



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년01월07일  
(11) 등록번호 10-2349242  
(24) 등록일자 2022년01월05일

(51) 국제특허분류(Int. Cl.)  
G06T 7/20 (2017.01) G06K 9/00 (2006.01)  
G06K 9/32 (2006.01) G11B 27/02 (2006.01)  
H04N 5/232 (2006.01)  
(52) CPC특허분류  
G06T 7/20 (2013.01)  
G06K 9/00765 (2013.01)  
(21) 출원번호 10-2021-7003332(분할)  
(22) 출원일자(국제) 2014년05월30일  
심사청구일자 2021년03월03일  
(85) 번역문제출일자 2021년02월02일  
(65) 공개번호 10-2021-0014774  
(43) 공개일자 2021년02월09일  
(62) 원출원 특허 10-2016-7000795  
원출원일자(국제) 2014년05월30일  
심사청구일자 2019년05월14일  
(86) 국제출원번호 PCT/US2014/040166  
(87) 국제공개번호 WO 2014/200714  
국제공개일자 2014년12월18일  
(30) 우선권주장  
61/835,414 2013년06월14일 미국(US)  
14/036,947 2013년09월25일 미국(US)  
(56) 선행기술조사문헌  
JP2002518723 A\*  
JP2010114801 A\*  
KR1020110128529 A\*  
JP5090474 B2\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
켈컴 인코포레이티드  
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775  
(72) 발명자  
종 신  
미국 92121 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
가오 다산  
미국 92121 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
(뒷면에 계속)  
(74) 대리인  
특허법인코리어나

전체 청구항 수 : 총 41 항

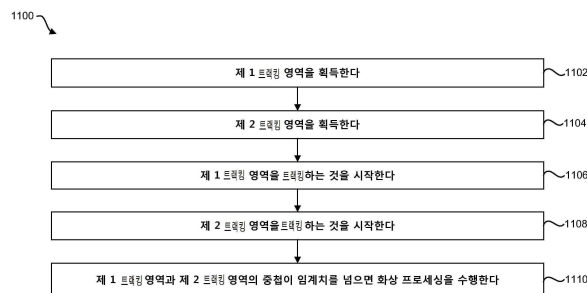
심사관 : 박금옥

(54) 발명의 명칭 트랙커 보조 이미지 캡처

(57) 요약

화상 프로세싱을 위한 방법이 설명된다. 제 1 트래킹 영역이 획득된다. 제 2 트래킹 영역이 또한 획득된다. 방법은 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하는 단계를 포함한다. 화상 프로세싱은 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치를 통과하면 수행된다.

대표도



(52) CPC특허분류

*G06K 9/3233* (2013.01)

*G11B 27/02* (2021.01)

*H04N 5/23222* (2018.08)

*H04N 5/23293* (2021.08)

(72) 발명자

**비 닝**

미국 92121 캘리포니아주 샌디에고 모어하우스 드  
라이브 5775

**레이버 스티븐 더글라스**

미국 92121 캘리포니아주 샌디에고 모어하우스 드  
라이브 5775

**베이스 스코트 데이비드**

미국 92121 캘리포니아주 샌디에고 모어하우스 드  
라이브 5775

**맥코맥 디어워드**

미국 92121 캘리포니아주 샌디에고 모어하우스 드  
라이브 5775

## 명세서

### 청구범위

#### 청구항 1

화상 프로세싱을 위한 방법으로서,

제 1 트래킹 영역의 사용자 선택을 수신하는 단계;

제 2 트래킹 영역의 사용자 선택을 수신하는 단계;

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하는 단계;

제 1 프레임에서, 상기 제 1 프레임에서 캡처된 제 1 트래킹된 오브젝트의 상기 제 1 트래킹 영역과 상기 제 1 프레임에서 캡처된 제 2 트래킹된 오브젝트의 상기 제 2 트래킹 영역 사이의 중첩의 양을 결정하는 단계로서, 상기 중첩의 양을 결정하는 단계는 상기 제 1 트래킹 영역이 상기 제 2 트래킹 영역보다 더 큰지 여부에 기초하는, 상기 중첩의 양을 결정하는 단계; 및

상기 중첩의 양이 임계치를 통과하는 것을 결정하는 것에 응답하여 화상 프로세싱을 수행하는 단계를 포함하는, 화상 프로세싱을 위한 방법.

#### 청구항 2

제 1 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 커지게 되면 수행되는, 화상 프로세싱을 위한 방법.

#### 청구항 3

제 1 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 작아지게 되면 수행되는, 화상 프로세싱을 위한 방법.

#### 청구항 4

제 1 항에 있어서,

상기 화상 프로세싱은 사진을 캡처하는 단계를 포함하고, 상기 사진은 사전에 레코딩된 비디오 영상으로부터 캡처되는, 화상 프로세싱을 위한 방법.

#### 청구항 5

제 1 항에 있어서,

상기 화상 프로세싱은 사진을 캡처하는 단계를 포함하고, 상기 사진은 라이브 영상으로부터 캡처되는, 화상 프로세싱을 위한 방법.

#### 청구항 6

제 1 항에 있어서,

상기 화상 프로세싱은 비디오 시퀀스를 편집하는 단계를 포함하고, 상기 제 1 트래킹된 오브젝트는 상기 비디오 시퀀스에서 제거되는, 화상 프로세싱을 위한 방법.

#### 청구항 7

제 6 항에 있어서,

상기 비디오 시퀀스의 제 2 프레임을 선택하는 단계로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않는, 상기 제 2 프레임을 선택하는 단계; 및

상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하는 단계를 추가적으로 포함하는, 화상 프로세싱을 위한 방법.

#### 청구항 8

제 7 항에 있어서,

상기 제 2 프레임은 상기 제 1 프레임보다 시간상 나중에 나오는, 화상 프로세싱을 위한 방법.

#### 청구항 9

제 7 항에 있어서,

상기 제 2 프레임은 상기 제 1 프레임보다 시간상 일찍 나오는, 화상 프로세싱을 위한 방법.

#### 청구항 10

제 7 항에 있어서,

상기 제 1 프레임을 편집된 비디오 시퀀스의 일부분으로서 저장하는 단계를 추가적으로 포함하는, 화상 프로세싱을 위한 방법.

#### 청구항 11

제 1 항에 있어서,

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역의 사용자 선택은 포커스 링을 통해 사용자에게 의해 입력되는, 화상 프로세싱을 위한 방법.

#### 청구항 12

제 11 항에 있어서,

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하는 것은 상기 사용자가 터치스크린으로부터 손가락을 릴리프한 후에 일어나는, 화상 프로세싱을 위한 방법.

#### 청구항 13

제 1 항에 있어서,

상기 제 2 트래킹 영역은 액션 라인을 포함하는, 화상 프로세싱을 위한 방법.

#### 청구항 14

화상 프로세싱을 위해 구성된 전자 디바이스로서,

프로세서;

상기 프로세서와 전자 통신 상태에 있는 메모리; 및

메모리에 저장된 명령들을 포함하고,

상기 명령들은,

제 1 트래킹 영역의 사용자 선택을 수신하고;

제 2 트래킹 영역의 사용자 선택을 수신하고;

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하고;

제 1 프레임에서, 상기 제 1 프레임에서 캡처된 제 1 트래킹된 오브젝트의 상기 제 1 트래킹 영역과 상기 제 1 프레임에서 캡처된 제 2 트래킹된 오브젝트의 상기 제 2 트래킹 영역 사이의 중첩의 양을 결정하는 것으로서, 상기 중첩의 양을 결정하는 것은 상기 제 1 트래킹 영역이 상기 제 2 트래킹 영역보다 더 큰지 여부에 기초하는, 상기 중첩의 양을 결정하고; 그리고

상기 중첩의 양이 임계치를 통과하는 것을 결정하는 것에 응답하여 화상 프로세싱을 수행하는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 15

제 14 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 커지게 되면 수행되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 16

제 14 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 작아지게 되면 수행되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 17

제 14 항에 있어서,

상기 화상 프로세싱은 사진을 캡처하는 것을 포함하고, 상기 사진은 사전에 레코딩된 비디오 영상으로부터 캡처되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 18

제 14 항에 있어서,

상기 화상 프로세싱은 사진을 캡처하는 것을 포함하고, 상기 사진은 라이브 영상으로부터 캡처되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 19

제 14 항에 있어서,

상기 화상 프로세싱은 비디오 시퀀스를 편집하는 것을 포함하고, 상기 제 1 트래킹된 오브젝트는 상기 비디오 시퀀스에서 제거되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 20

제 19 항에 있어서,

상기 명령들은,

상기 비디오 시퀀스의 제 2 프레임을 선택하는 것으로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않고; 그리고

상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하도록 더 실행가능한, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 21

제 20 항에 있어서,

상기 제 2 프레임은 상기 제 1 프레임보다 시간상 나중에 나오는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 22

제 20 항에 있어서,

상기 제 2 프레임은 상기 제 1 프레임보다 시간상 일찍 나오는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 23

제 20 항에 있어서,

상기 명령들은 상기 제 1 프레임을 편집된 비디오 시퀀스의 일부분으로서 저장하도록 더 실행가능한, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 24

제 14 항에 있어서,

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역의 사용자 선택은 포커스 링을 통해 사용자에게 의해 입력되는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 25

제 24 항에 있어서,

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하는 것은 상기 사용자가 터치스크린으로부터 손가락을 릴리즈한 후에 일어나는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 26

제 14 항에 있어서,

상기 제 2 트래킹 영역은 액션 라인을 포함하는, 화상 프로세싱을 위해 구성된 전자 디바이스.

#### 청구항 27

화상 프로세싱을 위한 장치로서,

제 1 트래킹 영역의 사용자 선택을 수신하기 위한 수단;

제 2 트래킹 영역의 사용자 선택을 수신하기 위한 수단;

상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하기 위한 수단;

제 1 프레임에서, 상기 제 1 프레임에서 캡처된 제 1 트래킹된 오브젝트의 상기 제 1 트래킹 영역과 상기 제 1 프레임에서 캡처된 제 2 트래킹된 오브젝트의 상기 제 2 트래킹 영역 사이의 중첩의 양을 결정하기 위한 수단으로서, 상기 중첩의 양을 결정하는 단계는 상기 제 1 트래킹 영역이 상기 제 2 트래킹 영역보다 더 큰지 여부에 기초하는, 상기 중첩의 양을 결정하기 위한 수단; 및

상기 중첩의 양이 임계치를 통과하는 것을 결정하는 것에 응답하여 화상 프로세싱을 수행하기 위한 수단을 포함하는, 화상 프로세싱을 위한 장치.

#### 청구항 28

제 27 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 커지게 되면 수행되는, 화상 프로세싱을 위한 장치.

#### 청구항 29

제 27 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 작아지게 되면 수행되는, 화상 프로세싱을 위한 장치.

#### 청구항 30

제 27 항에 있어서,

상기 화상 프로세싱은 비디오 시퀀스를 편집하는 것을 포함하고, 상기 제 1 트래킹된 오브젝트는 상기 비디오 시퀀스에서 제거되는, 화상 프로세싱을 위한 장치.

#### 청구항 31

제 30 항에 있어서,

상기 비디오 시퀀스의 제 2 프레임을 선택하는 수단으로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않는, 상기 제 2 프레임을 선택하는 수단; 및

상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하는 수단을 추가적으로 포함하는, 화상 프로세싱을 위한 장치.

### 청구항 32

명령들을 갖는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체로서,

상기 명령들은,

전자 디바이스로 하여금 제 1 트래킹 영역의 사용자 선택을 수신하게 하는 코드;

상기 전자 디바이스로 하여금 제 2 트래킹 영역의 사용자 선택을 수신하게 하는 코드;

상기 전자 디바이스로 하여금 상기 제 1 트래킹 영역 및 상기 제 2 트래킹 영역을 트래킹하기 시작하게 하는 코드;

상기 전자 디바이스로 하여금, 제 1 프레임에서, 상기 제 1 프레임에서 캡처된 제 1 트래킹된 오브젝트의 상기 제 1 트래킹 영역과 상기 제 1 프레임에서 캡처된 제 2 트래킹된 오브젝트의 상기 제 2 트래킹 영역 사이의 중첩의 양을 결정하게 하는 코드로서, 상기 중첩의 양을 결정하는 단계는 상기 제 1 트래킹 영역이 상기 제 2 트래킹 영역보다 더 큰지 여부에 기초하는, 상기 중첩의 양을 결정하게 하는 코드; 및

상기 전자 디바이스로 하여금, 상기 중첩의 양이 임계치를 통과하는 것을 결정하는 것에 응답하여 화상 프로세싱을 수행하게 하는 코드를 포함하는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

### 청구항 33

제 32 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 커지게 되면 수행되는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

### 청구항 34

제 32 항에 있어서,

상기 화상 프로세싱은 상기 중첩의 양이 상기 임계치보다 작아지게 되면 수행되는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

### 청구항 35

제 32 항에 있어서,

상기 화상 프로세싱은 비디오 시퀀스를 편집하는 것을 포함하고, 상기 제 1 트래킹된 오브젝트는 상기 비디오 시퀀스에서 제거되는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

### 청구항 36

제 32 항에 있어서,

상기 명령들은,

상기 전자 디바이스로 하여금, 비디오 시퀀스의 제 2 프레임을 선택하게 하는 코드로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않는, 상기 제 2 프레임을 선택하게 하는 코드; 및

상기 전자 디바이스로 하여금, 상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하게 하는 코드를 추가적으로 포함하는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

### 청구항 37

제 1 항에 있어서,

상기 제 1 트래킹 영역의 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 단계로서, 상기 제 1 사용자-인터페이스 엘리먼트는 트래킹의 상태에 따라 모습을 변화시키는, 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 단계를 추가적으로 포함하는, 화상 프로세싱을 위한 방법.

### 청구항 38

제 1 항에 있어서,

상기 제 1 트래킹 영역의 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 단계; 및

상기 제 2 트래킹 영역의 제 2 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 단계를 추가적으로 포함하는, 화상 프로세싱을 위한 방법.

### 청구항 39

제 14 항에 있어서,

상기 명령들은,

상기 제 1 트래킹 영역의 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하고; 그리고

상기 제 2 트래킹 영역의 제 2 사용자-인터페이스 엘리먼트 대표를 디스플레이하도록 더 실행가능한, 화상 프로세싱을 위해 구성된 전자 디바이스.

### 청구항 40

제 27 항에 있어서,

상기 제 1 트래킹 영역의 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 수단; 및

상기 제 2 트래킹 영역의 제 2 사용자-인터페이스 엘리먼트 대표를 디스플레이하는 수단을 추가적으로 포함하는, 화상 프로세싱을 위한 장치.

### 청구항 41

제 32 항에 있어서,

상기 전자 디바이스로 하여금, 상기 제 1 트래킹 영역의 제 1 사용자-인터페이스 엘리먼트 대표를 디스플레이하게 하는 코드; 및

상기 전자 디바이스로 하여금, 상기 제 2 트래킹 영역의 제 2 사용자-인터페이스 엘리먼트 대표를 디스플레이하게 하는 코드를 추가적으로 포함하는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

## 발명의 설명

## 기술 분야

[0001] 관련 출원들에 대한 상호 참조

[0002] 본 출원은 "TRACKER ASSISTED IMAGE CAPTURE" 로 2013 년 6 월 14 일에 출원된 미국 가출원 제 61/835,414 호 와 관계되고 그의 우선권을 주장한다.

[0003] 기술분야

[0004] 본 개시물은 일반적으로 전자 디바이스들에 관한 것이다. 좀더 구체적으로, 본 개시물은 트랙커 보조 이미지 캡처에 대한 시스템들 및 방법들에 관한 것이다.

## 배경 기술

[0005] 지난 수십 년간, 전자 디바이스들의 이용이 흔해졌다. 특히, 전자 기술의 진보들은 점점 더 복잡하고 유용



한 전자 디바이스들의 비용을 감소시켰다. 비용 감소 및 소비자 요구로 전자 디바이스들의 이용을 급증해, 전자 디바이스들은 현대 사회에서 사실상 어디에나 있다. 전자 디바이스들의 이용이 확대됨에 따라, 전자 디바이스들의 새롭고 향상된 피쳐들에 대한 요구도 확대되었다. 좀더 구체적으로, 새로운 기능들을 수행하고 하거나, 보다 빠르게, 보다 효율적으로, 또는 보다 높은 품질로 기능들을 수행하는 전자 디바이스들은 보통 많은 사람들이 원하는다.

[0006] 일부 전자 디바이스들 (예를 들어, 카메라들, 비디오 캠코더들, 디지털 카메라들, 셀룰러 폰들, 스마트 폰들, 컴퓨터들, 텔레비전들 등) 은 이미지들을 캡처하거나 활용한다. 예를 들어, 디지털 카메라는 디지털 이미지를 캡처할 수도 있다.

[0007] 전자 디바이스들의 새로운 및/또는 향상된 피쳐들은 보통 많은 사람들이 원하는다. 본 논의에서 관찰될 수 있는 바와 같이, 전자 디바이스들의 새로운 및/또는 향상된 특징들을 추가하는 시스템들 및 방법들은 이로울 수도 있다.

## 발명의 내용

### 해결하려는 과제

#### 과제의 해결 수단

[0008] 화상 프로세싱을 위한 방법이 설명된다. 제 1 트래킹 영역이 획득된다. 제 2 트래킹 영역이 또한 획득된다. 방법은 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하는 단계를 포함한다. 화상 프로세싱은 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치를 통과하면 수행된다.

[0009] 화상 프로세싱은 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치보다 커지게 되면 수행될 수도 있다. 화상 프로세싱은 또한 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치보다 작아지게 되면 수행될 수도 있다. 화상 프로세싱은 사진을 캡처하는 것을 포함할 수도 있다. 사진은 사전에 레코딩된 비디오 영상으로부터 또는 영상으로부터 캡처될 수도 있다. 화상 프로세싱은 또한 비디오 시퀀스를 편집하는 것을 포함할 수도 있다. 제 1 트래킹 영역에 의해 트래킹된 오브젝트는 비디오 시퀀스로부터 제거될 수도 있다.

[0010] 제 1 트래킹 영역이 비디오 시퀀스의 제 1 프레임에서 임계치보다 더 많이 제 2 트래킹 영역과 중첩된다고 결정될 수도 있다. 비디오 시퀀스의 제 2 프레임이 선택될 수도 있다. 제 1 트래킹 영역은 제 2 프레임에서 제 2 트래킹 영역과 중첩되지 않을 수도 있다. 제 1 프레임에서의 제 1 트래킹 영역은 제 2 프레임으로부터의 대응하는 대체 영역으로 대체될 수도 있다.

[0011] 제 2 프레임은 제 1 프레임보다 시간상 나중에 나올 수도 있다. 제 2 프레임은 또한 제 1 프레임보다 시간상 일찍 나올 수도 있다. 편집된 제 1 프레임은 편집된 비디오 시퀀스의 일부분으로서 저장될 수도 있다. 제 1 트래킹 영역 및 제 2 트래킹 영역은 포커스 링을 통해 사용자에게 의해 입력될 수도 있다. 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하는 것은 사용자가 터치스크린으로부터 손가락을 톨리스한 후에 일어날 수도 있다. 제 2 트래킹 영역은 액션 라인을 포함할 수도 있다.

[0012] 화상 프로세싱을 위해 구성된 전자 디바이스가 또한 설명된다. 전자 디바이스는 프로세서, 프로세서와 전자 통신 상태에 있는 메모리, 및 메모리에 저장된 명령들을 포함한다. 명령들은 제 1 트래킹 영역을 획득하도록 실행가능하다. 명령들은 또한 제 2 트래킹 영역을 획득하도록 실행가능하다. 명령들은 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하도록 더 실행가능하다. 명령들은 또한 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치를 통과하면 화상 프로세싱을 수행하도록 실행가능하다.

[0013] 화상 프로세싱을 위한 장치가 설명된다. 장치는 제 1 트래킹 영역을 획득하는 수단을 포함한다. 장치는 또한 제 2 트래킹 영역을 획득하는 수단을 포함한다. 장치는 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하는 수단을 더 포함한다. 장치는 또한 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치를 통과하면 화상 프로세싱을 수행하는 수단을 포함한다.

[0014] 화상 프로세싱을 위한 컴퓨터-프로그램 제품이 또한 설명된다. 컴퓨터-프로그램 제품은 명령들을 갖는 비일시적 컴퓨터-판독가능 매체를 포함한다. 명령들은 전자 디바이스로 하여금 제 1 트래킹 영역을 획득하게 하

는 코드를 포함한다. 명령들은 또한 전자 디바이스로 하여금 제 2 트래킹 영역을 획득하게 하는 코드를 포함한다. 명령들은 전자 디바이스로 하여금 제 1 트래킹 영역 및 제 2 트래킹 영역을 트래킹하기 시작하게 하는 코드를 더 포함한다. 명령들은 또한 전자 디바이스로 하여금 제 2 트래킹 영역과 중첩되는 제 1 트래킹 영역의 일부분이 임계치를 통과하면 화상 프로세싱을 수행하게 하는 코드를 포함한다.

### 도면의 간단한 설명

[0015]

- 도 1 은 본 시스템들 및 방법들에서의 이용을 위한 전자 디바이스를 도시하는 블록도이다;
- 도 2a 는 오브젝트 트래킹 및 검출 모듈을 도시하는 블록도이다;
- 도 2b 는 프로세서에 의해 구현되는 도 2a 의 시스템 내의 일부 컴포넌트들을 도시한다;
- 도 3 은 모션-기반 트래킹 및 오브젝트 검출을 수행하는 방법을 도시하는 흐름도이다;
- 도 4 는 모션-기반 트래킹을 수행하는 방법을 도시하는 흐름도이다;
- 도 5 는 포워드-백워드 에러에 기초하는 모션-기반 트래킹에서 트래킹 에러를 추정하는 방법을 도시하는 흐름도이다;
- 도 6 은 오브젝트 검출을 수행하는 방법을 도시하는 흐름도이다;
- 도 7 은 본 시스템들 및 방법들과 이용될 수도 있는 상이한 윈도우 사이즈들을 도시하는 블록도이다;
- 도 8 은 오브젝트 트래킹 및 검출 모듈의 다른 가능한 구성을 도시하는 블록도이다;
- 도 9 는 평활화 모듈을 도시하는 블록도이다;
- 도 10 은 모션 트래킹 결과들에서 지터를 평활화하는 방법을 도시하는 흐름도이다;
- 도 11 은 오브젝트 트래킹을 이용하여 화상 프로세싱을 수행하는 방법의 흐름도이다;
- 도 12a 는 오브젝트 트래킹을 이용하는 화상 프로세싱의 일 예를 도시한다;
- 도 12b 는 또한 오브젝트 트래킹을 이용하는 화상 프로세싱의 일 예를 도시한다;
- 도 13 은 오브젝트 트래킹을 이용하는 화상 프로세싱의 다른 예를 도시한다;
- 도 14 는 오브젝트 트래킹을 이용하여 비디오 시퀀스에 대해 화상 프로세싱을 수행하는 방법의 흐름도이다;
- 도 15 는 전자 디바이스 상에 디스플레이된 편집되지 않은 비디오 시퀀스 및 편집된 비디오 시퀀스 양자 모두의 다수의 프레임들을 도시한다; 그리고
- 도 16 은 전자 디바이스 디바이스 내에 포함될 수도 있는 소정의 컴포넌트들을 도시한다.

### 발명을 실시하기 위한 구체적인 내용

[0016]

모바일 플랫폼 (예를 들어, 태블릿들, 폰들) 으로부터의 카메라를 이용하여 이미지 또는 해당 이미지 내의 사용자-정의된 관심 구역 내의 오브젝트를 트래킹하는 것은 어려울 수도 있다. 실시간 수행 (초당 ~30 프레임 (frames per second; fps)) 이 요구될 수도 있다. 일부 구성들은 광 흐름-기반 트래커 및 이미지 콘텐츠-기반 검출기의 출력을 결합하여 강건한 트래킹을 획득할 수도 있다. 그러나, 기존의 알고리즘의 계산은 모바일 플랫폼들이 실시간 성능을 달성하지 못하게 할 수도 있다.

[0017]

본 시스템들 및 방법들은 알고리즘의 트래킹 및 검출의 속도를 향상시키기 위해 다음의 기법들을 구현할 수도 있다: (1) 각각의 프레임에서 가능한 검출 윈도우들의 일부분을 이용 (예를 들어, 윈도우 위치선들을 무작위로 선택한다); (2) 이전에 검출된 타겟 사이즈에 가까운 오브젝트 검출을 위해 오직 몇 개의 공간적 스케일들만을 선택; (3) 이전 트래킹의 신뢰도 값에 기초하여, 부분 또는 전체 이미지에서 오브젝트를 검색할지를 결정; (4) 이전 트래킹 결과들에 기초하여 검출 윈도우들의 개수를 다이내믹하게 조절; (5) 병렬로 트래커 및 오브젝트 검출기를 구동하는 대신에, 우선 트래커를 적용, 이것이 계산상 덜 비싸기 때문이다; 그리고 (6) 오직 트래커의 신뢰도가 소정의 임계치보다 낮은 경우에만 오브젝트 검출기를 구동. 기술적 이점들 중 하나의 이점은 타겟 오브젝트를 트래킹하고/하거나 검출하는데 이용되는 계산들을 감소시키는 것이다.

[0018]

트래킹 및 검출 알고리즘의 일 특징적 이용은 화상 프로세싱이다. 화상 프로세싱은 사진 찍기 및/또는 비디

오 편집을 포함할 수도 있다. 화상 프로세싱을 구현하는 것은 설명된 트래킹 및 검출 알고리즘의 실시간 이용 애플리케이션을 제공할 수도 있다.

[0019] 본원에서 이용되는 바와 같이, 용어 "트래킹하다" 및 그것의 변형들은 특정 오브젝트를 식별하지 않는 모션 기반인 프로세스를 지칭한다. 예를 들어, 오브젝트 트래킹 및 검출 모듈은 프레임마다 (from frame to frame) 모션을 트래킹하고, (예를 들어, 카메라가 패닝인 경우) 전자 디바이스의 이동 또는 프레임마다 오브젝트들의 이동들에 기초하여 타겟 오브젝트의 위치, 사이즈, 또는 프레임을 결정할 수도 있다. 용어 "검출하다" 및 그것의 변형들은, 예를 들어, 프레임의 일부를 참조 이미지와 비교함으로써 타겟 오브젝트를 식별하려고 시도하는 프로세스를 지칭한다. 예를 들어, 오브젝트 트래킹 및 검출 모듈은 타겟 오브젝트를 식별하기 위한 시도로 캡처된 프레임들의 일부들을 (타겟 오브젝트의) 참조 이미지와 비교할 수도 있다. 일 예에서, 타겟이 더 이상 트래킹될 수 없는 경우 (예를 들어, 오브젝트가 시계 (field of view) 의 외부에 오게 되는 경우) 검출이 이용될 수도 있다. 모션-기반 트래킹 및 오브젝트 검출을 수행하는 시스템들 및 방법들이 하기에서 보다 상세히 설명된다.

[0020] 도 1 은 본 시스템들 및 방법들에서의 이용을 위한 전자 디바이스 (102) 를 도시하는 블록도이다. 전자 디바이스 (102) 는 또한 무선 통신 디바이스, 모바일 디바이스, 모바일 스테이션, 가입자 스테이션, 클라이언트, 클라이언트 스테이션, 사용자 장비 (user equipment; UE), 원격 스테이션, 액세스 단말기, 모바일 단말기, 단말기, 사용자 단말기, 가입자 유닛 등으로 지칭될 수도 있다. 전자 디바이스들의 예들은 랩탑 컴퓨터 또는 데스크탑 컴퓨터, 셀룰러 전화기들, 스마트 폰들, 무선 모뎀들, e-리더들, 태블릿 디바이스들, 게임 시스템들 등을 포함한다. 이러한 디바이스들 중 일부 디바이스는 하나 이상의 산업 표준들에 따라 동작할 수도 있다.

[0021] 전자 디바이스 (102), 예컨대, 스마트폰 또는 태블릿 컴퓨터는 카메라를 포함할 수도 있다. 카메라는 이미지 센서 (114) 및, 광학 시스템 (118) 의 시계 내에 위치한 오브젝트들의 이미지들을 이미지 센서 (114) 상에서 초점을 맞추는 광학 시스템 (118) (예를 들어, 렌즈들) 을 포함할 수도 있다. 전자 디바이스 (102) 는 또한 카메라 소프트웨어 애플리케이션 및 디스플레이 스크린을 포함할 수도 있다. 카메라 애플리케이션이 구동 중인 경우, 광학 시스템 (118) 의 시계 내에 위치한 오브젝트들의 이미지들은 이미지 센서 (114) 에 의해 레코딩될 수도 있다. 이미지 센서 (114) 에 의해 레코딩되고 있는 이미지들은 디스플레이 스크린 상에 디스플레이될 수도 있다. 이러한 이미지들은, 임의의 주어진 순간에, 광학 시스템 (118) 의 시계 내에 위치한 오브젝트들이 디스플레이 스크린 상에 디스플레이되도록, 상대적으로 높은 프레임 레이트로 연속하여 디스플레이될 수도 있다. 본 시스템들 및 방법들이 캡처된 비디오 프레임들의 면에서 설명되나, 본원에서 논의된 기법들은 임의의 디지털 이미지에 대해 이용될 수도 있다. 따라서, 용어들 비디오 프레임 및 디지털 이미지는 본원에서 상호교환가능하게 이용될 수도 있다.

[0022] 카메라 애플리케이션의 사용자 인터페이스 (120) 는 디스플레이 스크린 상에 디스플레이되고 있는 하나 이상의 오브젝트들이 트래킹되는 것을 가능하게 할 수도 있다. 전자 디바이스 (102) 의 사용자는 트래킹될 오브젝트(들)를 선택하는 것이 허락될 수도 있다. 또한, 선택된 오브젝트(들)는 추후 오브젝트를 검출하는데 참조로서 이용될 수도 있다.

[0023] 일 구성에서, 디스플레이는, 예를 들어, 손가락, 스타일러스, 또는 다른 툴에 의한, 물리적 터치로부터 입력을 수신하는 터치스크린 (116) 이다. 터치스크린 (116) 은 트래킹될 타겟 오브젝트를 정의하는 터치 입력을 수신할 수도 있다. 예를 들어, 전자 디바이스 (102) 가 관심 동물을 포함하는 자연 장면을 캡처하는 경우, 사용자는 동물이 추적될, 또는 필요하다면 검출될 것이라는 요구를 나타내는 바운딩 박스를 동물 주위에 그릴 수도 있다. 타겟 오브젝트들은 임의의 적절한 방식으로 선택될 수도 있다. 예를 들어, 안면 인식, 보행자 인식 등이 트래킹되거나, 검출되거나, 또는 트래킹되고 검출되는 타겟 오브젝트를 선택하는데 이용될 수도 있다. 일 구성에서는, 다수의 오브젝트들이 트래킹될 수도 있다. 사용자 인터페이스 (120) 는 사용자가 오브젝트 트래킹 및 검출 모듈 (104) 과 상호작용하여, 예를 들어, 하나 이상의 타겟 오브젝트들을 선택 (즉, 정의) 하는 것을 허용할 수도 있다. 터치스크린 (116) 은 뷰파인더 (131) 를 포함할 수도 있다. 뷰파인더 (131) 는 비디오 스트림 또는 라이브 피드를 디스플레이하는 터치스크린 (116) 의 일부분을 지칭할 수도 있다. 예를 들어, 뷰파인더 (131) 는 전자 디바이스 (102) 상의 카메라에 의해 획득된 뷰를 디스플레이할 수도 있다.

[0024] 전자 디바이스 (102) 는 선택된 오브젝트를 트래킹하고/하거나 비디오 프레임에서 오브젝트를 검출하기 위한 오브젝트 트래킹 및 검출 모듈 (104) 을 포함할 수도 있다. 오브젝트 트래킹 및 검출 모듈 (104) 은 하나 이상의 오브젝트들을 트래킹하기 위한 모션 트랙커 (106) 를 포함할 수도 있다. 모션 트랙커 (106) 는 프레임

(예를 들어, 비디오 프레임) 마다 이미지에서의 지점들의 모션을 트래킹하기 위해 모션-기반하여 이전 비디오 프레임과 현재 비디오 프레임 사이의 타겟 오브젝트의 위치 및/또는 위치의 변화를 추정할 수도 있다.

[0025] 오브젝트 트래킹 및 검출 모듈 (104) 은 또한 비디오 프레임에서 오브젝트를 검출하기 위한 오브젝트 검출기 (108) 를 포함할 수도 있다. 오브젝트 검출기 (108) 는 모션-기반 모델보다는 오브젝트 모델을 이용하여, 현재 비디오 프레임의 모두 또는 일부분을 (예를 들어, 비디오 프레임들의 시퀀스에서) 캡처된 이전 비디오 프레임 (112) 의 일부분과 비교함으로써 오브젝트를 검출할 수도 있다. 오브젝트 검출기 (108) 는 비디오 프레임 내의 다수의 오브젝트들을 검출하는데 이용될 수도 있다.

[0026] 오브젝트 트래킹 및 검출 모듈 (104) 은 또한 메모리 버퍼 (110) 를 포함할 수도 있다. 메모리 버퍼 (110) 는 하나 이상의 캡처된 프레임들 및 캡처된 비디오 프레임들과 연관된 데이터를 저장할 수도 있다. 일 예에서, 메모리 버퍼 (110) 는 이전에 캡처된 비디오 프레임 (112) 을 저장할 수도 있다. 오브젝트 트래킹 및 검출 모듈 (104) 은 모션-기반 트래킹 및/또는 오브젝트 검출을 수행할 시에 캡처된 이전 비디오 프레임 (112) 에 대해 메모리 버퍼 (110) 로부터 제공된 데이터를 이용할 수도 있다. 데이터는 타겟 오브젝트를 보다 정확하게 트래킹하고/하거나 검출하기 위해 모션-기반 트래킹 및 오브젝트 검출을 조정하도록 메모리 버퍼 (110) 로부터 피드백을 통해 모션 트래커 (106) 또는 오브젝트 검출기 (108) 에 제공될 수도 있다. 예를 들어, 메모리 버퍼 (110) 는 오브젝트를 트래킹하거나 검출하는 경우 모션 트래커 (106) 및 오브젝트 검출기 (108) 에 오브젝트의 위치 및 사이즈를 보다 정확하게 짚어내는데 이용될 수도 있는 하나 이상의 파라미터들을 제공하기 위해 모션 트래커 (106) 및 오브젝트 검출기 (108) 에 위치 및 윈도우 사이즈 데이터를 제공할 수도 있다.

[0027] 위에서 언급된 바와 같이, 전자 디바이스 (102) 는 모션-기반 트래킹을 수행할 수도 있다. 모션-기반 트래킹은 다양한 방법들을 이용하여 수행될 수도 있다. 일 예에서, 트래킹은 중간값 흐름 방법에 의해 수행되며, 여기서 모션 트래커 (106) 는 한 쌍의 이미지들  $I_t$ ,  $I_{t+1}$  (예를 들어, 비디오 프레임들) 및 바운딩 박스 ( $\beta_t$ ) 를 받아들이고 바운딩 박스 ( $\beta_{t+1}$ ) 를 출력한다. 지점들의 세트가 바운딩 박스 ( $\beta_t$ ) 내의 직사각형 그리드에서 초기화되고 지점들을 트래킹하여  $I_t$  와  $I_{t+1}$  사이의 최소 모션 흐름을 발생시킬 수도 있다. 지점 예측의 품질이 추정될 수도 있고 각각의 지점에는 에러가 할당된다. 최악의 예측들 중 일부분 (예를 들어, 50%) 은 필터링될 수도 있고, 한편 남아 있는 예측들은 전체 바운딩 박스의 변위를 추정하는데 이용된다. 모션 트래커 (106) 는 전자 디바이스 (102) 에 의해 캡처된 각각의 비디오 프레임에 대해 모션-기반 트래킹을 수행할 수도 있다. 유사한 방법에서, 모션-기반 트래킹은 현재 비디오 프레임 내의 타겟 오브젝트를 정확하게 트래킹하기 위해 하나 이상의 경사도들 (예를 들어,  $x$  및  $y$  경사도들) 을 산출하고 시간 경사도를 산출하기 위해 한 쌍의 프레임 사이의 차이를 이용하고 다수의 경사도 값들을 이용함으로써 수행될 수도 있다. 모션-기반 트래킹에 관한 보다 상세한 사항들이 하기에서 제공된다.

[0028] 모션-기반 트래킹을 수행하는 경우, 모션 트래커 (106) 는 모션-트래킹 방법의 산출된 또는 추정된 정확도에 기초하여 트래킹 신뢰도 값을 결정할 수도 있다. 일부 구성들에서, 트래킹 신뢰도 값은 타겟 오브젝트가 현재 비디오 프레임 또는 비디오 프레임의 정의된 윈도우 내에 속하는 가능성 또는 확률에 대응하는 0 과 1 사이의 실수일 수도 있다. 트래킹 신뢰도 값은 트래킹 임계치와 비교될 수도 있다. 트래킹 신뢰도 값이 트래킹 임계치보다 크면, 타겟 오브젝트가 현재 비디오 프레임 내에서 발견되는 가능성이 높을 수도 있다. 그렇지 않고, 트래킹 신뢰도 값이 트래킹 임계치 이하이면, 타겟 오브젝트가 현재 비디오 프레임에서 발견되는 가능성이 낮을 수도 있거나 그 여부가 불확실하다. 트래킹 신뢰도 값을 결정하기 위한 다양한 방법들이 이용될 수도 있다. 일 구성에서, 트래킹 신뢰도 값은 현재 비디오 프레임에서의 트래킹된 윈도우 (예를 들어, 트래킹 패치 윈도우) 와 이전에 캡처된 비디오 프레임들로부터의 이전에 저장된 이미지 패치들 사이의 정규화된 교차상관 (normalized cross correlation; NCC) 을 산출함으로써 결정된다. 트래킹 신뢰도 값을 결정하는 것에 관한 추가적인 상세한 사항들이 하기에서 제공된다.

[0029] 전자 디바이스 (102) 는 또한 오브젝트 검출을 수행할 수도 있다. 오브젝트 검출은 다양한 방법들을 이용하여 수행될 수도 있다. 일 구성에서, 오브젝트 검출은 슬라이딩 윈도우 방법을 이용하여 수행되며, 여기서 타겟 오브젝트가 현재 비디오 프레임에서 또는 현재 비디오 프레임의 특정 윈도우나 윈도우들의 서브세트에서 발견되는지 여부를 결정하기 위해 비디오 프레임 내의 윈도우들의 다수의 서브세트들의 콘텐츠가 뷰잉된다. 모든 가능한 윈도우 위치들 및 사이즈들의 전부 또는 서브세트가 비디오 프레임에서 검색될 수도 있다. 예를 들어, 각각의 윈도우는 데이터의 픽셀들에 대응할 수도 있고, 오브젝트 검출기 (108) 는 데이터의 픽셀들을 이용하여 하나 이상의 계산들을 수행해 타겟 오브젝트가 특정 윈도우 또는 서브윈도우 내에 있는 신뢰도의 레벨 (예를 들어, 이진 표시자) 을 결정할 수도 있다. 하나 이상의 윈도우들과 연관된 신뢰도의 레벨에



기초하여, 검출기 신뢰도 값이 현재 비디오 프레임에 대해 획득될 수도 있다. 또한, 오브젝트 검출의 정확도 또는 효율을 증가시키기 위해 추가적인 기법들이 이용될 수도 있다. 이러한 기법들 중 일부 기법이 하기에 설명된다.

[0030] 일부 구성들에서, 모션 트랙커 (106) 및 오브젝트 검출기 (108) 는 병렬로 보다는 순차적으로 동작할 수도 있다. 예를 들어, 전자 디바이스 (102) 는 선택된 오브젝트 (예를 들어, 타겟 오브젝트) 의 모션-기반 트랙킹을 수행하고, 트랙킹된 파라미터에 기초하여 선택된 오브젝트의 오브젝트 검출을 순차적으로 수행할 수도 있다. 일 구성에서, 전자 디바이스 (102) 는 현재 비디오 프레임에 대해 모션-기반 트랙킹을 수행할 수도 있다. 전자 디바이스 (102) 는 그 다음에 트랙킹된 파라미터에 기초하여 현재 프레임에 대해 오브젝트 검출을 수행할 수도 있다. 일 구성에서, 트랙킹된 파라미터는 신뢰도 값과 임계치 사이의 비교에 기초할 수도 있다. 예를 들어, 트랙킹 신뢰도 값이 트랙킹 임계치보다 아래인 경우, 전자 디바이스 (102) 는 오브젝트 검출을 수행할 수도 있다. 그렇지 않고, 트랙킹 신뢰도 값이 트랙킹 임계치보다 위인 경우, 전자 디바이스 (102) 는 현재 비디오 프레임의 모션 트랙킹 결과들에 기초하여 현재 비디오 프레임에 대한 오브젝트 검출을 건너뛰고 다음 비디오 프레임에 대해 모션-기반 트랙킹을 계속 수행할 수도 있다. 다시 말해, 오브젝트 검출은 오직 모션-기반 트랙킹이 매우 좋지 않은 경우에만, 예를 들어, 트랙킹 신뢰도 값이 트랙킹 임계치 아래인 경우에만 수행될 수도 있다. 오브젝트 검출이 수행되는지 여부 및/또는 오브젝트 검출이 수행되는 방법을 고려하는 경우에 다른 트랙킹된 파라미터들이 이용될 수도 있다. 트랙킹된 파라미터들의 예들은 타겟 오브젝트의 구역, 윈도우 위치, 윈도우 사이즈, 스케일 레벨, 타겟 사이즈, 트랙킹 및/또는 검출 신뢰도 값, 또는 타겟 오브젝트의 효율적인 트랙킹 및/또는 검출을 가능하게 하는데 이용될 수도 있는 다른 파라미터를 포함할 수도 있다.

[0031] 모션-기반 트랙킹 및 트랙킹된 파라미터에 기초하여 오브젝트 검출을 순차적으로 수행하는 것은 전자 디바이스 (102) 가 아주 많은 계산들을 수행하지 않으면서 비디오 프레임 내에서 타겟 오브젝트를 트랙킹하고/하거나 검출하는 것을 가능하게 할 수도 있다. 구체적으로, 모션-기반 트랙킹이 오브젝트 검출보다 덜 계산 집약적일 수도 있기 때문에, 전자 디바이스 (102) 는 현재 비디오 프레임 내의 타겟 오브젝트를 정확하게 트랙킹하는데 모션-기반 트랙킹이 이용될 수도 있는 경우 오브젝트 검출을 수행하는 것을 건너뛸 수도 있다. 예를 들어, 트랙킹 신뢰도 값이 특정 타겟 임계치를 초과한다고 전자 디바이스 (102) 가 결정하면, 전자 디바이스 (102) 는 현재 비디오 프레임 내의 타겟 오브젝트의 위치 또는 존재를 정확하게 결정하기 위해 현재 비디오 프레임에 대해 오브젝트 검출이 필요하지 않다고 결정할 수도 있다. 또한, 오브젝트 검출이 많은 경우에 유리할 수도 있기 때문에, 전자 디바이스 (102) 는 트랙킹 임계 값에 대한 비교에 기초하여 모션-기반 트랙킹이 부적당한 경우에 타겟 오브젝트를 보다 정확하게 검출하거나 오브젝트 검출을 수행하기 위해 오브젝트 검출이 이용될 수도 있는 경우들을 결정할 수도 있다.

[0032] 일부 구성들에서는, 현재 비디오 프레임에 대한 오브젝트 검출을 건너뛰는 대신에, 메모리 버퍼 (110) 에 의해 제공되는 모션-기반 트랙킹의 결과들 및/또는 추가적인 정보가 오브젝트 검출을 수행하는 프로세스를 줄이거나 조정하는데 이용될 수도 있다. 예를 들어, 타겟 오브젝트가 모션-기반 트랙킹 방법을 이용하여 정확하게 트랙킹될 수 없는 경우, 전자 디바이스 (102) 는 모션-기반 트랙킹을 통해 제공되는 파라미터들 없이 보다 적은 계산력을 이용하여 오브젝트를 보다 정확하게 검출하기 위해 위치, 윈도우 스케일, 또는 오브젝트 검출 중에 이용될 수도 있는 타겟 오브젝트와 연관된 다른 트랙킹된 파라미터들을 여전히 추정하거나 획득할 수도 있다. 따라서, 모션-기반 트랙킹이 트랙킹 임계치를 초과하는 트랙킹 신뢰도 값을 제공하지 않는 경우일지라도, 후속하여 오브젝트 검출을 수행하는 경우에 모션-기반 트랙킹의 결과들이 이용될 수도 있다.

[0033] 전자 디바이스 (102) 상의 뷰파인더 (131) 는 제 1 트랙킹 영역 (133) 및 제 2 트랙킹 영역 (135) 을 포함할 수도 있다. 제 1 트랙킹 영역 (133) 및 제 2 트랙킹 영역 (135) 양자 모두는 터치스크린 (116) 을 이용하여 사용자에게 의해 명시될 수도 있다. 예를 들어, 사용자는 터치스크린 (116) 상에서 포커스 링을 제 1 트랙킹 영역 (133) 및 제 2 트랙킹 영역 (135) 의 원하는 위치들로 드래그할 수도 있다. 요구되지는 않으나, 트랙킹 영역들 중 하나의 트랙킹 영역은 고정되어 있을 수도 있다. 예를 들어, 제 1 트랙킹 영역 (133) 은 사람의 걸음걸이를 따라갈 수도 있고 제 2 트랙킹 영역 (135) 은 고정되어 있는 나무를 커버할 수도 있다. 일 구성에서, 제 2 트랙킹 영역 (135) 은 전자 디바이스 (102) 상의 전체 터치스크린 (116) 을 커버할 수도 있다.

[0034] 전자 디바이스 (102) 는 화상 프로세싱 모듈 (137) 을 포함할 수도 있다. 화상 프로세싱 모듈 (137) 은 상이한 유형의 화상 프로세싱, 예컨대 사진 찍기 또는 사전에 레코딩된 비디오 편집하기를 제공할 수도 있다. 화상 프로세싱 모듈 (137) 은 중첩 (overlap) (143) 을 포함할 수도 있다. 중첩 (143) 은 제 1 트랙킹 영역 (133) 과 제 2 트랙킹 영역 (135) 사이의 중첩의 양을 반영할 수도 있다. 예를 들어, 중첩 (143) 은 제

1 트랙킹 영역 (133) 과 제 2 트랙킹 영역 (135) 이 서로 전혀 중첩하지 않는다면 0% 일 수도 있다. 마찬가지로, 제 1 트랙킹 영역 (133) 이 제 2 트랙킹 영역 (135) 과 완전히 중첩된다면 (또는 어떤 트랙킹 영역이 더 큰지에 따라, 제 2 트랙킹 영역 (135) 이 제 1 트랙킹 영역 (133) 과 완전히 중첩된다면) 중첩 (143) 은 100% 일 수도 있다.

[0035] 화상 프로세싱 모듈 (137) 은 임계치 (145) 를 포함할 수도 있다. 중첩 (143) 은 임계치 (145) 와 비교되어 화상 프로세싱이 수행되어야 하는지 여부를 결정할 수도 있다. 예를 들어, 중첩 (143) 이 임계치 (145) 보다 크게 되는 경우 사진 (149) 이 찍힐 수도 있다. 다른 예로서, 중첩 (143) 이 임계치 (145) 보다 작게 되는 경우 사진 (149) 이 찍힐 수도 있다. 또 다른 예에서, 중첩 (143) 이 임계치 (145) 보다 더 커지거나 작아지게 되는 경우 비디오 편집이 수행될 수도 있다. 비디오 편집의 일 예에서, 편집되지 않은 비디오 시퀀스 (147) 에서의 프레임들은 편집된 비디오 시퀀스 (151) 를 획득하기 위해 편집될 수도 있다.

[0036] 도 2a 는 오브젝트 트랙킹 및 검출 모듈 (204) 을 도시하는 블록도이다. 오브젝트 트랙킹 및 검출 모듈 (204) 은 전자 또는 무선 디바이스 내에 구현될 수도 있다. 오브젝트 트랙킹 및 검출 모듈 (204) 은 광학 흐름 모듈 (226) 및 트랙킹 신뢰도 값 (228) 을 갖는 모션 트랙커 (206) 를 포함할 수도 있다. 오브젝트 트랙킹 및 검출 모듈 (204) 은 또한 스캐너 로케이터 (230), 스캐너 스케일러 (236), 분류기 (238), 및 검출 신뢰도 값 (240) 을 갖는 오브젝트 검출기 (208) 를 포함할 수도 있다. 메모리 버퍼 (210) 는 모션 트랙커 (206) 및 오브젝트 검출기 (208) 에 제공될 수도 있는 캡처된 이전 비디오 프레임 (212) 과 연관된 데이터를 저장할 수도 있다. 오브젝트 트랙킹 및 검출 모듈 (204), 모션 트랙커 (206), 오브젝트 검출기 (208), 및 메모리 버퍼 (210) 는 도 1 과 관련되어 위에서 설명된 오브젝트 트랙킹 및 검출 모듈 (104), 모션 트랙커 (106), 오브젝트 검출기 (108), 및 메모리 버퍼 (110) 의 구성들일 수도 있다.

[0037] 모션 트랙커 (206) 는 현재 비디오 프레임 (N) (224) 에 대해 모션-기반 트랙킹을 수행하는데 이용될 수도 있다. 예를 들어, 이전 비디오 프레임 (N-1) (222) 및 현재 비디오 프레임 (N) (224) 이 (예를 들어, 전자 디바이스 (102) 에 의해) 수신될 수도 있다. 이전 비디오 프레임 (N-1) (222) 은 비디오 프레임들의 시퀀스에서 현재 비디오 프레임 (N) (224) 에 바로 선행할 수도 있다. 추가적인 비디오 프레임들이 오브젝트 트랙킹 및 검출 모듈 (204) 에 의해 획득되고 프로세싱될 수도 있다. 이전 비디오 프레임 (N-1) (222) 은 모션 트랙커 (206) 에 제공될 수도 있다. 또한, 메모리 버퍼 (210) 는, 본원에서 캡처된 이전 비디오 프레임 (212) 이라고 지칭되는, 이전 비디오 프레임 (N-1) (222) 과 연관된 데이터를 저장할 수도 있다. 일부 구성들에서, 메모리 버퍼 (210) 는 전자 디바이스 (102) 로부터 (예를 들어, 카메라로부터) 직접적으로 이전 비디오 프레임 (N-1) (222) 에 관한 정보를 획득할 수도 있다. 메모리 버퍼 (210) 는 또한 이전 비디오 프레임 (N-1) (222) 에서 오브젝트가 트랙킹되었고/되었거나 검출된 곳을 명시할 수도 있는 이전 비디오 프레임 (N-1) (222) 에 관한 트랙킹 결과들을 통합 모듈 (260) 로부터 획득할 수도 있다. 이전 비디오 프레임 (N-1) (222) 또는 다른 이전에 캡처된 비디오 프레임들에 관한 이러한 정보는 메모리 버퍼 (210) 에 저장될 수도 있다.

[0038] 모션 트랙커 (206) 는 후속하여 비디오 프레임들의 시퀀스에서 현재 비디오 프레임 (N) (224) 을 수신할 수도 있다. 모션 트랙커 (206) 는 (예를 들어, 메모리 버퍼 (210) 로부터 제공된 정보를 이용하여) 현재 비디오 프레임 (N) (224) 을 이전 비디오 프레임 (N-1) (222) 과 비교할 수도 있다. 모션 트랙커 (206) 는 광학 흐름 모듈 (226) 을 이용하여 현재 비디오 프레임 (N) (224) 에 대해 오브젝트의 모션을 트랙킹할 수도 있다. 광학 흐름 모듈 (226) 은 현재 비디오 프레임 (N) (224) 에 대해 오브젝트의 모션-기반 트랙킹을 수행하기 위한 하드웨어 및/또는 소프트웨어를 포함할 수도 있다. 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 을 비교함으로써, 모션 트랙커 (206) 는 타겟이 현재 비디오 프레임 (N) (224) 에 있는 가능성과 연관된 트랙킹 신뢰도 값 (228) 을 결정할 수도 있다. 일 예에서, 트랙킹 신뢰도 값 (228) 은 타겟 오브젝트가 현재 비디오 프레임 (N) (224) 또는 현재 비디오 프레임 (N) (224) 내의 윈도우 내에 있는 확실성의 백분율에 기초한 (예를 들어, 0 과 1 사이의) 실수이다.

[0039] 오브젝트 검출기 (208) 는 현재 비디오 프레임 (N) (224) 에서 오브젝트를 검출하는데 이용될 수도 있다. 예를 들어, 오브젝트 검출기 (208) 는 비디오 프레임들의 시퀀스에서 현재 비디오 프레임 (N) (224) 을 수신할 수도 있다. 오브젝트 검출기 (208) 는 트랙킹된 파라미터에 기초하여 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행할 수도 있다. 트랙킹된 파라미터는 타겟 오브젝트가 정확하게 트랙킹되는 가능성에 대응하는 트랙킹 신뢰도 값 (228) 을 포함할 수도 있다. 보다 구체적으로, 트랙킹된 파라미터는 트랙킹 신뢰도 값 (228) 의 트랙킹 임계치 (250) 와의 비교를 포함할 수도 있다. 트랙킹된 파라미터는 또한 메모리 버퍼 (210) 로부터 제공된 정보를 포함할 수도 있다. 오브젝트를 검출할 경우에 이용될 수도 있는 트랙킹된

파라미터들의 일부 예들은 오브젝트 검출을 수행하는 경우에 파라미터로서 오브젝트 검출기 (208) 에 의해 이용될 수도 있는 구역, 윈도우 위치, 윈도우 사이즈, 또는 다른 정보를 포함한다.

[0040] 오브젝트 검출기 (208) 는 스캐너 로케이터 (230) 를 포함할 수도 있다. 스캐너 로케이터 (230) 는 윈도우 위치 선택기 (232) 및 랜덤화기 (234) 를 포함할 수도 있다. 윈도우 위치 선택기 (232) 는 비디오 프레임 내의 다수의 윈도우들을 선택할 수도 있다. 예를 들어, 비디오 프레임은 다수의 윈도우들을 포함할 수도 있으며, 다수의 윈도우들의 각각은 연관된 위치 및 사이즈를 갖는다. 일 구성에서, 각각의 비디오 프레임은 다수의 (예를 들어, 약 10,000 개의) 중첩되는 윈도우들로 나뉘지며, 각각의 중첩되는 윈도우들은 비디오 프레임에서의 전체 픽셀들의 일부분을 포함한다. 그렇지 않고, 임의의 적절한 개수의 윈도우들이 있을 수도 있고, 그것들은 중첩되지 않을 수도 있다. 스캐너 로케이터 (230) 내의 윈도우 위치 선택기 (232) 는 타겟 오브젝트를 식별하려고 시도하기 위한 윈도우의 위치를 선택할 수도 있다. 랜덤화기 (234) 는 오브젝트를 검출하기 위해 다양한 사이즈들 및 위치들의 윈도우들을 무작위로 선택할 수도 있다. 일부 구성들에서, 랜덤화기 (234) 는 비디오 프레임 내의 윈도우들을 무작위로 선택한다. 그렇지 않고, 랜덤화기 (234) 는 하나 이상의 인자들에 기초하여 윈도우들을 보다 정확하게 선택할 수도 있다. 예를 들어, 랜덤화기 (234) 는 구역, 사이즈, 또는 오브젝트가 위치될 가능성이 가장 높은 일반적인 위치에 기초하여 윈도우들의 선택을 제한할 수도 있다. 이러한 정보는, 메모리 버퍼 (210) 를 통해 획득될 수도 있거나 모션-기반 트래킹을 통해 획득될 수도 있으며, 전적으로 의지하기에는 충분히 정확하지 않을 수도 있으나, 오브젝트 검출을 수행하는 경우에 도움이 되는 정보를 제공할 수도 있다. 따라서, 랜덤화기 (234) 가 검색을 위해 다수의 윈도우들을 무작위로 선택할 수도 있으나, 오브젝트 검출기 (208) 에 의해 제공된 정보에 기초하여, 윈도우들의 선택은 좁혀질 수도 있고, 따라서 완전히 무작위는 아니다.

[0041] 오브젝트 검출기 (208) 는 또한 스캐너 스케일러 (236) 를 포함할 수도 있으며, 스캐너 스케일러 (236) 는 소정의 사이즈의 윈도우를 그리거나 선택하는데 이용될 수도 있다. 윈도우 사이즈는 이미지가 특정 윈도우 내에 있는지 여부를 검출하기 위해 오브젝트를 검출하거나 윈도우들의 선택을 원래의 이미지와 비교하는 경우에 윈도우들의 사이즈들을 줄이기 위해 스캐너 로케이터 (230) 에 의해 이용될 수도 있다. 스캐너 스케일러 (236) 는 오브젝트를 정의하는 경우에 초기에 소정의 사이즈들 또는 스케일 레벨들의 하나 이상의 윈도우들을 선택하거나, 그렇지 않고, 메모리 버퍼 (210) 로부터 제공된 정보에 기초하여 소정의 사이즈들 또는 스케일 레벨들의 하나 이상의 윈도우들을 그릴 수도 있다.

[0042] 분류기 (238) 는 타겟 오브젝트의 일부 또는 전부가 특정 윈도우에서 발견되는지 여부를 결정하는데 이용될 수도 있다. 일부 구성들에서, 분류기 (238) 는 타겟 오브젝트가 특정 윈도우 또는 서브윈도우 내에서 검출되는지 여부를 나타내기 위해 각각의 윈도우에 대한 이진 값을 생성할 수도 있다. 이러한 분류 (예를 들어, 이진 분류) 는 오브젝트 검출기 (208) 에 의해 검색된 각각의 윈도우에 대해 수행될 수도 있다. 구체적으로, 분류기 (238) 는 오브젝트가 검출되는 각각의 윈도우에 대해 이진수 1 을, 그리고 오브젝트가 검출되지 않는 각각의 윈도우에 대해 이진수 0 을 발생시킬 수도 있다. 1 들 및 0 들의 개수 또는 조합에 기초하여, 오브젝트 검출기 (208) 는 타겟 오브젝트가 현재 비디오 프레임 (N) (224) 내에 존재하는 가능성을 나타내는 검출 신뢰도 값 (240) 을 결정할 수도 있다. 일부 구성들에서, 검출 신뢰도 값 (240) 은 오브젝트가 정확하게 검출된 백분율 또는 확률을 나타내는 0 과 1 사이의 실수이다.

[0043] 오브젝트 검출기 (208) 는 구역, 타겟 사이즈, 윈도우 사이즈, 스케일 레벨, 윈도우 위치, 및 하나 이상의 신뢰도 값들을 포함하는, 다양한 트래킹된 파라미터들에 따라 오브젝트 검출을 수행할 수도 있다. 비디오 프레임 또는 윈도우들의 서브세트가 검색되고 오브젝트 검출기 (208) 가 각각의 검색된 윈도우에 대한 이진 값을 획득하면, 오브젝트 검출기 (208) 는 가장 높은 신뢰도를 갖는 윈도우 사이즈 뿐만 아니라 현재 비디오 프레임에서의 위치 또는 구역을 결정할 수도 있다. 이러한 위치 및 윈도우 사이즈는 타겟 오브젝트를 보다 정확하게 트래킹하고/하거나 검출하기 위해 후속하는 트래킹 및 검출에서 이용될 수도 있다.

[0044] 위에서 언급된 바와 같이, 다양한 방법들이 타겟 오브젝트를 검출할 시에 오브젝트 검출기 (208) 에 의해 이용될 수도 있다. 일 구성에서, 타겟 오브젝트를 검출하는 것은 모든 가능한 윈도우 위치에 있는 윈도우들 및 모든 가능한 윈도우 사이즈에 대해 이진 분류를 수행하는 것을 포함할 수도 있다. 그러나, 모든 가능한 윈도우를 검색하는 것은 자원 집약적이다. 따라서, 다른 구성에서, 오브젝트 검출기는, 비디오 프레임에서의 모든 가능한 윈도우들 보다는, 윈도우 위치들 및 사이즈들의 서브세트를 검색할 수도 있다. 예를 들어, 오브젝트 검출기 (208) 는 모든 가능한 윈도우들의 1% 를 검색할 수도 있다. 그 다음에, 검출이 성공적이지 않으면 (예를 들어, 검출 신뢰도 값 (240) 이 검출 임계치 (252) 보다 작으면), 보다 높은 백분율, 예를 들어, 2% 의 윈도우 위치들이 후속하는 캡처된 프레임에서 검색될 수도 있다. 검색된 백분율의 윈도우 위치들에서



의 단계는 균일하거나, 균일하지 않거나, 느리거나, 빠를 수 있는데, 즉, 연속적인 프레임들은 1%, 2%, 3%, 4% 또는 1%, 2%, 4%, 8% 를 가질 수도 있다. 일 구성에서, 검색된 프레임들의 백분율은 높은 검출 신뢰도 값에 응답하여, 즉, 타겟 오브젝트가 다음 비디오 프레임이도록, 매우 높게 (예를 들어, 80%, 90%, 100%) 설정될 수도 있다. 예를 들어, 검색된 프레임들의 백분율은 검출 및 트래킹 임계 값 (256) 을 초과하는 검출 및 트래킹 신뢰도 값에 응답하여 적어도 80% 로 점프할 수도 있다. 그렇지 않으면, 백분율은 60%, 70%, 90% 등으로 점프할 수도 있다. 또한, 검출 및 트래킹 임계 값에 대한 임의의 적절한 값, 예를 들어, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85 등이 이용될 수도 있다. 또한, 검색된 윈도우들의 백분율은, 랜덤화기 (234) (난수 발생기) 에 기초하여 무작위로 결정될 수도 있는데, 예를 들어, 1% 와 15% 사이의 무작위 백분율의 윈도우들이 캡처된 프레임에서 검색될 수도 있다. 모든 윈도우 위치들의 서브세트를 검색함으로써, 오브젝트 검출은 전자 디바이스 (102) 에서 보다 적은 자원들을 이용할 수도 있다.

[0045] 또한, 본 시스템들 및 방법들은 각각의 위치에 대한 윈도우 사이즈들의 서브세트를 검색할 수도 있다. 각각의 윈도우 사이즈는 본원에서 스케일 레벨이라고 지칭될 수도 있으며, 각각의 스케일 레벨은 특정 윈도우 사이즈에 대응한다. 예를 들어, 20 개의 가능한 스케일 레벨들이 있을 수도 있다. 모두 20 개의 스케일 레벨들을 검색하기 보다는, 스케일 레벨들 또는 윈도우 사이즈들의 서브셋이 각각의 윈도우 위치에서 검색될 수도 있다.

[0046] 본 시스템들 및 방법들은 또한 메모리 버퍼 (210) 로부터의 피드백을 이용하여 검색된 윈도우 위치들 및 사이즈들을 조정할 수도 있다. 다시 말해, 타겟 오브젝트가 성공적으로 검출되고/되거나 트래킹된 마지막 캡처된 비디오 프레임의 위치 및 사이즈는 현재 비디오 프레임 (N) (224) 을 검색하기 위한 시작 지점으로서 이용될 수도 있다. 예를 들어, 타겟 오브젝트가 최근의 비디오 프레임에서 검출되고 트래킹되었으면 (즉, 최근에 캡처된 비디오 프레임에 대한 검출 및 트래킹 신뢰도 값 (256) 이 검출 및 트래킹 임계치 위이면), 스캐너 로케이터가 최근의 프레임과 연관된 위치 및 사이즈에서 현재 캡처된 프레임을 검색하기 시작할 수도 있다. 예를 들어, 타겟 오브젝트가 광학 시스템의 시계를 벗어나거나 멀리로 사라지는 경우, 타겟 오브젝트는 타겟 오브젝트가 광학 시스템의 시계를 떠나거나 멀리 떨어진 경우와 동일한 사이즈로 다시 나타날 가능성이 높을 수도 있다. 따라서, 오브젝트 검출을 수행하는 경우 후속하는 비디오 프레임들에서 타겟 오브젝트를 검출하도록 사이즈 또는 사이즈들의 범위가 예측될 수도 있다.

[0047] 또한, 캡처된 비디오 프레임 (N) (224) 에서 검색된 윈도우 위치들 및 윈도우 사이즈들의 검색 범위는 최근의 비디오 프레임 (예를 들어, 이전 비디오 프레임 (N-1) (222)) 에서의 타겟 오브젝트와 연관된 윈도우 위치 및 윈도우 사이즈와 유사한 것으로 제한될 수도 있다. 본원에서 이용되는 바와 같이, 용어 "검색 범위" 는 비디오 프레임에서 타겟 오브젝트를 검출하고/하거나 트래킹할 경우에 사용될 수도 있는 후보 윈도우 위치들 또는 후보 윈도우 사이즈들 (또는 양자 모두) 의 세트를 지칭한다. 예를 들어, 검색된 윈도우 위치들의 서브셋은 최근의 비디오 프레임에서 타겟 오브젝트가 발견된 곳, 예를 들어, 현재 비디오 프레임 (N) (224) 의 사등분들, 또는 이등분들 중 하나에 기초하여 현재 비디오 프레임 (N) (224) 의 일부분 내에서부터 선택될 수도 있다. 다시 말해, 검색 공간은 타겟 오브젝트가 마지막으로 트래킹되었거나 검출된 곳 근처로 제한될 수도 있다. 유사하게, 각각의 윈도우 위치에 대해 검색된 프레임들의 사이즈들은 최근의 비디오 프레임에서 타겟팅된 오브젝트가 발견된 윈도우의 사이즈에 기초하여 제한될 수도 있다. 예를 들어, 8 의 스케일 레벨을 갖는 윈도우를 이용하여 현재 프레임에서 오브젝트가 검출되었으면, 스캐너 스케일러 (236) 는 오직 8 의 현재 비디오 프레임 (N) (224) 에 대한 윈도우 스케일 레벨들, 플러스 또는 마이너스 3, 즉, 스케일 레벨들 5-11 만을 선택할 수도 있다. 이는 또한 낮은 확률 검색을 없애고 오브젝트 검출의 효율을 증가시킬 수도 있다. 그렇지 않고, 최근의 (현재가 아닌) 비디오 프레임이 타겟 오브젝트를 검출하지 않으면 (즉, 현재 비디오 프레임에 대한 검출 및 트래킹 신뢰도 값 (256) 이 검출 및 트래킹 임계치보다 아래이다), 오브젝트 검출기 (208) 는 검색되는 검색 공간 (윈도우 위치들) 을 확장할 수도 있는데, 예를 들어, 보다 넓은 범위의 이미지 또는 전체 이미지가 검색의 대상이 될 수도 있다.

[0048] 오브젝트 트래킹 및 검출 모듈 (204) 은 다수의 윈도우들을 병합하여 단일 윈도우를 형성하기 위한 통합 모듈 (260) 을 포함할 수도 있다. 초기에 2 개의 신뢰도 값들: 오브젝트 검출기 (208) 로부터의 검출 신뢰도 값 (240) 및 모션 트랙커 (206) 로부터의 트래킹 신뢰도 값 (225) 이 있다. 통합 모듈 (260) 은 2 개의 신뢰도 값들 (예를 들어, 보다 큰 것 하나를 택한다) 을 검출 및 트래킹 신뢰도 값 (256) 에 결합할 수도 있다. 검출 및 트래킹 신뢰도 값 (256) 은 타겟 오브젝트가 비디오 프레임에서 식별되었는지 여부를 나타낼 수도 있다. 일 구성에서, 검출 및 트래킹 신뢰도 값 (256) 은 0 과 1 사이의 실수일 수도 있으며, 여기서 0 은 타겟 오브젝트가 특정 비디오 프레임에서 식별된 가능한 가장 낮은 신뢰도를 나타내고, 1 은 타겟 오브젝트가 특정 비



디오 프레임에서 식별된 가능한 가장 높은 신뢰도를 나타낸다. 다시 말해, 검출 및 트래킹 신뢰도 값 (256)은 타겟 오브젝트가 발견된 가능성의 종합 표시의 역할을 할 수도 있다. 또한, 검출 및 트래킹 신뢰도 값 (256)은 다음 비디오 프레임에서 검색하기 위한 윈도우 위치, 윈도우 사이즈, 윈도우들의 백분율을 결정하는데 이용된 파라미터일 수도 있다. 통합 모듈 (260)은 메모리 버퍼 (210)에 현재 비디오 프레임 (N) (224)에 관한 정보를 제공하는데 이용될 수도 있다. 일 예에서, 통합 모듈 (260)은 트래킹된 윈도우 (242) (예를 들어, 윈도우 위치 (244), 윈도우 사이즈 (246) 등)에 관한 정보 및 검출 및 트래킹 신뢰도 값 (256)을 메모리 버퍼 (210)에 제공할 수도 있다. 통합 모듈 (260)은 모션 트랙커 (206) 및 오브젝트 검출기 (208)로부터의 트래킹 결과들 (예를 들어, 바운딩 박스들)을 이용하여 결합된 트래킹 결과 (예를 들어, 바운딩 박스)를 형성하고 검출 및 트래킹 신뢰도 값 (256)을 산출할 수도 있다.

[0049] 메모리 버퍼 (210)는 이전 비디오 프레임 (N-1) (222), 현재 비디오 프레임 (N) (224), 또는 다른 캡처된 비디오 프레임들과 연관된 하나 이상의 값들을 저장할 수도 있다. 일 구성에서, 메모리 버퍼 (210)는 캡처된 이전 비디오 프레임 (212)을 저장하며, 캡처된 이전 비디오 프레임 (212)은 이전 비디오 프레임 (N-1) (222)에 대응하는 정보를 포함할 수도 있다. 캡처된 이전 비디오 프레임 (212)은 각각의 윈도우 (242)에 대한 위치 (244), 윈도우 사이즈 (246), 및 (예를 들어, 분류기 (238)로부터의) 이진 결정 (248)을 포함하여, 하나 이상의 윈도우들 (242)에 관한 정보를 포함할 수도 있다. 캡처된 이전 비디오 프레임 (212)은 또한 트래킹 임계치 (250), 검출 임계치 (252), 및 검출 및 트래킹 임계치 (254)를 포함할 수도 있다. 트래킹 임계치 (250)는 모션 트랙커 (206) 또는 오브젝트 트래킹 및 검출 모듈 (204) 상의 회로 (예를 들어, 신뢰도 레벨 비교기)에 제공되어 트래킹 신뢰도 레벨이 트래킹 임계치 (250)보다 큰지 여부를 결정할 (258) 수도 있다. 검출 임계치 (252)는 오브젝트 검출기 (208) 또는 오브젝트 트래킹 및 검출 모듈 (204) 상의 다른 회로에 제공되어 검출신뢰도 값 (240)이 검출 임계치 (252)보다 큰지 여부를 결정할 수도 있다. 검출 및 트래킹 임계치 (254)는 트래킹 임계치 (250) 및 검출 임계치 (252)에 기초하여 결합된 값일 수도 있다. 검출 및 트래킹 임계치 (254)는 검출 및 트래킹 신뢰도 값 (256)과 비교되어 모션-기반 트래킹 및 오브젝트 검출에 대한 결합된 신뢰도 값을 결정할 수도 있다. 임계치들의 각각은 타겟 오브젝트가 비디오 프레임 내에 위치되는 가능성에 기초할 수도 있다. 오브젝트 트래킹 및 검출 모듈 (204)은 특정 검출 및 트래킹 신뢰도 값 (256)이 획득될 때까지 현재 비디오 프레임 (N) (224)에 대해 모션-기반 트래킹 및/또는 검출을 수행할 수도 있다. 또한, 모션-기반 트래킹 및 오브젝트 검출은 다수의 비디오 프레임들에서의 각각의 비디오 프레임에 대해 수행될 수도 있다.

[0050] 모션-기반 트래킹 및 오브젝트 검출을 수행하는 것은 모션-기반 트래킹에 뒤이어 트래킹된 파라미터에 기초해 오브젝트 검출을 순차적으로 수행하는 것을 포함할 수도 있다. 특히, 본 시스템들 및 방법들은 2 단계 트래킹 및 검출 접근법을 수행할 수도 있다. 모션-기반 트래킹이 오브젝트 검출에 이용된 실제 오브젝트 식별보다는 장면의 상대적 모션에 기초하기 때문에, 모션-기반 트래킹은 오브젝트 검출을 수행하는 것보다 전자 디바이스에서 덜 자원-집약적일 수도 있다. 이에 따라, 오브젝트 검출기 (208) 대신에 모션 트랙커 (206)를 이용하는 것이 더 효율적일 수도 있으며, 여기서 타겟 오브젝트는 역시 오브젝트 검출을 수행하지 않으면서 정확하게 트래킹될 수도 있다.

[0051] 따라서, 오브젝트 검출기 (208)와 병렬로 모션 트랙커 (206)를 이용하기 보다는, 오브젝트 트래킹 및 검출 모듈 (204)은 오직 모션 트랙커 (206)가 불충분한 경우에만 오브젝트 검출기 (208)를 이용하는데, 즉, (전혀 수행되지 않았다면) 모션 트래킹 및 오브젝트 검출은 병렬로 대신에 순차적으로 수행된다. 트래킹이 수행된 각각의 비디오 프레임에 대해, 모션 트랙커 (206)는 트래킹 신뢰도 값 (228)을 생성할 수도 있으며, 트래킹 신뢰도 값 (228)은 타겟 오브젝트가 현재 비디오 프레임 (N) (224)에 있는 가능성을 나타내는 0과 1 사이의 실수일 수도 있다.

[0052] 2 단계 트래킹 및 검출 접근법의 일 구성에서, 모션 트랙커 (206)는 우선 현재 비디오 프레임 (N) (224)에 대해 모션-기반 트래킹을 수행할 수도 있다. 모션 트랙커 (206)는 모션-기반 트래킹 프로세스에 기초하여 트래킹 신뢰도 값 (228)을 결정할 수도 있다. 트래킹 신뢰도 값 (228) 및 메모리 버퍼 (210)에 의해 제공된 트래킹 임계치 (250)를 이용하여, 오브젝트 트래킹 및 검출 모듈 (204) 내의 회로 (예를 들어, 신뢰도 레벨 비교기)는 트래킹 신뢰도 값 (228)이 트래킹 임계치 (250)를 초과하는지 여부를 결정할 수도 있다 (258). 트래킹 신뢰도 값 (228)이 트래킹 임계치 (250)보다 크면, 오브젝트 트래킹 및 검출 모듈 (204)은 오브젝트 검출을 수행하는 것을 건너뛰고 통합 모듈 (260)에 트래킹 결과를 제공하여 출력 (262)을 생성할 수도 있다. 출력 (262)은 타겟 오브젝트가 현재 비디오 프레임 (N) (224)에 있다는 표시를 포함할 수도 있다. 또한, 출력 (262)은 타겟 오브젝트에 관한 추가적인 정보를 포함할 수도 있다.

- [0053] 트래킹 신뢰도 값 (228) 이 트래킹 임계치 (250) 를 초과하지 않는 경우, 오브젝트 검출기 (208) 는 후속하여 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행할 수도 있다. 오브젝트 검출은 현재 비디오 프레임 (N) (224) 내의 윈도우들 모두 또는 서브세트에 대해 수행될 수도 있다. 오브젝트 검출기 (208) 는 또한 모션-기반 트래킹의 결과들 및/또는 메모리 버퍼 (210) 로부터 제공된 정보에 기초하여 윈도우들의 서브세트, 윈도우 사이즈들, 또는 다른 검출 기준을 선택할 수도 있다. 오브젝트 검출은 오브젝트 검출기 (208) 에 제공된 하나 이상의 트래킹된 파라미터들에 기초하여 더 강건하거나 덜 강건한 프로세스를 이용해 수행될 수도 있다. 오브젝트 검출기 (208) 는 검출 신뢰도 값 (240) 을 결정하고 검출 신뢰도 값 (240) 을 검출 임계치 (252) 와 비교할 수도 있다. 검출 신뢰도 값 (240) 이 검출 임계치 (252) 보다 위이면, 오브젝트 검출기 (208) 는 통합 모듈 (260) 에 검출 결과를 제공하여 출력 (262) 을 생성할 수도 있다. 출력 (262) 은 타겟 오브젝트가 현재 비디오 프레임 (N) (224) 내에 있다는 표시를 포함하고/하거나 검출된 오브젝트에 관한 추가적인 정보를 포함할 수도 있다.
- [0054] 그렇지 않고, 검출 신뢰도 값 (240) 이 검출 임계치 (252) 이하이면, 오브젝트 검출기 (208) 는 보다 강건한 방법을 이용하여, 예컨대 현재 비디오 프레임 (N) (224) 내의 보다 많은 개수의 윈도우들을 검색하여, 다시 오브젝트 검출을 수행할 수도 있다. 오브젝트 검출기 (208) 는 만족스러운 검출 신뢰도 값 (240) 이 획득될 때까지 오브젝트 검출의 프로세스를 반복할 수도 있다. 현재 비디오 프레임 내에서 타겟 오브젝트가 식별되기에 만족스러운 검출 신뢰도 값 (240) 이 획득되면, 오브젝트 트래킹 및 검출 모듈 (204) 은 다음 비디오 프레임에 대한 트래킹 및 검출을 수행하는데 이용될 수도 있다.
- [0055] 도 2b 는 프로세서 (264) 에 의해 구현되는 도 2a 의 시스템 내의 일부 컴포넌트들을 도시한다. 도 2a 에 도시된 바와 같이, 오브젝트 트래킹 및 검출 모듈 (204) 은 프로세서 (264) 에 의해 구현될 수도 있다. 상이한 컴포넌트들을 구현하는데 상이한 프로세서들이 이용될 수도 있다 (예를 들어, 하나의 프로세서가 모션 트래커 (206) 를 구현할 수도 있고, 다른 프로세서가 오브젝트 검출기 (208) 를 구현하는데 이용될 수도 있고, 또 다른 프로세서가 메모리 버퍼 (210) 를 구현하는데 이용될 수도 있다).
- [0056] 도 3 은 모션-기반 트래킹 및 오브젝트 검출을 수행하는 방법 (300) 을 도시하는 흐름도이다. 방법 (300) 은 전자 디바이스 (102), 예를 들어, 오브젝트 트래킹 및 검출 모듈 (104) 에 의해 구현될 수도 있다. 전자 디바이스 (102) 는 이전 비디오 프레임 (N-1) (222) 을 현재 비디오 프레임 (N) (224) 과 비교함으로써 현재 비디오 프레임 (N) (224) 에 대해 모션-기반 트래킹을 수행할 수도 있다 (302). 오브젝트를 트래킹하는 것은 이미지의 쌍들 사이의 지점들을 트래킹함으로써 중간값 흐름 방법을 이용하여 수행될 수도 있다. 모션-기반 트래킹의 다른 방법들이 또한 이용될 수도 있다. 또한, 모션-기반 트래킹은 메모리 버퍼 (110) 를 통해 제공된 캡처된 이전 비디오 프레임 (112) 에 관한 정보를 이용하여 현재 비디오 프레임 (N) (224) 에 대해 수행될 수도 있다.
- [0057] 전자 디바이스 (102) 는 트래킹 신뢰도 값 (228) 을 결정할 수도 있다 (304). 트래킹 신뢰도 값 (228) 은 타겟 오브젝트가 정확하게 트래킹된 가능성 또는 확신도를 나타낼 수도 있다. 전자 디바이스 (102) 는 트래킹 신뢰도 값 (228) 이 트래킹 임계치 (250) 보다 큰지 여부를 결정할 수도 있다 (306). 트래킹 신뢰도 값 (228) 이 트래킹 임계치 (250) 보다 크면, 전자 디바이스 (102) 는 다음 비디오 프레임에 대해 모션-기반 트래킹을 수행할 수도 있다 (308). 또한, 전자 디바이스 (102) 는 모션-기반 트래킹의 결과에 기초하여 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행하는 것을 건너뛸 수도 있다. 다시 말해, 오브젝트 검출은 오직 모션 트래킹이 매우 좋지 않은 경우에만, 즉, 트래킹 신뢰도 값 (228) 이 트래킹 임계치 (250) 보다 크지 않으면, 현재 비디오 프레임 (N) (224) 에 대해 수행될 수도 있다. 그러나, 트래킹 신뢰도 값 (228) 이 트래킹 임계치 (250) 보다 크지 않으면, 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행할 수도 있다 (310). 전자 디바이스 (102) 는 모션-기반 트래킹 다음에 오브젝트 검출을 수행할 수도 있다. 일부 구성들에서, 오브젝트 검출은 다양한 강건성으로 다수 회 수행되어 보다 높은 검출 신뢰도 값 (240) 을 획득할 수도 있다.
- [0058] 도 4 는 모션-기반 트래킹을 수행하는 방법 (400) 을 도시하는 흐름도이다. 방법 (400) 은 전자 디바이스 (102), 예를 들어, 오브젝트 트래킹 및 검출 모듈 (104) 에 의해 구현될 수도 있다. 전자 디바이스 (102) 는 바운딩 박스를 이용하여 타겟 오브젝트를 식별할 수도 있다 (402). 식별하는 것 (402) 은 터치스크린 (116) 또는 관심 오브젝트가 선택되는 다른 입력 방법을 이용하여 수동으로 수행될 수도 있다. 다수의 오브젝트들이 유사한 방식으로 식별될 수도 있다. 또한, 트래킹될 오브젝트를 식별하는데 다른 입력 방법들이 이용될 수도 있다. 일 예에서, 오브젝트는 타겟 오브젝트 주위에 바운딩 박스를 수동으로 그림으로써 식별

된다.

[0059] 전자 디바이스 (102) 는 바운딩 박스 내의 그리드에서의 지점들을 초기화할 수도 있다 (404). 그리드에서의 지점들은 바운딩 박스에 걸쳐 균일하게 이격될 수도 있다. 또한, 지점들은 2 개의 이미지들 (예를 들어, 이전 비디오 프레임 (N-1) (222) 및 현재 비디오 프레임 (N) (224)) 사이의 그리드에서 트랙킹될 수도 있다 (406). 일 예에서, 지점들은 이미지들 사이에 최소 모션 흐름을 발생시키는 루카스-카나데 (Lucas-Kanade) 트랙커에 의해 트랙킹된다. 전자 디바이스 (102) 는 2 개의 이미지들 (예를 들어, 이전 비디오 프레임 (N-1) (222) 및 현재 비디오 프레임 (N) (224)) 사이의 트랙킹 에러를 추정할 수도 있다 (408). 트랙킹 에러를 추정하는 것 (408) 은 트랙킹된 지점들의 각각의 지점에 에러 값을 할당하는 것을 포함할 수도 있다. 또한, 트랙킹 에러를 추정하는 것 (408) 은, 예를 들어, 포워드-백워드 에러, 정규화된 교차 상관 (NCC), 및 제곱합 차이들을 포함하여, 다양한 방법들을 이용해 수행될 수도 있다. 추정된 트랙킹 에러는 트랙킹 신뢰도 값 (228) 을 획득하고 궁극적으로 타겟 오브젝트가 현재 비디오 프레임 (N) (224) 에 있는 가능성을 결정하는데 이용될 수도 있다. 일 구성에서, 트랙킹 신뢰도 값 (228) 은 현재 비디오 프레임 (N) (224) 과 이전 비디오 프레임 (N-1) (222) 에서의 트랙킹된 윈도우 사이의 정규화된 교차 상관 (NCC) 을 산출함으로써 획득될 수도 있다. 트랙킹 에러는 또한 도 5 와 관련하여 하기에서 보다 상세히 설명된 포워드-백워드 에러 추정을 포함하여, 추가적인 기법들을 이용해 추정될 수도 있다. 또한, 전자 디바이스 (102) 는 외곽 지점 예측들을 필터링할 수도 있다 (410). 예를 들어, 전자 디바이스는 최악의 예측들의 50% 를 필터링할 수도 있다. 남아 있는 예측들은 바운딩 박스의 변위를 추정하는데 이용될 수도 있다.

[0060] 전자 디바이스 (102) 는 바운딩 박스를 업데이트할 수도 있다 (412). 바운딩 박스를 업데이트하는 것 (412) 은 업데이트된 바운딩 박스가 다음 비디오 프레임에 대한 새로운 바운딩 박스가 되도록 수행될 수도 있다. 모션-기반 트랙킹 프로세스는 그 다음에 다음 비디오 프레임에 대해 반복될 수도 있거나, 트랙킹 신뢰도 값 (228) 이 트랙킹 임계치 (250) 이하이면, 모션-기반 트랙킹 프로세스는 타겟 오브젝트가 정확하게 트랙킹될 수도 있을 때까지 다음 비디오 프레임에 대해 중단될 수도 있다. 일부 구성들에서, 현재 비디오 프레임 (N) (224) 에 대한 모션-기반 트랙킹이 만족스러운 결과를 제공하지 않는 경우, 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행하여 타겟 오브젝트의 위치를 찾을 시에 보다 높은 레벨의 신뢰도를 획득할 수도 있다. 일부 구성들에서, 모션-기반 트랙킹이 만족스러운 결과를 생성할 수 없는 경우 (예를 들어, 타겟 오브젝트가 비디오 프레임의 범위에서 벗어나는 경우), 오브젝트 검출은 타겟 오브젝트가 검출될 때까지 임의의 후속하는 비디오 프레임들에 대해 수행될 수도 있다.

[0061] 도 5 는 포워드-백워드 에러에 기초하는 모션-기반 트랙킹에서 트랙킹 에러를 추정하는 방법 (500) 을 도시하는 흐름도이다. 방법 (500) 은 전자 디바이스 (102) (예를 들어, 오브젝트 트랙킹 및 검출 모듈 (104)) 에 의해 구현될 수도 있다. 일부 구성들에서, 전자 디바이스 (102) 는 트랙킹된 윈도우들 사이의 정규화된 교차 상관 (NCC) 을 산출할 수도 있다. 정규화된 교차 상관 (NCC) 은 트랙킹 신뢰도 값 (228) 을 결정하는데 이용될 수도 있다. 전자 디바이스 (102) 는 또한 정규화된 교차 상관 (NCC) 을 보완하는 다양한 트랙킹 에러 추정 기법들 (예를 들어, 포워드-백워드 에러, 제곱합 차) 을 이용할 수도 있다. 포워드-백워드 에러 추정을 이용하는 예에서, 전자 디바이스 (102) 는 이전 비디오 프레임 (N-1) (222) 및 현재 비디오 프레임 (N) (224) 사이에 포워드 트랙킹을 수행하여 포워드 궤도를 결정할 수도 있다 (502). 포워드 트랙킹은 k 단계들에 대해 앞서 이미지를 트랙킹하는 것을 포함할 수도 있다. 결과적인 포워드 궤도는  $(x_t, x_{t+1}, \dots, x_{t+k})$  와 동일할 수도 있으며, 여기서  $x_t$  는 시간에서의 지점 위치이고 k 는 이미지들의 시퀀스의 길이를 나타낸다. 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 과 이전 비디오 프레임 (N-1) (222) 사이에 백워드 트랙킹을 수행하여 백워드 궤도를 결정할 수도 있다 (504). 결과적인 백워드 궤도는  $(\hat{x}_t, \hat{x}_{t+1}, \dots, \hat{x}_{t+k})$  와 동일하며, 여기서  $\hat{x}_{t+k} = x_{t+k}$  이다.

[0062] 전자 디바이스 (102) 는 포워드 궤도와 백워드 궤도 사이의 포워드-백워드 에러를 결정할 수도 있다 (506). 포워드-백워드 에러는 포워드 궤도와 백워드 궤도 사이의 거리로 정의될 수도 있다. 또한, 다양한 거리들이 궤도 비교를 위해 정의될 수도 있다. 일 구성에서, 포워드-백워드 에러를 결정하는 경우 유효 궤도의 초기 지점과 종료 지점 사이의 유클리드 거리가 이용될 수도 있다. 일 구성에서, 포워드-백워드 에러는 트랙킹 에러로서 이용될 수도 있으며, 트랙킹 에러는 트랙킹 신뢰도 값 (228) 을 결정하는데 이용될 수도 있다.

[0063] 도 6 은 오브젝트 검출을 수행하는 방법 (600) 을 도시하는 흐름도이다. 방법 (600) 은 전자 디바이스 (102) (예를 들어, 오브젝트 트랙킹 및 검출 모듈 (104)) 에 의해 구현될 수도 있다. 전자 디바이스 (102)



는 현재 비디오 프레임 (N) (224) 에서 윈도우 위치들 및 사이즈들의 서브세트를 검색함으로써 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출 및 모션-기반 트래킹을 수행할 수도 있다 (602).

[0064] 전자 디바이스 (102) 는 검출 및 트래킹 신뢰도 값 (256) 을 결정할 수도 있다 (604). 검출 및 트래킹 신뢰도 값 (256) 은 타겟 오브젝트가 현재 비디오 프레임 (N) (224) 에서 또는 특정 윈도우 내에서 발견되는지 여부의 신뢰도 레벨을 제공할 수도 있다. 전자 디바이스 (102) 는 또한 검출 및 신뢰도 값 (256) 이 검출 및 트래킹 임계치 (254) 보다 큰지 여부를 결정할 수도 있다 (606). 검출 및 신뢰도 값 (256) 이 검출 및 트래킹 임계치 (254) 보다 크면, 전자 디바이스 (102) 는 다음 비디오 프레임에서 윈도우들의 서브세트 (예를 들어, 동일한 서브세트) 및 사이즈들을 이용하여 다음 비디오 프레임에 대해 오브젝트 검출을 수행할 수도 있다 (608). 그렇지 않고, 검출 및 신뢰도 값 (256) 이 검출 및 트래킹 임계치 (254) 보다 작으면, 전자 디바이스 (102) 는 다음 비디오 프레임에서 윈도우 위치들 및 사이즈들의 보다 큰 서브세트를 이용하여 다음 비디오 프레임에 대해 오브젝트 검출을 수행할 수도 있다 (610). 일부 구성들에서, 신뢰도 값 (256) 이 검출 및 트래킹 임계치 (254) 보다 작은 경우, 전자 디바이스 (102) 는 다음 비디오 프레임의 전체 검색 공간 및/또는 모든 윈도우들을 이용하여 다음 비디오 프레임에 대해 오브젝트 검출을 수행할 수도 있다 (610).

[0065] 도 7 은 본 시스템 및 방법들과 이용될 수도 있는 상이한 윈도우 사이즈들 (766) 을 갖는 이미지 윈도우 (700) 를 도시하는 블록도이다. 구체적으로, 도 7 은 10 개의 가능한 윈도우 사이즈들 (766a-j) 의 세트를 도시한다. 각각의 윈도우 사이즈 (766) 는 스케일 레벨 (예를 들어, 1-10) 에 대응할 수도 있다. 본원에서는 직사각형으로 도시되나, 검색된 윈도우들은 임의의 형상, 예를 들어, 정사각형, 직사각형, 원형, 타원형, 사용자-정의 등일 수도 있다. 또한, 임의의 개수의 윈도우 사이즈들 (766) 또는 스케일 레벨들, 예를 들어, 5, 15, 20, 30 등이 이용가능할 수도 있다.

[0066] 위에서 설명된 바와 같이, 검색 범위는 특정 위치에 대해 이용된 윈도우 사이즈들의 서브세트에 의해 표기될 수도 있는데, 예를 들어, 현재 비디오 프레임 (N) (224) 에서 검색된 윈도우 사이즈들은 최근의 프레임에서의 타겟 오브젝트와 연관된 윈도우 위치 및 윈도우 사이즈와 유사하게 제한될 수도 있다. 예를 들어, 피드백 없이, 오브젝트 검출기 (208) 는 각각의 선택된 윈도우 위치에 대해 모두 10 개의 윈도우 사이즈들 (766a-j) 을 검색할 수도 있다. 그러나, 오브젝트가 제 5 윈도우 사이즈 (766e) 를 갖는 윈도우를 이용하여 최근의 (현재가 아닌) 비디오 프레임에서 검출되었으면, 스캐너 스케일러 (236) 는 현재 캡처된 프레임에 대해 5, 플러스 또는 마이너스 3 의 윈도우 사이즈들, 즉, 윈도우 사이즈들 2-8 만을 선택할 수도 있다. 다시 말해, 제 1 윈도우 사이즈 (766a), 제 9 윈도우 사이즈 (766i), 및 제 10 윈도우 사이즈 (766j) 를 갖는 윈도우들은 최근의 또는 이전 비디오 프레임 (N-1) (222) 으로부터의 피드백에 기초하여 검색되지 않을 수도 있다. 이는 또한 낮은 확률 검색을 없애고 오브젝트 검출의 효율을 증가시킬 수도 있다. 다시 말해, 최근의 비디오 프레임으로부터의 피드백을 이용하는 것은 수행되는 계산들을 감소시키는데 도움이 될 수도 있다. 그렇지 않고, 최근의 비디오 프레임이 타겟 오브젝트를 검출하지 않으면 (즉, 최근에 캡처된 프레임에 대한 검출 및 트래킹 신뢰도 값 (256) 이 검출 및 트래킹 임계치 (254) 보다 작으면), 오브젝트 검출기 (208) 는 사이즈 레벨들의 서브세트를 이용함으로써 검색 범위를 제한하지 않을 수도 있다.

[0067] 도 8 은 오브젝트 트래킹 및 검출 모듈 (804) 의 다른 가능한 구성을 도시하는 블록도이다. 도 8 에 도시된 오브젝트 트래킹 및 검출 모듈 (804) 은 도 2 에 도시된 오브젝트 트래킹 및 검출 모듈 (204) 과 유사한 모듈들을 포함하고 유사한 기능을 수행할 수도 있다. 구체적으로, 도 8 에 도시된 오브젝트 검출기 (808), 모션 트래커 (806), 스캐너 로케이터 (830), 윈도우 위치 선택기 (832), 랜덤화기 (834), 스캐너 스케일러 (836), 분류기 (838), 통합 모듈 (860), 메모리 버퍼 (810), 캡처된 이전 비디오 프레임 (812), 윈도우 (842), 위치 (844), 사이즈 (846), 이진 결정 (848), 트래킹 임계치 (850), 검출 임계치 (852), 검출 및 트래킹 임계치 (854), 검출 신뢰도 값 (840), 트래킹 신뢰도 값 (828), 및 검출 및 트래킹 신뢰도 값 (856) 은 도 2 에 도시된 오브젝트 검출기 (208), 모션 트래커 (206), 스캐너 로케이터 (230), 윈도우 위치 선택기 (232), 랜덤화기 (234), 스캐너 스케일러 (236), 분류기 (238), 통합 모듈 (260), 메모리 버퍼 (210), 캡처된 이전 비디오 프레임 (212), 윈도우 (242), 위치 (244), 사이즈 (246), 이진 결정 (248), 트래킹 임계치 (250), 검출 임계치 (252), 검출 및 트래킹 임계치 (254), 검출 신뢰도 값 (240), 트래킹 신뢰도 값 (228), 및 검출 및 트래킹 신뢰도 값 (256) 에 대응하고 유사한 기능을 가질 수도 있다.

[0068] 또한, 오브젝트 트래킹 및 검출 모듈 (804) 은 타겟 모션 및 트래킹 에러로 인한 지터링 효과를 감소시키는데 이용되는 평활화 모듈 (861) 을 포함할 수도 있다. 다시 말해, 평활화 모듈 (861) 은 트래킹 결과들을 평활화하여, 검색 윈도우로 하여금 위치 (x, y) (844) 및 사이즈 (폭, 높이) (846) 양자 모두에서 보다 평활화된 궤도를 갖게 한다. 평활화 모듈 (861) 은 단순한 이동 평균 (moving average; MA) 필터들 또는 자기 회귀

(auto regression; AR) 필터들일 수 있다. 위치 (844) 및 사이즈 (846) 에 대한 평활화 정도는 상이할 수 있다. 칼만 (Kalman) 필터와 같은 예측 필터들이 또한 위치 (844) 평활화에 적합할 수도 있다. 따라서, 평활화 모듈 (861) 은 입력으로서 평활화되지 않은 위치 (863) 및 평활화되지 않은 사이즈 (865) 를 수신하여, 평활화된 위치 (867) 및 평활화된 사이즈 (869) 를 출력할 수도 있다.

[0069] 도 9 는 평활화 모듈 (961) 을 도시하는 블록도이다. 평활화 모듈 (961) 은 타겟 모션 및 트래킹 에러로 인한 지터링 효과를 감소시키는데 이용될 수도 있는데, 즉, 따라서 트래킹 결과들 (바운딩 박스) 은 위치 (x, y) 및 사이즈 (폭, 높이) 양자 모두에서 보다 평활한 궤도를 갖는다. 일 구성에서, 위치 평활화 필터 (971) 및 사이즈 평활화 필터 (973) 는 입력으로서 평활화되지 않은 위치 (963) 및 평활화되지 않은 사이즈 (965) 를 수신하여 평활화된 위치 (967) 및 평활화된 사이즈 (969) 를 출력하기 위해 자기 회귀 (AR) 모델을 이용하여 구현된다.

[0070] 자기 회귀 (AR) 모델에서,  $X$  는 위치 또는 사이즈의 어느 일방에서 평활화될 변수라고 가정한다. 또한,  $X'$  는 오브젝트 트래커에 의한  $X$  의 출력이라고 하자. 이러한 구성에서, 시간  $t$  에서의  $X$  의 평활화된 필터링,  $X_t$  는 식 (1) 에 따라 설명될 수 있다:

$$X_t = W * X'_t + (1-W) * X_{t-1} \quad (1)$$

[0072] 여기서  $X'_t$  는 시간  $x$  에서의  $X$  의 트래커 출력이며,  $X_{t-1}$  는 시간  $t-1$  에서의  $X$  의 평활화된 결과이고,  $W$  ( $0 \leq W \leq 1$ ) 는 평활화 효과를 제어하는 평활화 가중치이다. 예를 들어,  $X'_t$  는 현재 비디오 프레임 (N) (224) 에 대해 선택된 윈도우 위치 또는 윈도우 사이즈일 수도 있고,  $X_{t-1}$  는 이전 비디오 프레임 (N-1) (222) 에 대해 이용된 윈도우 위치 또는 윈도우 사이즈일 수도 있다.

[0073] 상이한 평활화 가중치,  $W$  가 위치 평활화 필터 (971) 및 사이즈 평활화 필터 (973) 에 이용될 수 있다. 예를 들어, 일 구현에서, 윈도우 위치에 대해서는 보다 적은 평활화 효과가 있으나 윈도우 사이즈에 대해서는 보다 강한 평활화 효과가 있도록  $W_{\text{위치}} = 0.8$  및  $W_{\text{사이즈}} = 0.4$  이다. 이러한 평활화 가중치들의 선택은 보다 적은 트래킹 지연 및 보다 적은 지터링을 생성할 것이다.

[0074] 나아가, 평활화 가중치의 선택은 또한 검출 및 트래킹 신뢰도 값 (856) 이 소정의 임계치 (예를 들어, 검출 및 트래킹 임계치 (854)) 아래로 떨어지는 경우 감소될 수도 있다. 이는 잠재적인 트래킹 및 검출 에러들이 높을 경우 보다 강한 필터링을 야기할 수도 있다. 예를 들어, 낮은 트래킹 신뢰도 (예를 들어, 검출 및 트래킹 신뢰도 값 (856) 이 검출 및 트래킹 임계치 (854) 아래이다) 에 응답하여, 위치 및 사이즈에 대한 평활화 가중치들은 각각  $W_{\text{위치}}=0.65$  및  $W_{\text{사이즈}}=0.2$  로 설정될 수도 있다. 다시 말해, 가중치들 중 하나 또는 양자 모두가 줄어들 수도 있으며, 이는 윈도우 위치 및 사이즈 선택이 현재 비디오 프레임의 윈도우 위치들 및 사이즈들 보다 이전 비디오 프레임들의 윈도우 위치들 및 사이즈들에 보다 크게 기대도록 할 수도 있다.

[0075] 또한, 가중치는 검출 및 트래킹 신뢰도 값 (856) 보다는 트래킹 신뢰도 값 (828) 또는 검출 신뢰도 값 (840) 에 기초할 수도 있다. 예를 들어, 평활화 가중치들,  $W_{\text{위치}}$ , 및  $W_{\text{사이즈}}$  는 트래킹 신뢰도 값 (828) 이 트래킹 임계치 (850) 아래로 떨어지는 것에 응답하여 줄어들 수도 있는데, 즉, 열악한 모션 트래킹에 응답하여 보다 강한 필터링이 이용될 수도 있다. 그렇지 않으면, 평활화 가중치들은 검출 신뢰도 값 (840) 이 검출 임계치 (852) 아래로 떨어지는 것에 응답하여 줄어들 수도 있는데, 즉, 열악한 오브젝트 검출에 응답하여 보다 강한 필터링이 이용될 수도 있다.

[0076] 다른 구성에서, 칼만 필터링이 윈도우 위치를 평활화하는데 이용될 수도 있다. 그러한 구성에서, 필터링은 식 (2) 내지 식 (7) 에 따라 정의될 수도 있다:

$$x_k = F_k x_{k-1} + w_k \quad (2)$$

$$z_k = H x_{k-1} + v_k \quad (3)$$

[0079] 여기서,  $x_{k-1}$  는 시간  $k-1$  에서의 이전 상태이며,  $x_k$  는  $x_k = [x, y, \dot{x}, \dot{y}]$  에 의해 정의된 현재 상태이며, 여기서 (x,y) 는 바운딩 박스 중심 위치이며,  $(\dot{x}, \dot{y})$  은 각각의 방향에서의 속도이다. 또한, 상태 전이

모델,  $F_k$  및 관측 모델,  $H$  는 각각 식 (4) 및 식 (5) 에 의해 정의될 수도 있다:

$$F_k = \begin{bmatrix} 1, & 0, & \Delta t, & 0 \\ 0, & 1, & 0, & \Delta t \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} \quad (4)$$

$$H = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \end{bmatrix} \quad (5)$$

여기서  $\Delta t$  는 조율할 수 있는 파라미터이다. 또한,  $w_k$  는 식 (6) 에 따라 공분산  $Q$  (즉,  $w_k \sim N(0, Q)$ ) 을 갖는 제로평균 다변수 정규 분포로부터 인출되는 것으로 가정되는 프로세스 노이즈이다:

$$Q = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} \sim \sigma_1^2 \quad (6)$$

여기서  $\sigma_1$  는 조율가능한 파라미터이다. 유사하게,  $w_k$  는 식 (7) 에 따른 공분산  $R$  (즉,  $v_k \sim N(0, R)$ ) 을 갖는 제로 평균 가우시안 화이트 노이즈인 것으로 가정되는 관측 노이즈이다:

$$R = \begin{bmatrix} 1, & 0, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & 0 \\ 0, & 0, & 0, & 1 \end{bmatrix} \sim \sigma_2^2 \quad (7)$$

여기서  $\sigma_2$  는 조율할 수 있는 파라미터이다.

도 10 은 모션 트래킹 결과들에서 지터를 평활화하는 방법 (1000) 을 도시하는 흐름도이다. 방법 (1000) 은 전자 디바이스 (102), 예를 들어, 전자 디바이스 (102) 에서의 오브젝트 트래킹 및 검출 모듈 (804) 에 의해 수행될 수도 있다. 전자 디바이스 (102) 는 현재 비디오 프레임 (224) 과 연관된 하나 이상의 윈도우 위치들 및 하나 이상의 윈도우 사이즈들, 예를 들어, 평활화되지 않은 위치 (863) 및 평활화되지 않은 사이즈 (865) 를 결정할 수도 있다 (1002). 전자 디바이스 (102) 는 또한 하나 이상의 윈도우 위치들 및 하나 이상의 윈도우 사이즈들을 필터링하여 하나 이상의 평활화된 윈도우 사이즈들 (867) 및 하나 이상의 평활화된 윈도우 사이즈들 (869) 을 생성할 수도 있다 (1004). 예를 들어, 이는 이동 평균 필터, 자기 회귀 필터, 또는 칼만 필터를 이용하는 것을 포함할 수도 있다. 일 구성에서, 낮은 트래킹 신뢰도 (예를 들어, 검출 및 트래킹 신뢰도 값 (856) 이 검출 및 트래킹 임계치 (854) 아래이다) 에 응답하여, 위치 및 사이즈에 대한 평활화 가중치들은 감소될 수도 있다. 그렇지 않으면, 평활화 가중치들은 검출 신뢰도 값 (840) 또는 트래킹 신뢰도 값 (828) 에 기초하여 감소될 수도 있다. 전자 디바이스는 또한 하나 이상의 평활화된 윈도우 위치들 (867) 및 하나 이상의 평활화된 사이즈들 (869) 에 의해 정의된 하나 이상의 윈도우들을 이용하여 현재 비디오 프레임 (224) 내에서 타겟 오브젝트를 검출할 수도 있다 (1006).

도 11 은 오브젝트 트래킹을 이용하여 화상 프로세싱을 수행하는 방법 (1100) 의 흐름도이다. 방법 (1100) 은 전자 디바이스 (102) 에 의해 수행될 수도 있다. 전자 디바이스 (102) 는 제 1 트래킹 영역 (133) 을 획득할 수도 있다 (1102). 전자 디바이스 (102) 는 또한 제 2 트래킹 영역 (135) 을 획득할 수도 있다 (1104). 일 구성에서, 트래킹 영역들의 각각은 터치스크린 (116) 또는 뷰파인더 (131) 를 이용하여 전자 디바이스 (102) 에 의해 획득될 수도 있다. 일 예로서, 제 2 트래킹 영역 (135) 은 전체 뷰파인더 (131) 를

커버할 수도 있다.

- [0089] 터치스크린 (116) 내의 영역 (통상적으로 정사각형 또는 원형이긴 하나, 다른 형상들이 또한 이용될 수도 있다) 은 사용자에게 의해 정의될 수도 있다. 이러한 영역은 트래킹 영역 또는 포커스 링이라고 지칭될 수도 있다. 포커스 링은 전자 디바이스 (102) 의 사용자가 오브젝트 또는 트래킹을 위한 영역을 빠르게 선택하는 것을 허용하는 사용자 인터페이스 (user interface; UI) 엘리먼트일 수도 있다. 일 예로서, 사용자는 영역 또는 오브젝트 위에 포커스 링을 둠으로써, 포커스 링을 오브젝트에 연관지을 수도 있다. 사용자의 손가락이 터치스크린으로부터 제거되면, 포커스 링이 오브젝트를 트래킹하기 시작할 수도 있다.
- [0090] 포커스 링은 오브젝트 트래킹의 상태 (예를 들어, 오브젝트가 트래킹된다, 오브젝트가 트래킹되지 않는다, 트래킹되긴 하나 오브젝트가 분실되었다) 에 따라 모습을 변화시킬 수도 있다. 포커스 링은 임의의 형상의 오브젝트들의 트래킹을 가능하게 하기 위해 (예를 들어, 원형으로부터 타원형으로 또는 직사각형으로) 형상이 변경되거나 사이즈가 다시 정해질 수도 있다. 일 구성에서, 트래킹된 오브젝트 상의 포커스 링을 터치하는 것은 전자 디바이스 (102) 로 하여금 해당 오브젝트를 트래킹하는 것을 중지하게 할 수도 있다. 포커스 링은 터치스크린 (116) 또는 뷰파인더 (131) 주위의 오브젝트를 따라갈 수도 있다.
- [0091] 전자 디바이스 (102) 는 제 1 트래킹 영역 (133) 을 트래킹하기 시작할 수도 있다 (1106). 전자 디바이스 (102) 는 또한 제 2 트래킹 영역 (135) 을 트래킹하기 시작할 수도 있다 (1108). 제 1 트래킹 영역 (133) 및 제 2 트래킹 영역 (135) 의 중첩 (143) 이 임계치 (145) 를 통과하면 전자 디바이스 (102) 는 화상 프로세싱 (1110) 을 수행할 수도 있다. 구성에 따라, 화상 프로세싱은 중첩 (143) 이 임계치 (145) 위로 되는 경우 또는 중첩 (143) 이 임계치 (145) 아래로 되는 경우 일어날 수도 있다. 화상 프로세싱은 사진을 찍는 것 및/또는 비디오 편집 (예를 들어, 비디오 프레임으로부터 오브젝트를 제거하는 것) 을 수행하는 것을 포함할 수도 있다.
- [0092] 도 12a 는 오브젝트 트래킹을 이용하는 화상 프로세싱의 일 예를 도시한다. 다수의 프레임들 (1253a-b) 이 도시된다. 프레임들 (1253) 은 사전에 레코딩된 비디오 시퀀스 (147) 또는 뷰파인더 (131) 를 통해 뷰잉되는 라이브 프레임들의 일부분일 수도 있다. 프레임 m (1253a) 에서, 제 1 트래킹 영역 (1233a) 은 걸어가는 사람 주위로 도시되고 제 2 트래킹 영역 (1235a) 은 고정되어 있는 나무 주위로 도시된다. 사용자는 걸어가는 사람이 고정되어 있는 나무의 앞에 있으면 사진 (149) 을 찍길 원할 수도 있다. 프레임 m (1253a) 에서, 제 1 트래킹 영역 (1233a) 은 제 2 트래킹 영역 (1235a) 과 중첩되지 않는다 (즉, 중첩 (143) 이 0% 이다). 전자 디바이스 (102) 는 중첩 (143) 이 50% 에 도달하면 화상 프로세싱을 수행하도록 구성될 수도 있다. 이러한 구성에서, 전자 디바이스 (102) 는 중첩 (143) 이 50% 에 도달하면 사진 (149) 을 찍도록 구성될 수도 있다.
- [0093] 프레임 n (1253b) 에서, 프레임 m (1253a) 이후로 시간이 경과되었다. 제 1 트래킹 영역 (1233b) 은 걸어가는 사람에 대해 남아 있고, 제 2 트래킹 영역 (1235b) 은 고정되어 있는 나무에 대해 남아 있다. 걸어가는 사람이 이동했기 때문에, 제 1 트래킹 영역 (1233b) 은 이제 50% 이상으로 제 2 트래킹 영역 (1235b) 과 중첩된다 (1243a). 따라서, 중첩 (1243a) 이 50% 에 도달하면, 전자 디바이스 (102) 는 사진 (149) (이 경우에, 고정되어 있는 나무 앞에 있는 걸어가는 사람의 사진 (149)) 을 찍도록 구성된다.
- [0094] 도 12b 는 또한 오브젝트 트래킹을 이용하는 화상 프로세싱의 일 예를 도시한다. 다수의 프레임들 (1253c-d) 이 도시된다. 프레임들 (1253) 은 사전에 레코딩된 비디오 시퀀스 (147) 또는 뷰파인더 (131) 를 통해 뷰잉되는 라이브 프레임들의 일부분일 수도 있다. 프레임 m (1253c) 에서, 제 1 트래킹 영역 (1233c) 은 걸어가는 사람 주위로 도시되고 액션 라인 (1287) 은 고정되어 있는 나무 주위로 도시된다. 액션 라인 (1287) 은 수직 라인, 수평 라인, 또는 (곡선의 라인과 같은) 다른 유형의 라인일 수도 있다. 제 1 트래킹 영역 (1233c) 및 액션 라인 (1287) 양자 모두는 사용자에게 의해 설정될 수도 있다. 걸어가는 사람이 액션 라인 (1287) 을 건너면 (즉, 중첩 (1243b) 이 일어나는 경우), 사용자는 사진 (또는 사진들의 버스트) 를 찍거나 다른 비디오 프로세싱을 수행하길 원할 수도 있다.
- [0095] 프레임 n (1253d) 에서, 프레임 m (1253c) 이후로 시간이 경과되었다. 제 1 트래킹 영역 (1233d) 은 걸어가는 사람에 대해 남아 있고, 액션 라인 (1287) 은 고정되어 있는 나무에 대해 남아 있다. 걸어가는 사람이 이동했기 때문에, 제 1 트래킹 영역 (1233d) 은 이제 액션 라인 (1287) 과 중첩된다 (1243b). 제 1 트래킹 영역 (1233d) 이 액션 라인 (1287) 과 교차하면, 전자 디바이스 (102) 는 사진 (149) 을 찍거나 다른 화상 프로세싱을 수행하도록 구성될 수도 있다.



- [0096] 도 13 은 오브젝트 트래킹을 이용하는 화상 프로세싱의 다른 예를 도시한다. 다수의 프레임들 (1253a-b) 이 도시된다. 프레임들 (1253) 은 사전에 레코딩된 비디오 시퀀스 (147) 또는 뷰파인더 (131) 를 통해 뷰잉되는 라이브 프레임들의 일부분일 수도 있다. 프레임 m (1353a) 에서, 제 1 트래킹 영역 (1333a) 은 걸어가는 사람 주위로 도시되고 제 2 트래킹 영역 (1335a) 은 고정되어 있는 나무 주위로 그리고 나무 주의의 영역에 도시된다. 걸어가는 사람이 더 이상 뷰 (예를 들어, 자연 샷) 에 있지 않으면 사용자는 사진 (149) 을 찍기를 원할 수도 있다. 프레임 m (1353a) 에서, 제 2 트래킹 영역 (1335a) 은 제 1 트래킹 영역 (1333a) 과 완전히 중첩된다 (즉, 중첩 (1343) 이 100% 이다). 전자 디바이스 (102) 는 중첩 (1343) 이 0% 에 도달하면 화상 프로세싱을 수행하도록 구성될 수도 있다. 이러한 구성에서, 전자 디바이스 (102) 는 중첩 (1343) 이 0% 에 도달하면 사진 (149) 을 찍도록 구성될 수도 있다.
- [0097] 프레임 n (1353b) 에서, 프레임 m (1353a) 이후로 시간이 경과되었다. 제 1 트래킹 영역 (1333b) 은 걸어가는 사람에 대해 남아 있고, 제 2 트래킹 영역 (1335b) 은 고정되어 있는 나무에 대해 남아 있다. 걸어가는 사람이 이동했기 때문에, 제 1 트래킹 영역 (1333a) 은 제 2 트래킹 영역 (1335b) 과 더 이상 중첩되지 않는다. 따라서, 중첩 (1343) 이 0% 에 도달하면, 전자 디바이스 (102) 는 사진 (149) (이 경우에, 걸어가는 사람이 없는 고정되어 있는 나무의 사진 (149)) 을 찍도록 구성된다.
- [0098] 도 14 는 오브젝트 트래킹을 이용하여 비디오 시퀀스 (147) 에 대해 화상 프로세싱을 수행하는 방법 (1400) 의 흐름도이다. 방법 (1400) 은 전자 디바이스 (102) 에 의해 수행될 수도 있다. 사용자는 화상 프로세싱을 위해 전자 디바이스 (102) 상에서 다수의 트래킹 영역들을 선택할 수도 있다. 전자 디바이스 (102) 는 제 1 트래킹 영역 (133) 이 비디오 시퀀스 (147) 제 1 프레임에서 임계치 (145) 보다 많이 제 2 트래킹 영역 (135) 과 중첩 (143) 된다고 결정할 수도 있다 (1402). 전자 디바이스 (102) 는 비디오 시퀀스 (147) 로부터 제 2 프레임을 선택할 수도 있다 (1404). 제 2 프레임은 제 2 프레임의 제 1 트래킹 영역 (133) 이 제 1 프레임의 제 1 트래킹 영역 (133) 에 대응하는, 제 2 프레임의 대체 영역과 중첩되지 않도록 선택될 수도 있다. 제 2 프레임의 대체 영역은 제 1 프레임에서 제 1 트래킹 영역 (133) 의 포지션을 반영할 수도 있다. 따라서, 대체 영역은 제 1 트래킹 영역 (133) 뒤쪽의 배경을 보여줄 수도 있다. 제 2 프레임은 제 1 프레임 전에 또는 후에 나오는 프레임일 수도 있다.
- [0099] 전자 디바이스 (102) 는 제 1 프레임의 제 1 트래킹 영역 (133) 을 제 2 프레임의 대응하는 대체 영역으로 대체할 수도 있다 (1406). 전자 디바이스 (102) 는 편집된 제 1 프레임을 편집된 비디오 시퀀스 (151) 의 일부분으로서 저장할 수도 있다 (1408).
- [0100] 도 15 는 전자 디바이스 (102) 상에 디스플레이된 편집되지 않은 비디오 시퀀스 (1547) 및 편집된 비디오 시퀀스 (1551) 양자 모두의 다수의 프레임들 (1553a-d) 을 도시한다. 오브젝트 트래킹을 이용하는 화상 프로세싱이 편집되지 않은 비디오 시퀀스 (1547) 에 대해 수행되어 편집된 비디오 시퀀스 (1551) 를 획득할 수도 있다. 편집되지 않은 비디오 시퀀스 (1547) 의 프레임 m (1553a), 프레임 n (1553b), 및 프레임 o (1553c) 가 도시된다. 프레임들 (1553a-c) 이 순차적이긴 하나 (프레임 n (1553b) 은 프레임 m (1553a) 다음에 나온다), 추가적인 프레임들 (미도시) 이 프레임들 (1553a-c) 사이에 나올 수도 있다 (예를 들어, 프레임 n (1553b) 은 프레임 m (1553a) 에 바로 뒤따르는 프레임이 아닐 수도 있다).
- [0101] 프레임 m (1553a) 은 걸어가는 사람 및 고정되어 있는 나무를 포함한다. 사용자는 걸어가는 사람을 포함하는 제 1 트래킹 영역 (1533a) 및 고정되어 있는 나무와 걸어가는 사람을 포함하는 제 2 트래킹 영역 (1535a) 을 전자 디바이스 (102) 를 이용하여 선택할 수도 있다. 일 구성에서, 제 2 트래킹 영역 (1535a) 은 고정되어 있는 것으로 구성될 수도 있다. 사용자는 또한 편집되지 않은 비디오 시퀀스 (1547) 로부터 걸어가는 사람을 제거하도록 전자 디바이스 (102) 를 구성할 수도 있다.
- [0102] 화상 프로세싱은 제 1 트래킹 영역 (1533) 이 제 2 트래킹 영역 (1535) 과 중첩되지 않으면 프레임 m (1553a) 의 제 1 트래킹 영역 (1533a) 을 다른 프레임 (1553) 으로부터의 대체 영역 (1555) 으로 교체하도록 구성될 수도 있다. 다시 말해, 프레임 m (1553a) 에서의 걸어가는 사람은 걸어가는 사람이 배경에 노출되도록 충분히 이동했으면 걸어가는 사람 뒤쪽의 배경으로 대체될 수도 있다. 이 구성에서, 화상 프로세싱은 중첩 (143) 이 0% 에 도달하면 수행될 수도 있다.
- [0103] 프레임 n (1553b) 은 (이동하는 제 1 트래킹 영역 (1533b) 에 의해 둘러싸이는) 걸어가는 사람 및 (고정되어 있는 제 2 트래킹 영역 (1535b) 에 의해 둘러싸이는) 고정되어 있는 나무를 포함한다. 프레임 n (1553b) 의 제 1 트래킹 영역 (1533b) 이 프레임 n (1553b) 의 제 2 트래킹 영역 (1535b) 과 중첩되기 때문에, 프레임 n



(1553b)은 프레임 m (1553a)에서의 대체를 위한 적합한 프레임으로 선택되지 않을 수도 있다.

[0104] 프레임 o (1553c)은 (이동하는 제 1 트랙킹 영역 (1533c)에 의해 둘러싸이는) 걸어가는 사람 및 (고정되어 있는 제 2 트랙킹 영역 (1535c)에 의해 둘러싸이는) 고정되어 있는 나무를 포함한다. 프레임 o (1553c)의 제 1 트랙킹 영역 (1533c)이 프레임 o (1553c)의 제 2 트랙킹 영역 (1535c)과 중첩되지 않기 때문에, 프레임 o (1553c)는 프레임 m (1553a)에서의 대체를 위해 선택될 수도 있다. 프레임 o (1553c)는 대체 영역 (1555)을 포함한다. 대체 영역 (1555)은 프레임 m (1553a)의 제 1 트랙킹 영역 (1533a)과 대응할 수도 있다. 따라서, 대체 영역 (1555)은 프레임 m (1553a)에서 모호한, 걸어가는 사람 뒤쪽의 배경을 포함할 수도 있다. 화상 프로세싱은 프레임 m (1553a)의 제 1 트랙킹 영역 (1533a)을 프레임 o (1553c)의 대체 영역 (1555)으로 대체할 수도 있다. 따라서, 편집된 비디오 시퀀스 (1551)에서, 프레임 m (1553d)은 걸어가는 사람이 제거된 것으로 도시된다.

[0105] 도 16은 전자 디바이스 디바이스 (1602)내에 포함될 수도 있는 소정의 컴포넌트들을 도시한다. 전자 디바이스 (1602)는, 도 1에서 도시된 전자 디바이스 (102)와 같은, 모바일 스테이션, 사용자 장비 (UE), 액세스 포인트 등일 수도 있다. 전자 디바이스 (1602)는 프로세서 (1603)를 포함한다. 프로세서 (1603)는 범용 단일칩 또는 다중칩 마이크로프로세서 (예를 들어, ARM), 특수 목적용 마이크로프로세서 (예를 들어, 디지털 신호 프로세서 (digital signal processor; DSP)), 마이크로제어기, 프로그램가능 게이트 어레이 등일 수도 있다. 프로세서 (1603)는 중앙 처리 장치 (CPU)로 지칭될 수도 있다. 전자 디바이스 (1602)에 단지 단일 프로세서 (1603)로 도시되었으나, 대안적인 구성에서, 프로세서들 (1603) (예를 들어, ARM 및 DSP)의 조합이 이용될 수 있다.

[0106] 전자 디바이스 (1602)는 또한 메모리 (1605)를 포함한다. 메모리 (1605)는 전자적 정보를 저장할 수 있는 임의의 전자 컴포넌트일 수도 있다. 메모리 (1605)는 랜덤 액세스 메모리 (random access memory; RAM), 판독 전용 메모리 (read-only memory; ROM), 자기 디스크 스토리지 매체, 광학 스토리지 매체, RAM의 플래시 메모리 디바이스, 프로세서와 함께 포함된 온보드 메모리, EPROM 메모리, EEPROM 메모리, 레지스터들 등, 및 이들의 조합으로서 구체화될 수도 있다.

[0107] 데이터 (1607a) 및 명령들 (1609a)이 메모리 (1605)에 저장될 수도 있다. 명령들 (1609a)은 본원에서 개시된 방법들을 구현하기 위해 프로세서 (1603)에 의해 실행가능할 수도 있다. 명령들 (1609a)을 실행하는 것은 메모리 (1605)에 저장되어 있는 데이터 (1607a)의 이용을 수반한다. 프로세서 (1603)가 명령들 (1609a)을 실행하는 경우, 명령들 (1609b)의 다양한 부분들이 프로세서 (1603) 상으로 로딩될 수도 있고, 데이터 (1607a)의 다양한 부분들이 프로세서 (1603) 상으로 로딩될 수도 있다.

[0108] 전자 디바이스 (1602)는 또한 전자 디바이스 (1602)로 그리고 전자 디바이스 (1602)로부터의 신호들의 송수신을 가능하게 하는 송신기 (1611) 및 수신기 (1613)를 포함할 수도 있다. 송신기 (1611) 및 수신기 (1613)는 집합적으로 송수신기 (1615)라고 지칭될 수도 있다. 안테나 (1617)가 송수신기 (1615)에 전기적으로 커플링될 수도 있다. 전자 디바이스는 (도시되지 않은) 복수의 송신기들, 복수의 수신기들, 복수의 송수신기들 및/또는 추가 안테나들을 또한 포함할 수도 있다.

[0109] 전자 디바이스 (1602)는 디지털 신호 프로세서 (DSP) (1621)를 포함할 수도 있다. 전자 디바이스 (1602)는 또한 통신 인터페이스 (1623)를 포함할 수도 있다. 통신 인터페이스 (1623)는 사용자가 전자 디바이스 (1602)와 상호작용하는 것을 허용할 수도 있다.

[0110] 전자 디바이스 (1602)의 여러 컴포넌트들은 하나 이상의 버스들에 의해 함께 커플링될 수도 있는데, 상기 버스들은 전력 버스, 제어 신호 버스, 상태 신호 버스, 데이터 버스 등을 포함할 수도 있다. 명확화를 위해, 여러 버스들은 도 16에서 버스 시스템 (1619)으로서 도시된다.

[0111] 본원에서 설명된 기술들은 직교 다중화 기법에 기초한 통신 시스템들을 포함하여 다양한 통신 시스템들에 대해 사용될 수도 있다. 이러한 통신 시스템들의 실시형태들은 직교 주파수 분할 다중 액세스 (Orthogonal Frequency Division Multiple Access; OFDMA) 시스템들, 단일-캐리어 주파수 분할 다중 액세스 (Single-Carrier Frequency Division Multiple Access; SC-FDMA) 시스템들 등을 포함한다. OFDMA 시스템은 직교 주파수 분할 다중화 (orthogonal frequency division multiplexing; OFDM)를 활용하는데, 이것은 전체 시스템 대역폭을 복수의 직교 서브캐리어들로 분할하는 변조 기술이다. 이들 서브캐리어들은 톤들, 빈들 등으로 또한 칭해질 수도 있다. OFDM을 통해, 각각의 서브 캐리어는 데이터와 함께 독립적으로 변조될 수도 있다. SC-FDMA 시스템은 시스템 대역폭에 걸쳐 분배된 서브캐리어들 상에서 전송되는 인터리브된 FDMA

(interleaved FDMA; IFDMA), 인접한 서브캐리어들의 블록 상에서 전송되는 국소화된 FDMA (localized FDMA; LFDMA), 또는 인접한 서브캐리어들의 복수의 블록들 상에서 전송되는 향상된 FDMA (enhanced FDMA; EFDMA)를 활용할 수도 있다. 일반적으로, 변조 심볼들은 OFDM에 의해 주파수 도메인에서 SC-FDMA에 의해 시간 도메인에서 전송된다.

[0112] 본 개시물에 따르면, 전자 디바이스에서의 회로는 이전 비디오 프레임과 현재 비디오 프레임을 비교함으로써 현재 비디오 프레임에 대해 모션-기반 트래킹을 수행하도록 적응될 수도 있다. 동일한 회로, 상이한 회로, 동일하거나 상이한 회로의 제 2 섹션이 트래킹된 파라미터에 기초하여 현재 비디오 프레임에서 오브젝트 검출을 수행하도록 적응될 수도 있다. 제 2 섹션은 유리하게는 제 1 섹션에 커플링될 수도 있거나, 제 1 섹션과 동일한 회로에 구체화될 수도 있다. 또한, 동일한 회로, 상이한 회로, 동일하거나 상이한 회로의 제 3 부분이 위에서 설명된 기능을 제공하는 회로(들) 또는 회로(들)의 섹션(들)의 구성을 제어하도록 적응될 수도 있다.

[0113] 용어 "결정하기"는 매우 다양한 작동들을 망라하므로, "결정하기"는 산출하기, 계산하기, 처리하기, 도출하기, 조사하기, 검색하기 (예를 들어, 테이블, 데이터베이스, 또는 다른 데이터 구조 내 검색하기), 확인하기 등을 포함할 수 있다. 또한, "결정하기"는 수신하기 (예를 들어, 정보 수신하기), 액세스하기 (예를 들어, 메모리 내의 데이터에 액세스하기) 등을 포함할 수 있다. 또한, "결정하기"는 해결하기, 선택하기, 고르기, 설정하기 등을 포함할 수 있다.

[0114] 구절 "~에 기초하는"은 달리 명백히 명시되지 않는 한 "오직 ~에만 기초하는"을 의미하지 않는다. 다시 말해, 구절 "~에 기초하는"은 "오직 ~에만 기초하는" 및 "적어도 ~에 기초하는" 양자 모두를 말한다.

[0115] 용어 "프로세서"는 범용 프로세서, 중앙 처리 유닛 (CPU), 마이크로프로세서, 디지털 신호 프로세서 (DSP), 제어기, 마이크로제어기, 상태 머신 등을 포괄하도록 광의적으로 해석되어야 한다. 어떤 상황을 하에서, "프로세서"는 주문형 반도체 (ASIC), 프로그램가능 로직 디바이스 (PLD), 필드 프로그램가능 게이트 어레이 (FPGA) 등을 지칭할 수도 있다. 용어 "프로세서"는 컴퓨팅 디바이스들의 조합, 예를 들어, DSP와 마이크로프로세서의 조합, 복수의 마이크로프로세서들, DSP 코어와 연계한 하나 이상의 마이크로프로세서들, 또는 임의의 다른 이러한 구성을 지칭할 수도 있다.

[0116] 용어 "메모리"는 전자적 정보를 저장할 수 있는 임의의 전자적 컴포넌트를 포괄하도록 광의적으로 해석되어야 한다. 용어 메모리는 여러 형태들의 프로세서 판독 가능한 매체, 예컨대 랜덤 액세스 메모리 (RAM), 판독-전용 메모리 (ROM), 비휘발성 랜덤 액세스 메모리 (NVRAM), 프로그램가능 판독 전용 메모리 (PROM), 소거가 가능한 프로그램가능 판독 전용 메모리 (EPROM), 전기적으로 소거가능한 PROM (EEPROM), 플래시 메모리, 자기 또는 광학 데이터 스토리지, 레지스터 등을 가리킬 수도 있다. 메모리는, 프로세서가 이 메모리로부터 정보를 판독하고 이 메모리로 정보를 기록할 수 있다면, 프로세서와 전자적으로 통신하고 있다고 말해진다. 프로세서에 일체형인 메모리는 프로세서와 전자적으로 통신한다.

[0117] 용어, "명령들" 및 "코드"는 임의의 형태의 컴퓨터-판독가능 명령문(들)을 포함하도록 광의적으로 해석되어야 한다. 예를 들어, "명령들" 및 "코드"의 용어들은 하나 이상의 프로그램들, 루틴들, 서브루틴들, 함수들, 절차들 등을 가리킬 수도 있다. "명령들" 및 "코드"들은 단일의 컴퓨터-판독가능 명령문 또는 많은 컴퓨터-판독가능 명령문을 포함할 수도 있다.

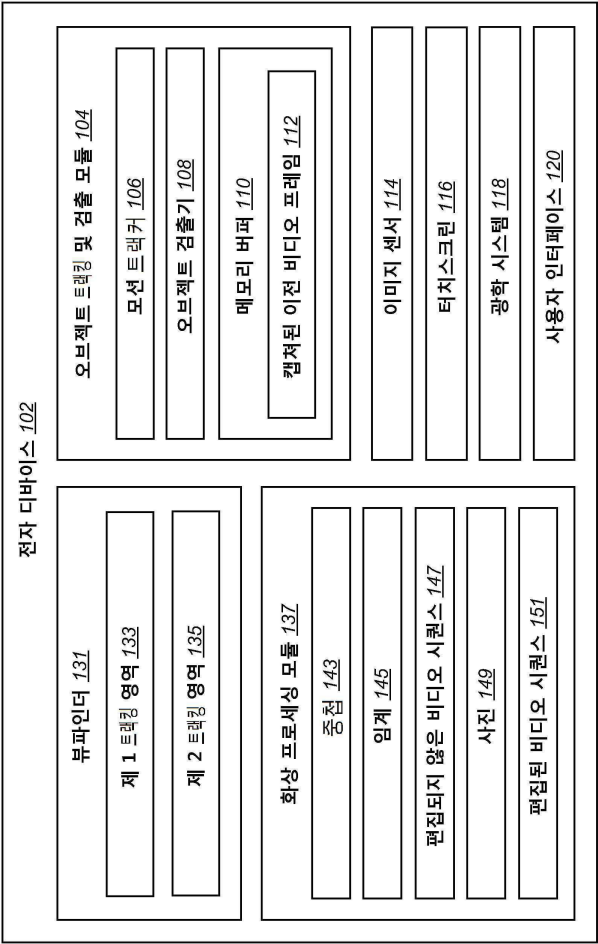
[0118] 본원에서 설명된 기능들은 하드웨어에 의해 실행되는 소프트웨어 또는 펌웨어에서 구현될 수도 있다. 상기 기능들은 컴퓨터 판독가능 매체 상에 하나 이상의 명령들로서 저장될 수도 있다. 용어 "컴퓨터-판독가능 매체" 또는 "컴퓨터-프로그램 제품"은 컴퓨터 또는 프로세서에 의해 액세스될 수 있는 임의의 유형의 저장 매체를 가리킨다. 비제한적인 예로서, 이러한 컴퓨터 판독 가능한 매체는 RAM, ROM, EEPROM, CD-ROM 또는 다른 광학 디스크 스토리지, 자기 디스크 스토리지 또는 다른 자기 스토리지 디바이스들, 또는 요구되는 프로그램 코드를 명령들 또는 데이터 구조들의 형태로 이송 또는 저장하기 위해 사용될 수 있으며 컴퓨터에 의해 액세스될 수 있는 임의의 다른 매체를 포함할 수도 있다. 본원에서 사용된 디스크 (disk) 및 디스크 (disc)는 컴팩트 디스크 (CD), 레이저 디스크, 광학 디스크, DVD (digital versatile disc), 플로피 디스크, 및 Blu-ray<sup>®</sup> 디스크를 포함하며, 여기서 디스크 (disk)들은 보통 데이터를 자기적으로 재생하는데 반해, 디스크 (disc)들은 레이저를 이용하여 데이터를 광학적으로 재생한다. 컴퓨터-판독가능 매체들은 유형적이고 일시적이지 않을 수도 있음에 유의해야 한다. 용어 "컴퓨터-판독가능 제품"은 컴퓨팅 디바이스 또는 프로세서에 의해 실행, 처리, 또는 컴퓨팅될 수도 있는 코드 또는 명령들 (예를 들어, "프로그램")과 조합하는 컴퓨팅 디바이스 또는 프로세서를 말한다. 본원에서 이용된 바와 같은 용어 "코드"는 소프트웨어, 명령들, 코드들, 또는 컴퓨팅

디바이스나 프로세서에 의해 실행될 수 있는 데이터를 지칭할 수도 있다.

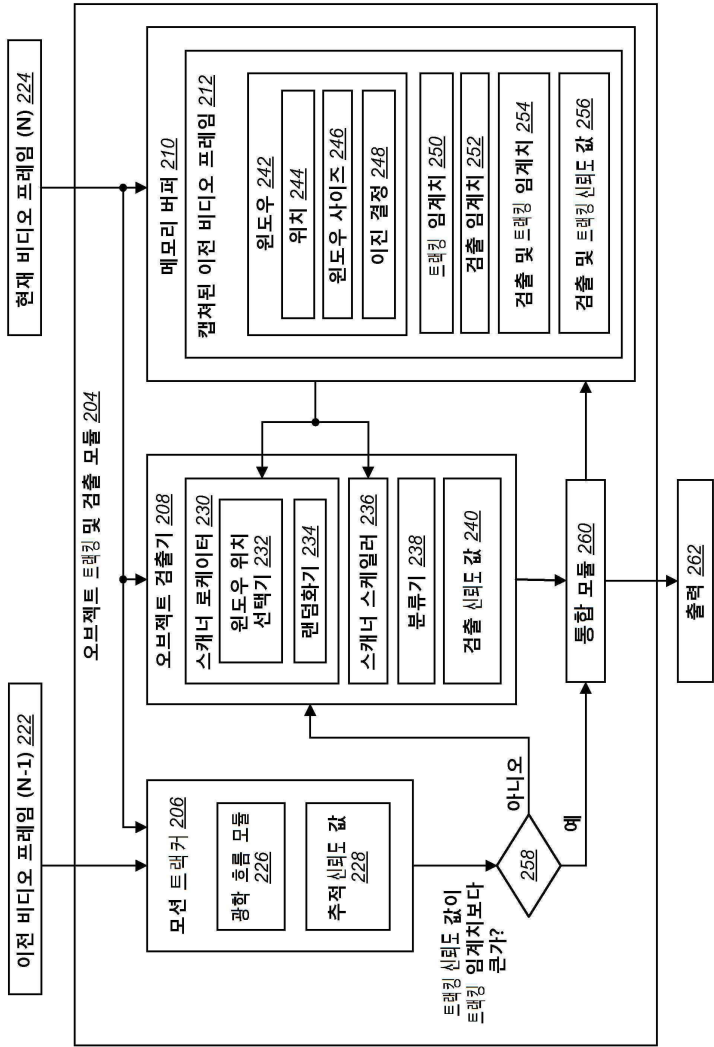
- [0119] 소프트웨어 또는 명령들은 또한 송신 매체를 통해 송신될 수도 있다. 예를 들어, 소프트웨어가 동축 케이블, 광 섬유 케이블, 연선, 디지털 가입자 회선 (DSL), 또는 적외선, 무선, 및/또는 마이크로파와 같은 무선 기술들을 이용하여 웹사이트, 서버, 또는 다른 원격 소스로부터 송신되는 경우, 동축 케이블, 광 섬유 케이블, 연선, DSL, 또는 적외선, 무선, 및 마이크로파와 같은 무선 기술들은 송신 매체의 정의 내에 포함된다.
- [0120] 본원에 개시된 방법들은 설명된 방법을 달성하기 위한 하나 이상의 단계들 또는 작동들을 포함한다. 방법 단계들 및/또는 액션들은 청구항들의 범위를 벗어나지 않으면서 서로 상호교환될 수도 있다. 다시 말해, 설명된 방법들의 적절한 동작을 위해 단계들 또는 작동들의 특정한 순서가 요구되지 않는 한, 특정한 단계들 및/또는 작동들의 순서 및/또는 이용은 청구항들의 범위를 벗어나지 않으면서 수정될 수도 있다.
- [0121] 또한, 도 2a, 도 2b, 도 3 내지 도 6, 도 10, 도 11, 및 도 14 에서 도시된 것들과 같이, 본원에서 설명된 방법들 및 기술들을 수행하기 위한 모듈들 및/또는 다른 적절한 수단들은 다운로드될 수 있고 및/또는 디바이스에 의해 다르게 획득될 수도 있음을 주지해야 한다. 예를 들어, 본원에서 설명된 방법들을 수행하기 위한 수단의 전송을 용이하게 하기 위한 서버에 디바이스가 커플링될 수도 있다. 그렇지 않고, 본원에서 설명된 여러 방법들은 스토리지 수단 (예를 들어, 랜덤 액세스 메모리 (RAM), 판독-전용 메모리 (ROM), 콤팩트 디스크 (CD) 또는 플로피디스크 등과 같은 물리적 저장 매체)을 통해 제공될 수 있고, 따라서, 스토리지 수단이 디바이스에 커플링되거나 제공되면, 디바이스는 여러 방법들을 얻을 수도 있다.
- [0122] 청구항들은 상기에서 예시된 정확한 구성 및 컴포넌트들로 제한되지 않는 것으로 이해되어야 한다. 청구항들의 범위를 벗어나지 않으면서 본원에 개시된 시스템들, 방법들, 및 장치들의 배열, 동작, 및 세부사항들에 다양한 수정, 변경, 및 변형이 이루어질 수도 있다.

도면

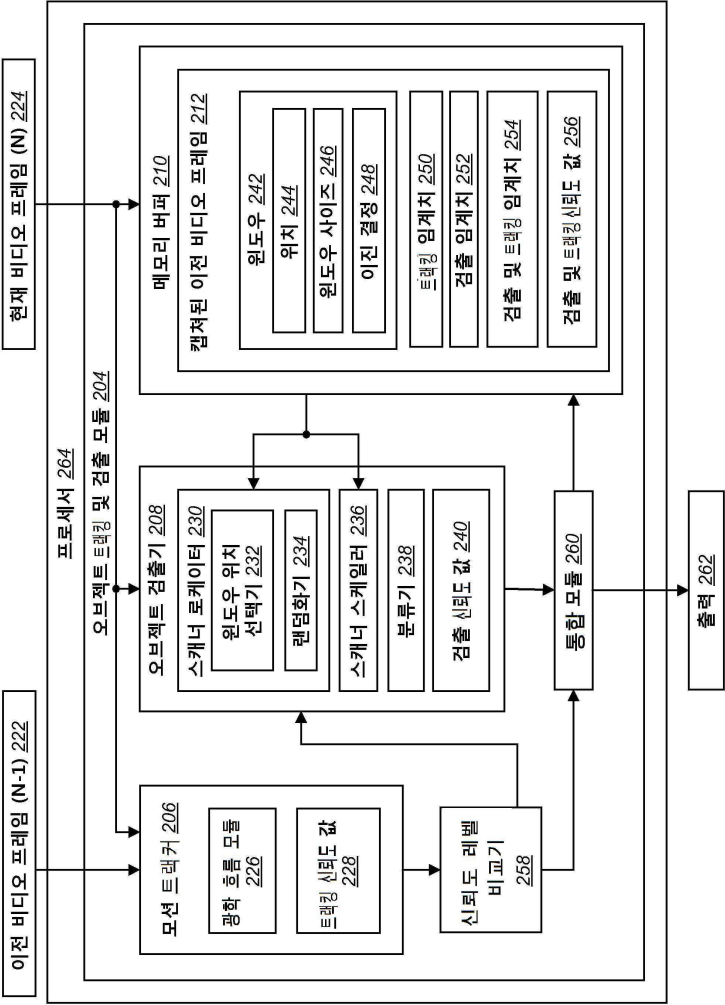
도면1



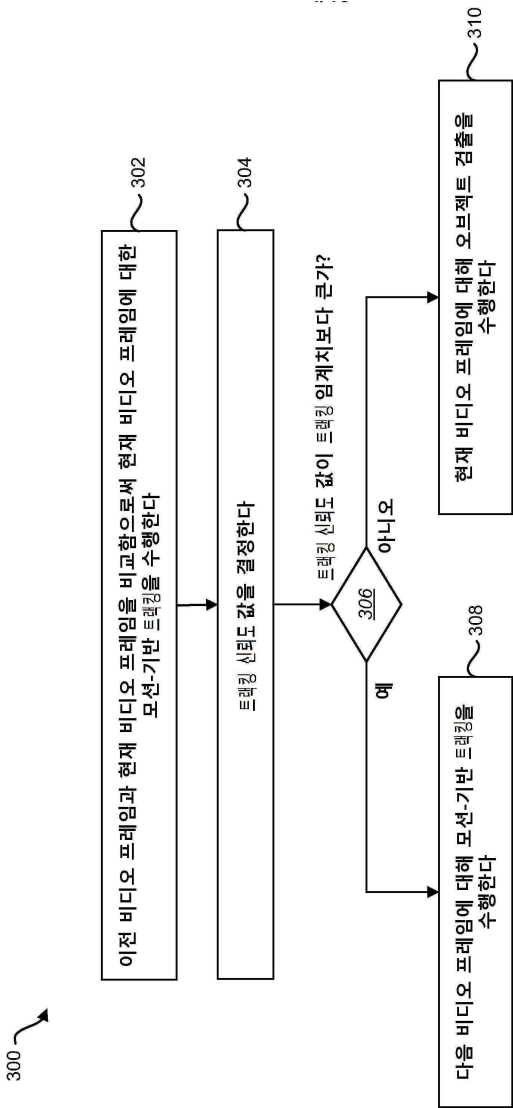
도면 2a



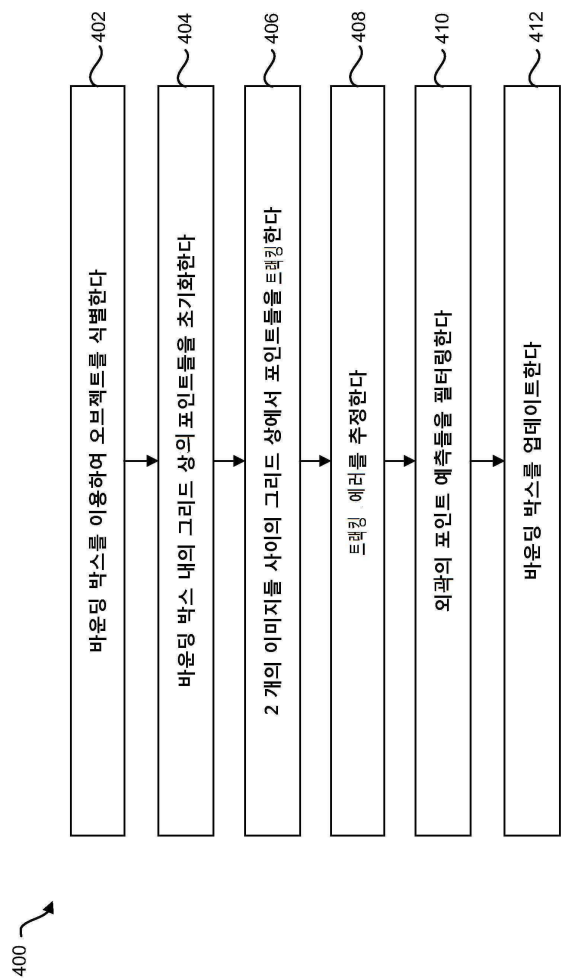
도면2b



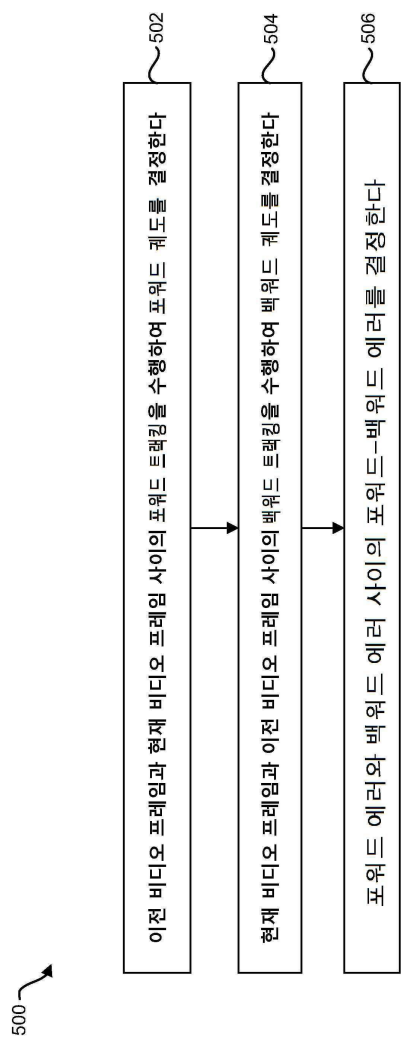
도면3



도면4

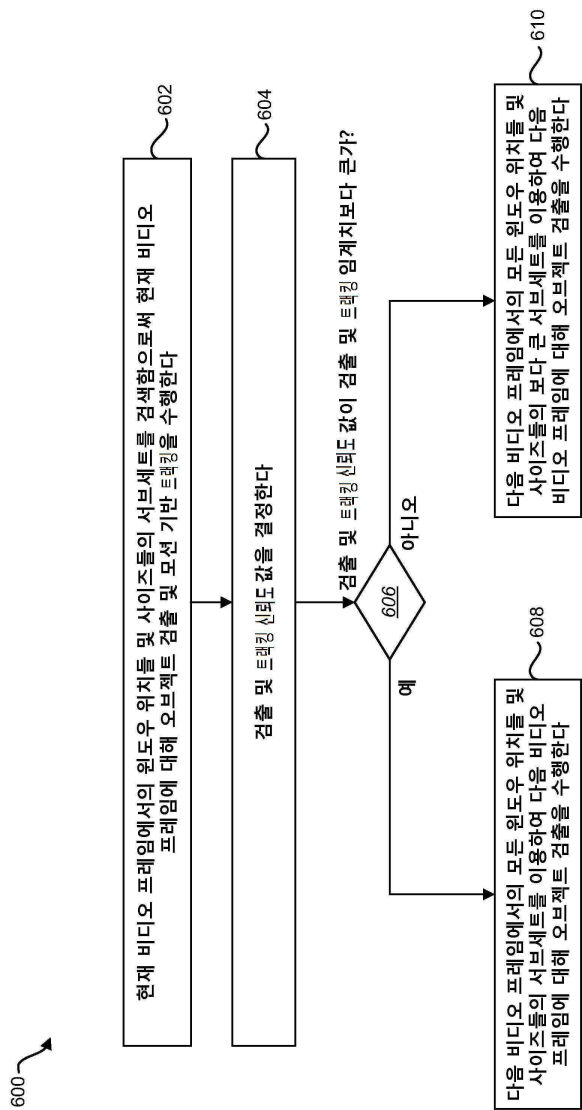


도면5

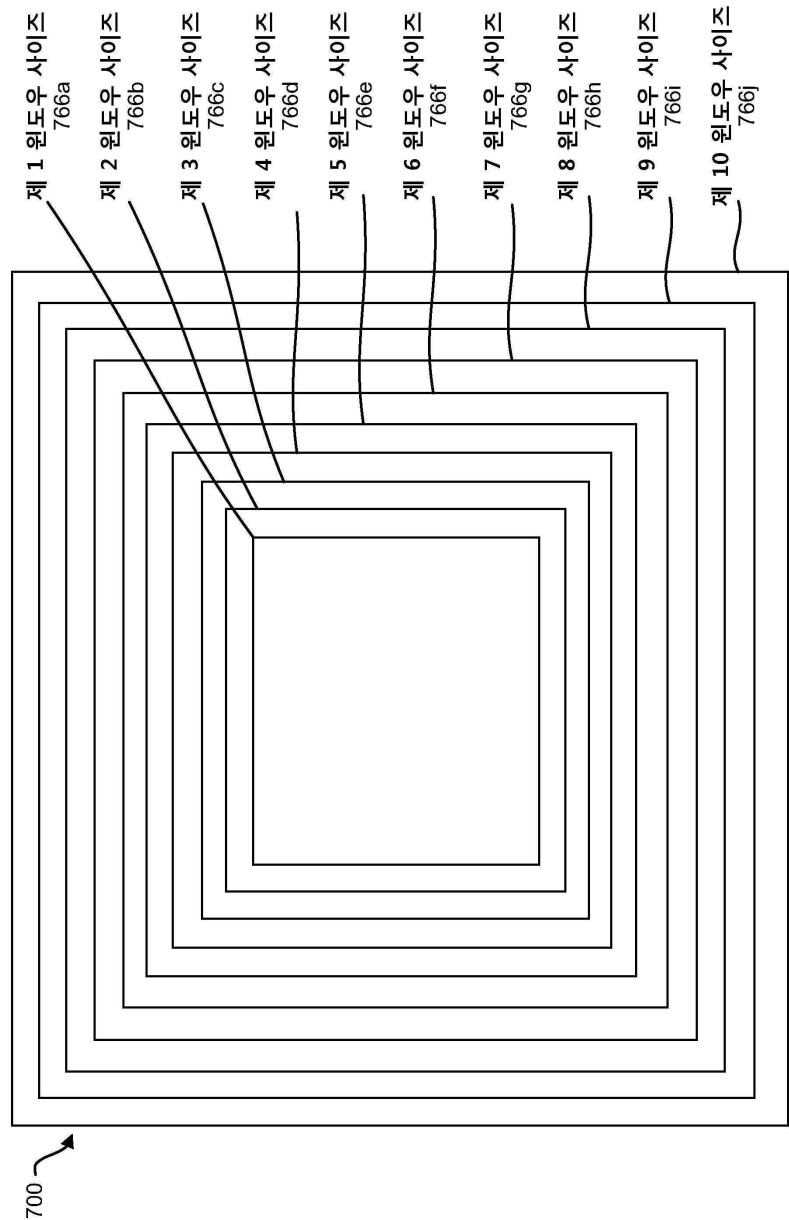




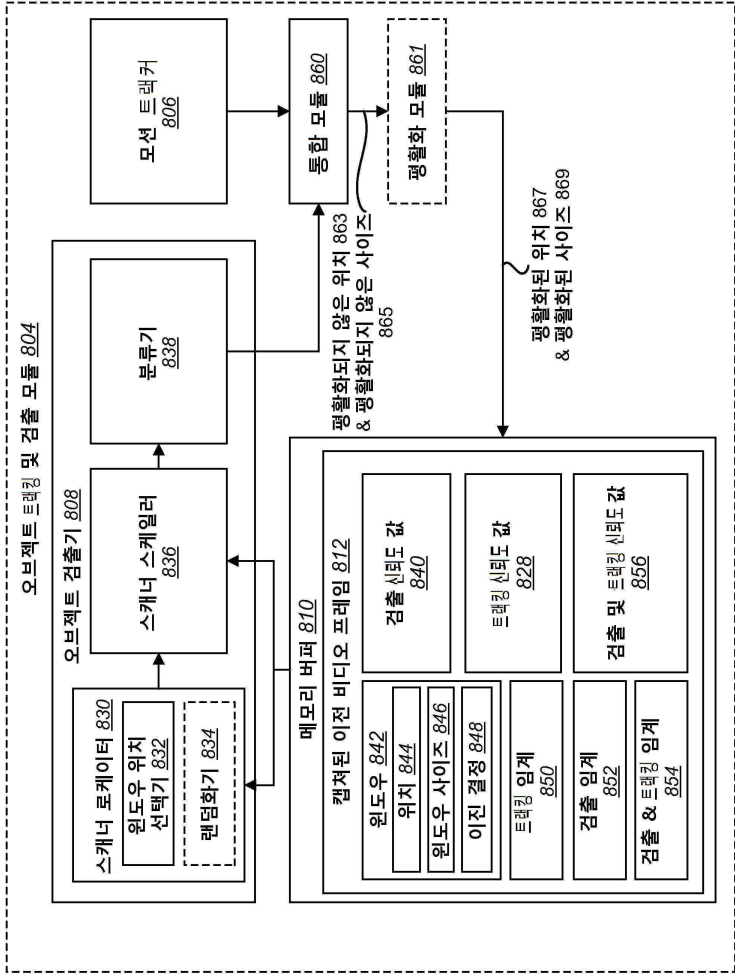
도면6



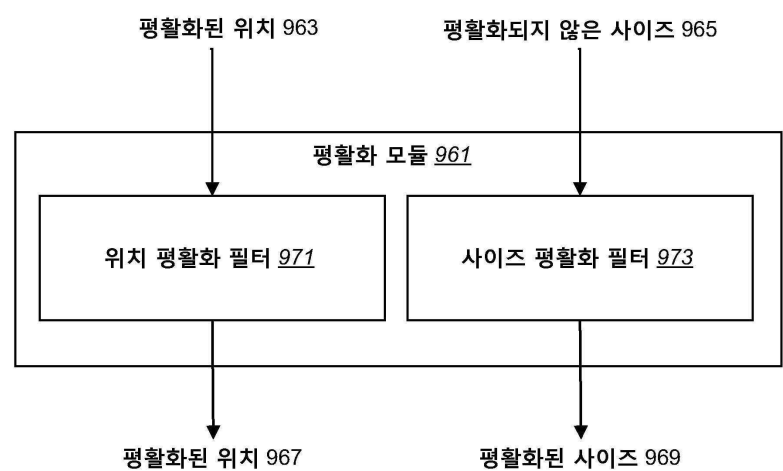
도면7



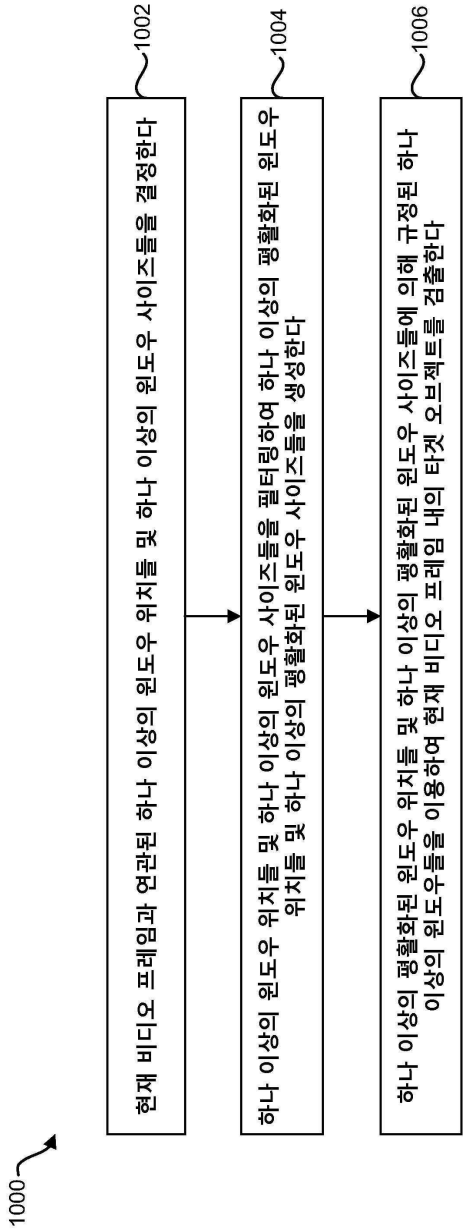
도면8



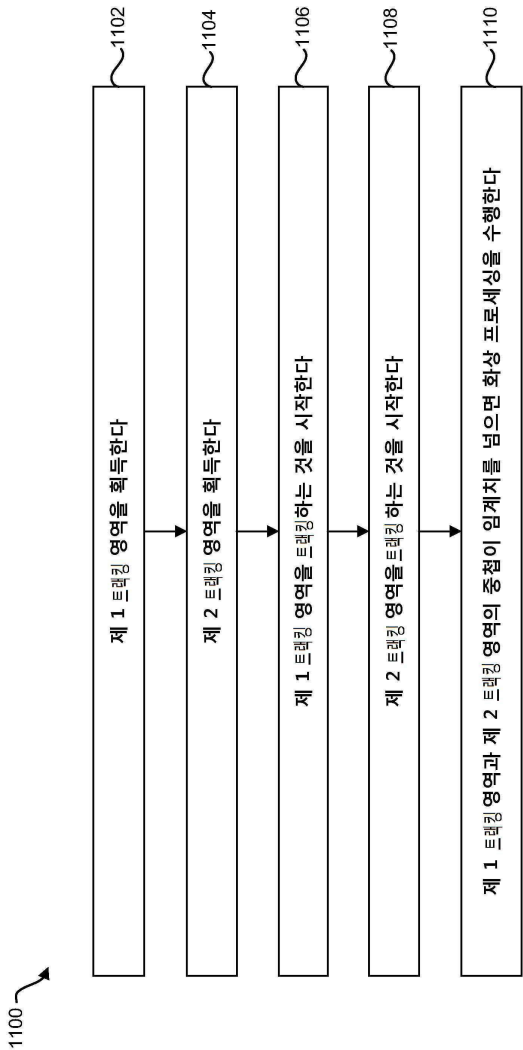
도면9



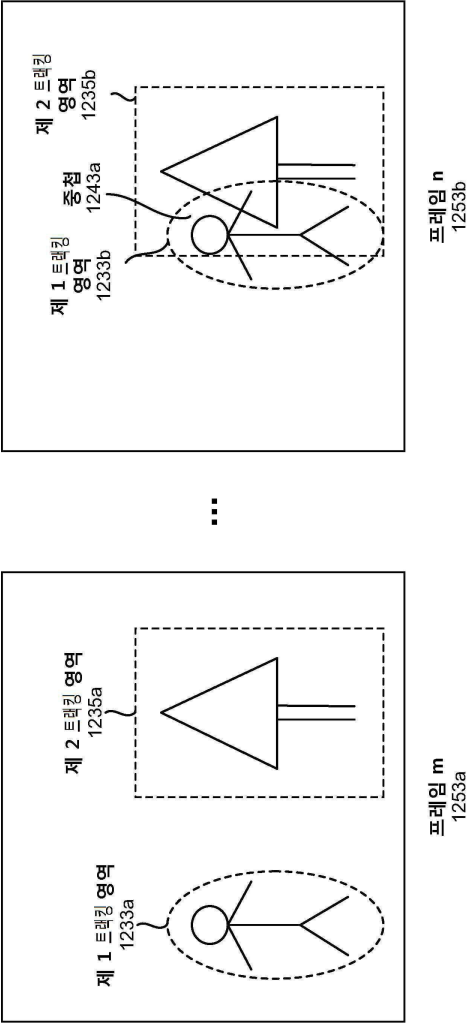
도면10



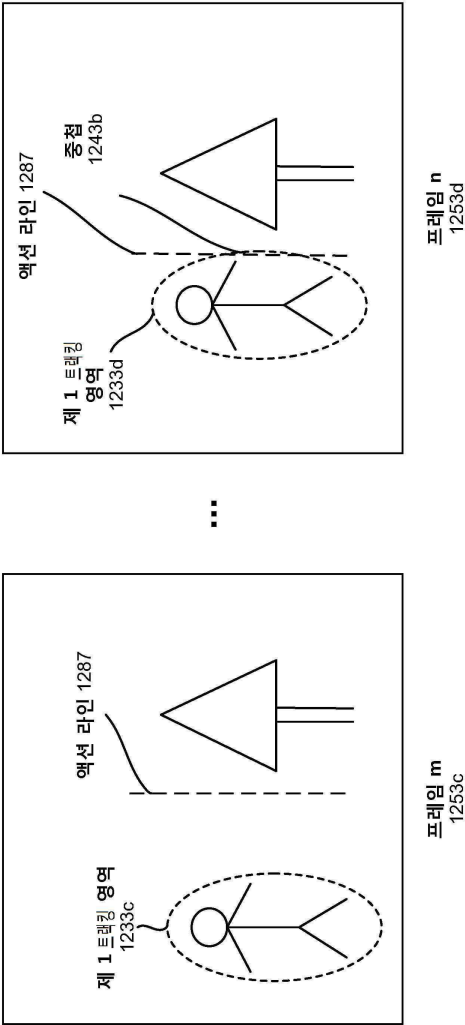
도면11



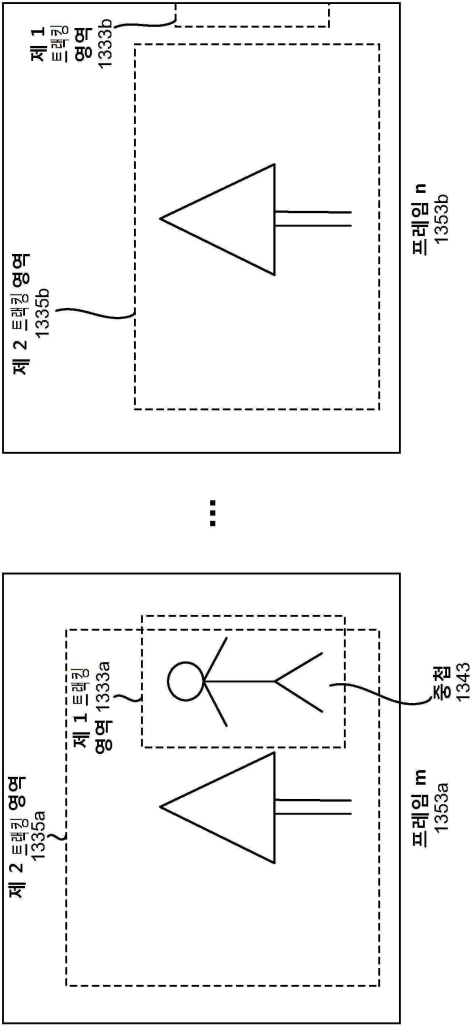
도면12a



도면12b

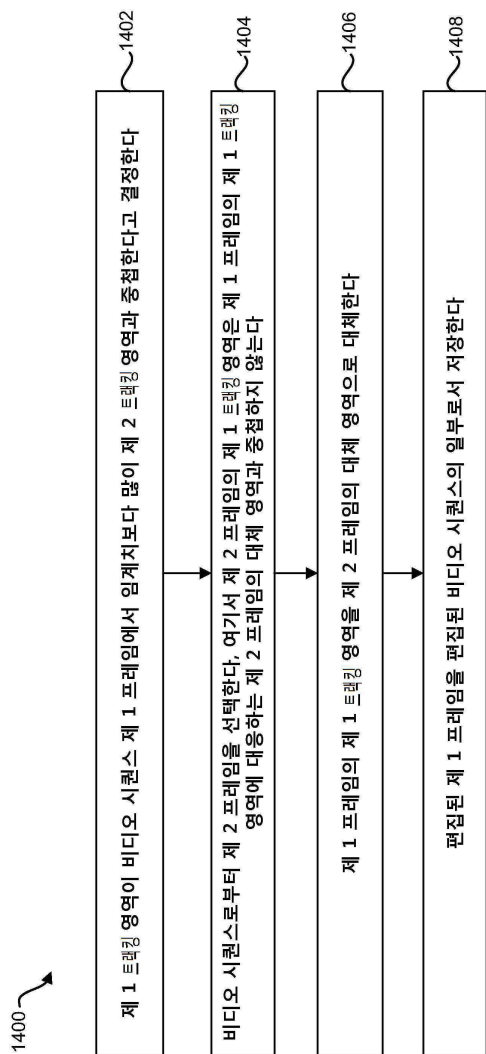


도면13

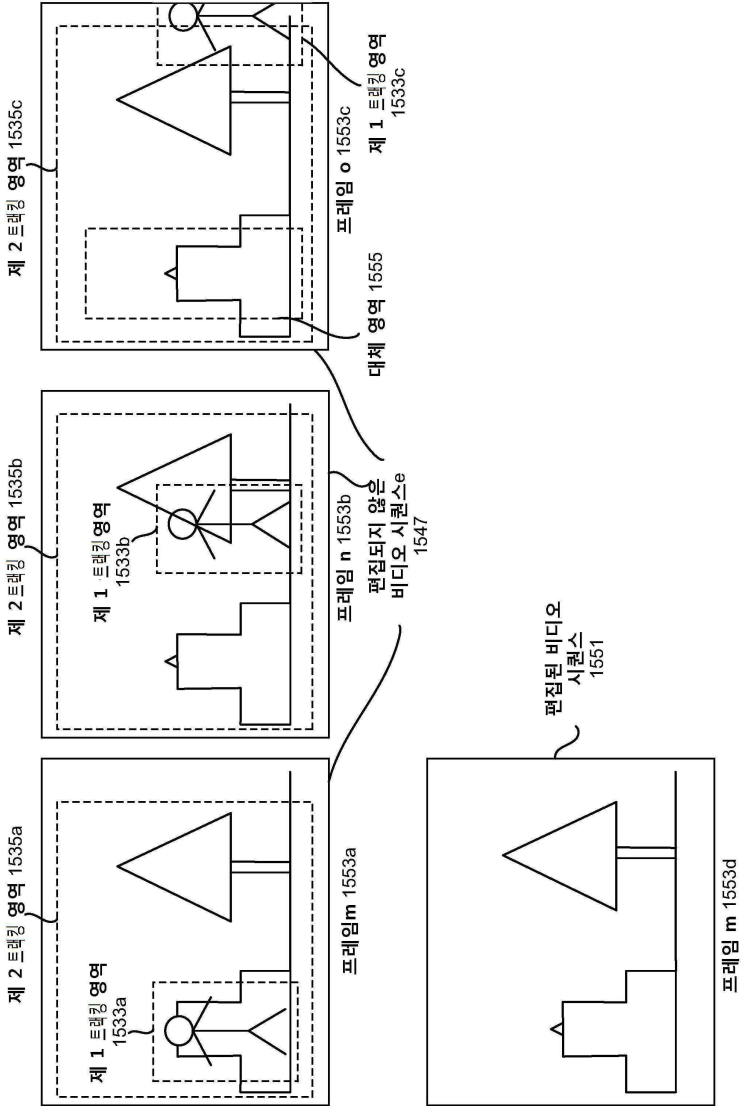




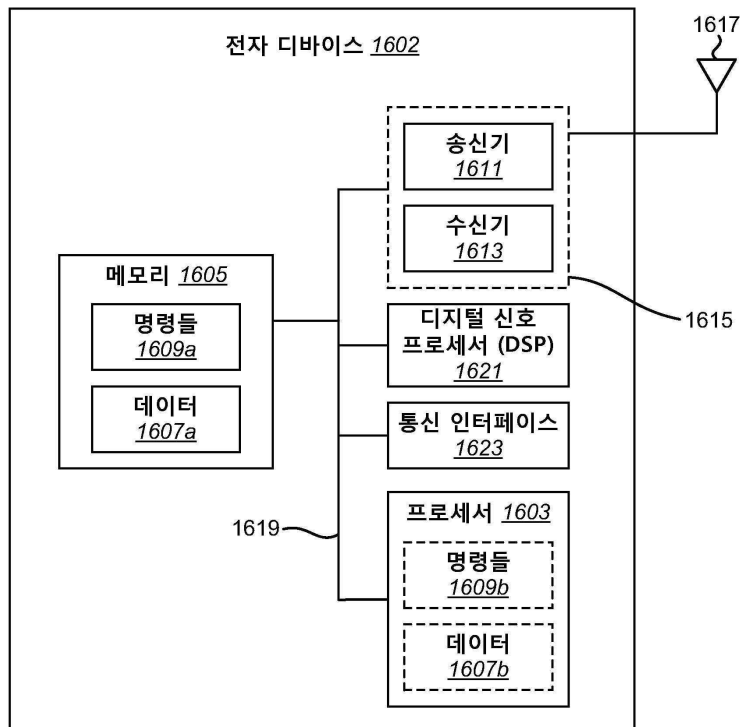
도면14



도면15



도면16



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 36

【변경전】

제 32 항에 있어서,

상기 명령들은,

상기 전자 디바이스로 하여금, 상기 비디오 시퀀스의 제 2 프레임을 선택하게 하는 코드로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않는, 상기 제 2 프레임을 선택하게 하는 코드; 및

상기 전자 디바이스로 하여금, 상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하게 하는 코드를 추가적으로 포함하는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.

【변경후】

제 32 항에 있어서,

상기 명령들은,

상기 전자 디바이스로 하여금, 비디오 시퀀스의 제 2 프레임을 선택하게 하는 코드로서, 상기 제 1 트래킹 영역은 상기 제 2 프레임 내의 대응하는 대체 영역과 중첩하지 않는, 상기 제 2 프레임을 선택하게 하는 코드; 및

상기 전자 디바이스로 하여금, 상기 제 1 프레임 내의 상기 제 1 트래킹 영역을 상기 제 2 프레임으로부터의 상기 대응하는 대체 영역으로 대체하게 하는 코드를 추가적으로 포함하는, 화상 프로세싱을 위한 비일시적 컴퓨터-판독가능 저장매체.