호는, 좌이(左耳) 및 우이(右耳)에 서로 다른 음성을 출력하도록 하는 신호일 수 있다. 서로 다른 음성은, 서로 다른 위상의 음성, 서로 다른 음량의 음성 및/또는 서로 다른 시점의 음성으로 이해될 수 있다. 즉, 좌이에 도달하는 음성 및 우이에 도달하는 음성 사이에 위상차, 음량차 및/또는 시간차가 형성된 오디오 신호일 수 있다. 다양한 실시예에 따르면, 스테레오 오디오 신호의 경우 적어도 두개의 음상(예: 좌측 제2음상(920L) 및 우측 제2음상(920R))을 형성할 수 있다.
- [97] 도 9의 [b]를 참조하면, 스테레오 오디오 신호가 두개의 음상(예: 좌측 제2음상(920L) 및 우측 제2음상(920R))을 형성한 경우일 수 있다. 일 실시예에 따르면, 좌측 제2음상(920L) 및 우측 제2음상(920R)은 청취자(90)의 정면으로부터 동일한 각도(예: θ_2)만큼 이격되어 있을 수 있다. 일 실시예에 따르면, 좌측 제2음상(920L) 및 우측 제2음상(920R)은 청취자(90)로부터 실질적으로 동일한 거리만큼 이격된 위치에 형성될 수 있다. 다양한 실시예에 따르면, 전자 장치(예: 도 3의 전자 장치(300))는 스테레오 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 전자 장치(300)는 오디오 신호를 처리하여, 스테레오 오디오 신호를 생성할 수 있다. 스테레오 오디오 신호를 생성하는 경우, 전자 장치(300)는 제2위치 정보에 기초하여 스테레오 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 제2위치 정보는 타겟 객체가 형성하는 좌우 방위각 또는 좌우 거리만을 포함할 수 있다. 좌우 방위각 또는 좌우 거리는, 타겟 객체가 전자 장치(300)의 중심으로부터 좌우 방향으로 떨어진 각도 또는 거리를 의미할 수 있다. 일 실시예에 따르면, 전자 장치(300)는 타겟 객체의 좌우 방위각 또는 좌우 거리에 기초하여 오디오 신호를 처리할 수 있다. 전자 장치(300)는 좌우 방위각 또는 좌우 거리에 기초하여 좌우 음향을 서로 다르게 생성, 즉, 패닝(panning)할 수 있다.
- [98] 도 10을 참조하면, 스테레오 오디오 신호의 음상(900 및 901)은 청취자(90)를 중심으로 한 모든 방향에 형성될 수 있다. 이 경우 2보다 큰 개수의 음상(900 및

901)이 형성될 수 있다. 이 경우 스테레오 오디오 신호는 바이노럴(binaural) 신호를 형성하는 것으로 이해될 수 있다. 다양한 실시예에 따르면, 바이노럴 오디오 신호는 방향감을 형성할 수 있다. 도 10을 참조하면, 특정한 위치의 음상(예: 제4음상(901))로부터 청취자(90)의 좌이(91) 및 우이(92)까지의 거리는 각각 d_l 및 d_r 에 해당할 수 있다. 이 경우 좌이(91) 및 우이(92)에 각각 도달하는 음성의 음량은, 각각 d_l 및 d_r 에 반비례하여 결정될 수 있다. 또한, 좌이(91) 및 우이(92)로 음성이 도달하는 시간은 각각 d_l 및 d_r 에 비례하여 결정될 수 있다. 다양한 실시예에 따르면, 전자 장치(300)는 오디오 신호에서, 특정한 위치의 음상(예: 제4음상(901))을 형성하기 위하여, 설정된 좌이(91) 및 우이(92)까지의 거리에 기초하여, 좌이(91) 및 우이(92) 각각에 도달하는 음향에, 음량차 및/또는 시간차를 형성하여 오디오 신호를 처리할 수 있다. 이 경우 전자 장치(300)는 전자 장치(300)는 제2위치 정보에 기초하여 방향감이 부여된 오디오 신호를 생성할 수 있다. 일 실시예에 따르면, 제2위치 정보는 타겟 객체가 형성하는 좌우 방위각, 좌우 거리 및/또는 전자 장치(300)로부터의 거리를 포함할 수 있다. 일 실시예에 따르면, 전자 장치(300)는 제2위치 정보에 기초하여 오디오 신호에 방향감을 부여, 즉, 렌더링(rendering)할 수 있다.

[99] 도 11을 참조하면, 오디오 신호는 상하 방향감 및 거리감을 형성할 수 있다. 도 11의 [a]를 참조하면, 청취자(90)는 상하 방향의 방향감이 형성된 오디오 신호를 청취할 수 있고, 상측 및 하측에 음상(예: 상부 음상(900H) 및 하부 음상(900L))이 형성될 수 있다. 다양한 실시예에 따르면, 전자 장치(300)는 오디오 신호의 전체적인 진동수 특성을 조절(예: 이퀄라이징(equalizing))하여 고진동수 영역으로 처리하는 경우 상측에 음상(상부 음상(900H))을 형성할 수 있고, 저진동수 영역으로 처리하는 경우 하부 음상(900L)을 형성할 수 있다. 이 경우, 전자 장치(300)는 상하 위치를 포함하는 제2위치 정보에 기초하여 오디오 신호를 상하 렌더링 할 수 있다.

[100] 도 11의 [b]를 참조하면, 음상(예: 근거리 음상(900C) 및 원거리 음상(900F))은 거리감을 형성할 수 있다. 도 11의 [b]를 참조하면, 청취자(90)의 정면을 기준으로 동일한 각도(θ)만큼 이격된 두 개의 음상(예: 근거리 음상(900C) 및 원거리 음상(900F))을 예시적으로 표현한 것일 수 있다. 일 실시예에 따르면, 근거리 음상(900C)은 청취자(90)로부터 거리 d_c 만큼 이격되어 있고, 원거리 음상(900F)은 청취자(90)로부터 거리 d_f 만큼 이격되어 있을 수 있다. 일 실시예에 따르면, 전자 장치(300)는 각 거리(예: d_c 및 d_f)에 반비례하는 음량을 적용하여 좌이 및 우이에 도달하는 음성을 포함하도록 오디오 신호를 처리할 수 있다. 일 실시예에 따르면 전자 장치(300)는 거리 정보를 포함하는 제2위치 정보를 생성하고, 서로 다른 거리 정보에 기초하여 서로 다른 거리에 해당하는 음상을 포함하도록 오디오 신호를 처리할 수 있다. 다양한 실시예에 따르면, 전자 장치(300)는 음상이 형성된 거리에 따라 좌우 편향의 정도가 상이하도록 오디오 신호를 처리할 수 있다. 예를 들면, 근거리 음상(900C) 및 원거리 음상(900F)은

각각 청취자(90)를 기준으로 동일한 각도(θ3)만큼 이격되어 있더라도 거리에 따라 청취자(90)는 좌우 방향 편향의 정도를 상이하게 느낄 수 있다. 청취자(90)는 가까이 있는 근거리 음상(900C)의 좌우 편향 정도를 원거리 음상(900F)에 비하여 크게 느낄 수 있다. 일 실시예에 따르면, 전자 장치(300)는 근거리 음상과의 거리(dc) 및 원거리 음상과의 거리(df)에 반비례하여 편향 정도를 결정하고, 오디오 신호를 처리(예: 패닝)할 수 있다.

- [101] 본 문서에 개시된 다양한 실시예에 따른 전자 장치(예: 전자 장치(300))는, 근거리 무선 통신을 지원하는 통신 모듈(310), 동영상 이미지를 촬영하는 카메라 모듈(320), 상기 촬영하는 동영상 이미지를 표시하는 디스플레이(330), 및 상기 통신 모듈, 카메라 모듈 및 디스플레이와 작동적으로(operatively) 연결되는 프로세서(350)를 포함하고, 상기 프로세서는, 상기 통신 모듈을 이용하여 외부 전자 장치(예: 외부 전자 장치(220))와 연결을 수립하고, 상기 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오 신호를 수신하고, 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체(예: 제1객체(211a), 제2객체(211b) 및/또는 제3객체(211c)) 가운데 타겟이 되는 타겟 객체를 확인하고, 상기 타겟 객체가 상기 디스플레이 상에 표시되는 위치에 관한 제1위치 정보를 확인하고, 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여, 실제 위치에 관한 제2위치 정보를 생성하고, 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하도록 설정될 수 있다.
- [102] 또한, 상기 프로세서는, 상기 촬영 중인 동영상 이미지에 포함되는 적어도 하나의 객체를 인식하고, 상기 인식된 적어도 하나의 객체 가운데 타겟 객체를 확인하도록 설정될 수 있다.
- [103] 또한, 상기 디스플레이는, 터치 입력을 수신하는 입력 모듈(331)을 더 포함하고, 상기 프로세서는, 상기 터치 입력을 수신하고, 상기 수신한 터치 입력의 상기 디스플레이 상의 위치를 확인하고, 상기 확인된 터치 입력의 위치(예: 터치 영역(212))에 기초하여 상기 타겟 객체를 확인하도록 설정될 수 있다.
- [104] 또한, 상기 프로세서는, 상기 적어도 하나의 객체의 이미지를 분석하고, 상기 이미지 분석에 기초하여 상기 타겟 객체를 확인하도록 설정될 수 있다.
- [105] 또한, 상기 프로세서는, 상기 이미지 분석에 기초하여 상기 외부 전자 장치의 이미지 및 상기 적어도 하나의 객체의 얼굴 이미지 가운데 적어도 하나에 대한 이미지를 인식하고, 상기 인식된 이미지에 기초하여 상기 타겟 객체를 확인하도록 설정될 수 있다.
- [106] 또한, 상기 프로세서는, 상기 카메라 모듈의 확대 배율 정보, 상기 카메라 모듈의 화각 정보를 포함하는 부가 정보를 더 확인하고, 상기 확인된 부가 정보 및 상기 제1위치 정보에 기초하여 상기 제2위치 정보를 생성하도록 설정될 수 있다.
- [107] 또한, 상기 외부 전자 장치로부터 상기 외부 전자 장치의 위치 정보를 더 수신하고, 상기 프로세서는, 상기 외부 전자 장치의 위치 정보에 기초하여 상기

제2위치 정보를 생성하도록 설정될 수 있다.

- [108] 또한, 센서(예: 센서 모듈(176))를 더 포함하고, 상기 프로세서는, 상기 센서를 이용하여 상기 외부 전자 장치로부터 발생하는 신호를 감지하고, 상기 감지한 신호에 기초하여 상기 제2위치 정보를 생성하도록 설정될 수 있다.
- [109] 또한, 상기 제2위치 정보는 좌우 거리 또는 좌우 방위각을 포함하고, 상기 프로세서는, 상기 오디오 신호를 상기 제2위치 정보에 기초해 패닝(panning)하여 상기 오디오 신호를 처리하도록 설정될 수 있다.
- [110] 또한, 상기 제2위치 정보는 상하를 더 포함하고, 상기 프로세서는, 상기 오디오 신호를 상기 제2위치 정보에 기초하여 3차원 렌더링 하여 상기 오디오 신호를 처리하도록 설정될 수 있다.
- [111] 또한, 상기 제2위치 정보는 상기 타겟 객체의 상기 전자 장치와의 거리를 더 포함하고, 상기 프로세서는, 상기 오디오 신호를, 상기 실제 위치에 기초하여 상기 오디오 신호의 음량을 조절 하도록 설정될 수 있다.
- [112] 또한, 데이터를 저장하고, 상기 프로세서와 작동적으로 연결된 메모리(예: 메모리(340))를 더 포함하고, 상기 프로세서는, 상기 처리된 오디오 신호 및 상기 촬영된 동영상 이미지를 인코딩하여 상기 메모리에 저장하도록 설정될 수 있다.
- [113] 본 문서에 개시된 다양한 실시예에 따른 전자 장치(예: 전자 장치(300))가 오디오 신호를 처리하는 방법에 있어서, 외부 전자 장치(예: 외부 전자 장치(220))와 연결을 수립하는 동작, 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오 신호를 수신하는 동작, 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체(예: 제1객체(211a), 제2객체(211b) 및/또는 제3객체(211c)) 가운데 타겟이 되는 타겟 객체를 확인하는 동작, 상기 타겟 객체가 상기 전자 장치의 디스플레이 상에 표시되는 위치에 관한 제1위치 정보를 확인하는 동작, 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여, 실제 위치에 관한 제2위치 정보를 생성하는 동작, 및 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하는 동작을 포함할 수 있다.
- [114] 또한, 상기 타겟 객체를 확인하는 동작은, 상기 촬영 중인 동영상 이미지에 포함되는 적어도 하나의 객체를 인식하는 동작, 및 상기 인식된 적어도 하나의 객체 가운데 타겟 객체를 확인하는 동작을 포함할 수 있다.
- [115] 또한, 상기 타겟 객체를 확인하는 동작은, 터치 입력을 수신하는 동작, 상기 수신한 터치 입력의 상기 디스플레이 상의 위치를 확인하는 동작, 및 상기 확인된 터치 입력의 위치에 기초하여 상기 타겟 객체를 확인하는 동작을 포함할 수 있다.
- [116] 또한, 상기 타겟 객체를 확인하는 동작은, 상기 적어도 하나의 객체의 이미지를 분석하는 동작, 및 상기 이미지 분석에 기초하여 상기 타겟 객체를 확인하는 동작을 포함할 수 있다.
- [117] 또한, 상기 제2위치 정보를 생성하는 동작은, 상기 전자 장치에 포함된 카메라

- 모듈(예: 카메라 모듈(320))의 확대 배율 정보, 상기 카메라 모듈의 화각 정보를 포함하는 부가 정보를 더 확인하는 동작, 및 상기 확인된 부가 정보 및 상기 제1위치 정보에 기초하여 상기 제2위치 정보를 생성하는 동작을 포함할 수 있다.
- [118] 또한, 상기 제2위치 정보는 좌우 거리 또는 좌우 방위각을 포함하고, 상기 오디오 신호를 처리하는 동작은, 상기 오디오 신호를 상기 제2위치 정보에 기초해 패닝(panning)하여 상기 오디오 신호를 처리하는 동작을 포함할 수 있다.
- [119] 또한, 상기 제2위치 정보는 높이를 더 포함하고, 상기 오디오 신호를 처리하는 동작은, 상기 오디오 신호를 상기 제2위치 정보에 기초하여 3차원 렌더링 하여 상기 오디오 신호를 처리하는 동작을 포함할 수 있다.
- [120] 또한, 상기 제2위치 정보는 상기 타겟 객체의 상기 전자 장치와의 거리를 더 포함하고, 상기 오디오 신호를 처리하는 동작은, 상기 제2위치 정보에 기초하여 상기 오디오 신호의 음량을 조절하는 동작을 포함할 수 있다.
- [121] 본 문서에 개시된 다양한 실시예들에 따른 전자 장치는 다양한 형태의 장치가 될 수 있다. 전자 장치는, 예를 들면, 휴대용 통신 장치(예: 스마트폰), 컴퓨터 장치, 휴대용 멀티미디어 장치, 휴대용 의료 기기, 카메라, 웨어러블 장치, 또는 가전 장치를 포함할 수 있다. 본 문서의 실시예에 따른 전자 장치는 전술한 기기들에 한정되지 않는다.
- [122] 본 문서의 다양한 실시예들 및 이에 사용된 용어들은 본 문서에 기재된 기술적 특징들을 특정한 실시예들로 한정하려는 것이 아니며, 해당 실시예의 다양한 변경, 균등물, 또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 또는 관련된 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 아이템에 대응하는 명사의 단수 형은 관련된 문맥상 명백하게 다르게 지시하지 않는 한, 상기 아이템 한 개 또는 복수 개를 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다. "제 1", "제 2", 또는 "첫째" 또는 "둘째"와 같은 용어들은 단순히 해당 구성요소를 다른 해당 구성요소와 구분하기 위해 사용될 수 있으며, 해당 구성요소들을 다른 측면(예: 중요성 또는 순서)에서 한정하지 않는다. 어떤(예: 제 1) 구성요소가 다른(예: 제 2) 구성요소에, "기능적으로" 또는 "통신적으로"라는 용어와 함께 또는 이런 용어 없이, "커플드" 또는 "커넥티드"라고 언급된 경우, 그것은 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로(예: 유선으로), 무선으로, 또는 제 3 구성요소를 통하여 연결될 수 있다는 것을 의미한다.
- [123] 본 문서의 다양한 실시예들에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구현된 유닛을 포함할 수 있으며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로와 같은 용어와 상호 호환적으로 사용될 수 있다. 모듈은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는, 상기 부품의 최소 단위

또는 그 일부가 될 수 있다. 예를 들면, 일실시에에 따르면, 모듈은 ASIC(application-specific integrated circuit)의 형태로 구현될 수 있다.

- [124] 본 문서의 다양한 실시예들은 기기(machine)(예: 전자 장치(101)) 의해 읽을 수 있는 저장 매체(storage medium)(예: 내장 메모리(136) 또는 외장 메모리(138))에 저장된 하나 이상의 명령어들을 포함하는 소프트웨어(예: 프로그램(140))로서 구현될 수 있다. 예를 들면, 기기(예: 전자 장치(101))의 프로세서(예: 프로세서(120))는, 저장 매체로부터 저장된 하나 이상의 명령어들 중 적어도 하나의 명령을 호출하고, 그것을 실행할 수 있다. 이것은 기기가 상기 호출된 적어도 하나의 명령어에 따라 적어도 하나의 기능을 수행하도록 운영되는 것을 가능하게 한다. 상기 하나 이상의 명령어들은 컴파일러에 의해 생성된 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장 매체는, 비일시적(non-transitory) 저장 매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장 매체가 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장 매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다.
- [125] 일실시에에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory(CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어™)를 통해 또는 두 개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.
- [126] 다양한 실시예들에 따르면, 상기 기술한 구성요소들의 각각의 구성요소(예: 모듈 또는 프로그램)는 단수 또는 복수의 개체를 포함할 수 있으며, 복수의 개체 중 일부는 다른 구성요소에 분리 배치될 수도 있다. 다양한 실시예들에 따르면, 전술한 해당 구성요소들 중 하나 이상의 구성요소들 또는 동작들이 생략되거나, 또는 하나 이상의 다른 구성요소들 또는 동작들이 추가될 수 있다. 대체적으로 또는 추가적으로, 복수의 구성요소들(예: 모듈 또는 프로그램)은 하나의 구성요소로 통합될 수 있다. 이런 경우, 통합된 구성요소는 상기 복수의 구성요소들 각각의 구성요소의 하나 이상의 기능들을 상기 통합 이전에 상기 복수의 구성요소들 중 해당 구성요소에 의해 수행되는 것과 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따르면, 모듈, 프로그램 또는 다른 구성요소에 의해 수행되는 동작들은 순차적으로, 병렬적으로, 반복적으로, 또는 휴리스틱하게 실행되거나, 상기 동작들 중 하나 이상이 다른 순서로 실행되거나,

생략되거나, 또는 하나 이상의 다른 동작들이 추가될 수 있다.

청구범위

- [청구항 1] 전자 장치에 있어서,
 근거리 무선 통신을 지원하는 통신 모듈;
 동영상 이미지를 촬영하는 카메라 모듈;
 상기 촬영하는 동영상 이미지를 표시하는 디스플레이; 및
 상기 통신 모듈, 카메라 모듈 및 디스플레이와 작동적으로(operatively)
 연결되는 프로세서를 포함하고,
 상기 프로세서는,
 상기 통신 모듈을 이용하여 외부 전자 장치와 연결을 수립하고,
 상기 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오
 신호를 수신하고,
 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체 가운데
 타겟이 되는 타겟 객체를 확인하고,
 상기 타겟 객체가 상기 디스플레이 상에 표시되는 위치에 관한 제1위치
 정보를 확인하고,
 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여,
 실제 위치에 관한 제2위치 정보를 생성하고,
 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하도록
 설정된 전자 장치.
- [청구항 2] 제1항에 있어서,
 상기 프로세서는,
 상기 촬영 중인 동영상 이미지에 포함되는 적어도 하나의 객체를
 인식하고,
 상기 인식된 적어도 하나의 객체 가운데 타겟 객체를 확인하도록 설정된
 전자 장치.
- [청구항 3] 제1항에 있어서,
 상기 디스플레이는, 터치 입력을 수신하는 입력 모듈을 더 포함하고,
 상기 프로세서는,
 상기 터치 입력을 수신하고,
 상기 수신한 터치 입력의 상기 디스플레이 상의 위치를 확인하고,
 상기 확인된 터치 입력의 위치에 기초하여 상기 타겟 객체를 확인하도록
 설정된 전자 장치.
- [청구항 4] 제1항에 있어서,
 상기 프로세서는,
 상기 적어도 하나의 객체의 이미지를 분석하고,
 상기 이미지 분석에 기초하여 상기 타겟 객체를 확인하도록 설정된 전자
 장치.

- [청구항 5] 제4항에 있어서,
 상기 프로세서는,
 상기 이미지 분석에 기초하여 상기 외부 전자 장치의 이미지 및 상기 적어도 하나의 객체의 얼굴 이미지 가운데 적어도 하나에 대한 이미지를 인식하고,
 상기 인식된 이미지에 기초하여 상기 타겟 객체를 확인하도록 설정된 전자 장치.
- [청구항 6] 제1항에 있어서,
 상기 프로세서는,
 상기 카메라 모듈의 확대 배율 정보, 상기 카메라 모듈의 화각 정보를 포함하는 부가 정보를 더 확인하고,
 상기 확인된 부가 정보 및 상기 제1위치 정보에 기초하여 상기 제2위치 정보를 생성하도록 설정된 전자 장치.
- [청구항 7] 제1항에 있어서,
 상기 외부 전자 장치로부터 상기 외부 전자 장치의 위치 정보를 더 수신하고,
 상기 프로세서는,
 상기 외부 전자 장치의 위치 정보에 기초하여 상기 제2위치 정보를 생성하도록 설정된 전자 장치.
- [청구항 8] 제1항에 있어서,
 센서를 더 포함하고,
 상기 프로세서는,
 상기 센서를 이용하여 상기 외부 전자 장치로부터 발생하는 신호를 감지하고,
 상기 감지한 신호에 기초하여 상기 제2위치 정보를 생성하도록 설정된 전자 장치.
- [청구항 9] 제1항에 있어서,
 상기 제2위치 정보는 좌우 거리 또는 좌우 방위각을 포함하고,
 상기 프로세서는,
 상기 오디오 신호를 상기 제2위치 정보에 기초해 패닝(panning)하여 상기 오디오 신호를 처리하도록 설정된 전자 장치.
- [청구항 10] 제9항에 있어서,
 상기 제2위치 정보는 상하를 더 포함하고,
 상기 프로세서는,
 상기 오디오 신호를 상기 제2위치 정보에 기초하여 3차원 렌더링 하여 상기 오디오 신호를 처리하도록 설정된 전자 장치.
- [청구항 11] 제9항에 있어서,
 상기 제2위치 정보는 상기 타겟 객체의 상기 전자 장치와의 거리를 더

포함하고,
 상기 프로세서는,
 상기 오디오 신호를,
 상기 실제 위치에 기초하여 상기 오디오 신호의 음량을 조절 하도록
 설정된 전자 장치.

- [청구항 12] 제1항에 있어서,
 데이터를 저장하고, 상기 프로세서와 작동적으로 연결된 메모리를 더
 포함하고,
 상기 프로세서는,
 상기 처리된 오디오 신호 및 상기 촬영된 동영상 이미지를 인코딩하여
 상기 메모리에 저장하도록 설정된 전자 장치.
- [청구항 13] 전자 장치가 오디오 신호를 처리하는 방법에 있어서,
 외부 전자 장치와 연결을 수립하는 동작;
 동영상 이미지 촬영과 동시에 상기 외부 전자 장치로부터 오디오 신호를
 수신하는 동작;
 상기 촬영되는 동영상 이미지에 포함되는 적어도 하나의 객체 가운데
 타겟이 되는 타겟 객체를 확인하는 동작;
 상기 타겟 객체가 상기 전자 장치의 디스플레이 상에 표시되는 위치에
 관한 제1위치 정보를 확인하는 동작;
 상기 제1위치 정보에 기초하여 상기 타겟 객체의 실제 위치를 추정하여,
 실제 위치에 관한 제2위치 정보를 생성하는 동작; 및
 상기 생성된 제2위치 정보에 기초하여 상기 오디오 신호를 처리하는
 동작을 포함하는 방법.
- [청구항 14] 제13항에 있어서,
 상기 타겟 객체를 확인하는 동작은,
 상기 촬영 중인 동영상 이미지에 포함되는 적어도 하나의 객체를
 인식하는 동작; 및
 상기 인식된 적어도 하나의 객체 가운데 타겟 객체를 확인하는 동작을
 포함하는 방법.
- [청구항 15] 제13항에 있어서,
 상기 타겟 객체를 확인하는 동작은,
 터치 입력을 수신하는 동작;
 상기 수신한 터치 입력의 상기 디스플레이 상의 위치를 확인하는 동작; 및
 상기 확인된 터치 입력의 위치에 기초하여 상기 타겟 객체를 확인하는
 동작을 포함하는 방법.