



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0100324
 (43) 공개일자 2016년08월23일

- | | |
|--|--|
| (51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01) G06F 17/27 (2006.01)
G06T 11/60 (2006.01) G06T 7/00 (2006.01)
(52) CPC특허분류
G06F 17/30268 (2013.01)
G06F 17/2765 (2013.01)
(21) 출원번호 10-2016-7018233
(22) 출원일자(국제) 2014년12월19일
심사청구일자 없음
(85) 번역문제출일자 2015년07월07일
(86) 국제출원번호 PCT/US2014/071606
(87) 국제공개번호 WO 2015/095762
국제공개일자 2015년06월25일
(30) 우선권주장
61/919,627 2013년12월20일 미국(US)
(뒷면에 계속) | (71) 출원인
켈컴 인코포레이티드
미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775
(72) 발명자
가오 다산
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
종 신
미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775
(뒷면에 계속)
(74) 대리인
특허법인코리아나 |
|--|--|

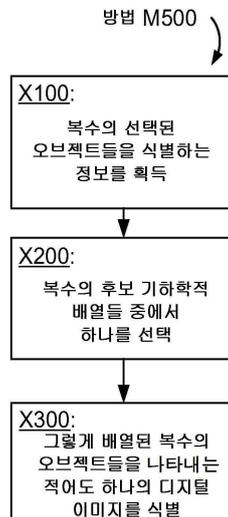
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **이미지 취출을 위한 시스템들, 방법들, 및 장치**

(57) 요약

이미지 취출의 방법은 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 단계 및 복수의 후보 기하학적 배열들 중에서 하나를 선택하는 단계를 포함한다. 이 방법은 또한, 적어도 하나의 프로세서에 의해, 그리고 상기 선택하는 단계에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 단계를 포함한다.

대표도



(52) CPC특허분류

G06F 17/30271 (2013.01)
G06F 17/30705 (2013.01)
G06F 17/3079 (2013.01)
G06F 17/30817 (2013.01)
G06F 17/30831 (2013.01)
G06F 17/30867 (2013.01)
G06T 11/60 (2013.01)
G06T 7/0044 (2013.01)
G06T 7/0048 (2013.01)

(72) 발명자

칸다다이 아난타파드마나반 아라사니팔라이

미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775

레이버 스티븐 더글라스

미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775

(30) 우선권주장

62/036,502	2014년08월12일	미국(US)
62/063,498	2014년10월14일	미국(US)
14/576,006	2014년12월18일	미국(US)

명세서

청구범위

청구항 1

이미지 추출의 방법으로서,

복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 단계;

복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택하는 단계; 및

적어도 하나의 프로세서에 의해, 그리고 상기 선택하는 단계에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 상기 후보 기하학적 배열에 따라 배열된 상기 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 단계

를 포함하는, 이미지 추출의 방법.

청구항 2

제 1 항에 있어서,

획득된 상기 정보는 상기 복수의 선택된 오브젝트들의 각각의 오브젝트에 대해, 상기 오브젝트와 연관되는 라벨을 포함하고,

상기 식별하는 단계는 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 상기 라벨들을 검색하는 단계를 포함하는, 이미지 추출의 방법.

청구항 3

제 1 항에 있어서,

상기 식별하는 단계는 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 상기 후보 기하학적 배열과 연관되는 라벨을 검색하는 단계를 포함하는, 이미지 추출의 방법.

청구항 4

제 1 항에 있어서,

상기 적어도 하나의 디지털 이미지는 제 1 비디오 파일로부터의 복수의 프레임들 및 제 2 비디오 파일로부터의 복수의 프레임들을 포함하는, 이미지 추출의 방법.

청구항 5

제 1 항에 있어서,

상기 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 단계는 (A) 스피치 인식 및 (B) 터치스크린 중 적어도 하나를 이용하여 수행되는, 이미지 추출의 방법.

청구항 6

명령들을 저장하는 비일시적 컴퓨터 판독가능 매체로서,

상기 명령들은 프로세서에 의해 실행될 때, 상기 프로세서로 하여금:

복수의 선택된 오브젝트들을 식별하는 정보를 획득하게 하고;

복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택하게 하고; 그리고

상기 선택하는 것에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 상기 후보 기하학적 배열에 따라 배열된 상기 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하게 하는, 비일시적 컴퓨터 판독가능 매체.

청구항 7

제 6 항에 있어서,

획득된 상기 정보는 상기 복수의 선택된 오브젝트들의 각각의 오브젝트에 대해, 상기 오브젝트와 연관되는 라벨을 포함하고,

상기 비밀시적 컴퓨터 판독가능 매체는 상기 프로세서에 의해 실행될 때, 상기 프로세서로 하여금, 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 상기 라벨들을 검색하게 하는 명령들을 저장하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 8

제 6 항에 있어서,

상기 비밀시적 컴퓨터 판독가능 매체는 상기 프로세서에 의해 실행될 때, 상기 프로세서로 하여금, 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 상기 후보 기하학적 배열과 연관되는 라벨을 검색하게 하는 명령들을 저장하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 9

제 6 항에 있어서,

상기 적어도 하나의 디지털 이미지는 제 1 비디오 파일로부터의 복수의 프레임들 및 제 2 비디오 파일로부터의 복수의 프레임들을 포함하는, 비밀시적 컴퓨터 판독가능 매체.

청구항 10

제 6 항에 있어서,

상기 복수의 선택된 오브젝트들을 식별하는 정보는 (A) 스피치 인식 및 (B) 터치스크린 중 적어도 하나를 이용하여 획득되는, 비밀시적 컴퓨터 판독가능 매체.

청구항 11

이미지 추출을 위한 장치로서,

복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 수단;

복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택하는 수단; 및

상기 선택하는 것에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 상기 후보 기하학적 배열에 따라 배열된 상기 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 수단

을 포함하는, 이미지 추출을 위한 장치.

청구항 12

제 11 항에 있어서,

획득된 상기 정보는 상기 복수의 선택된 오브젝트들의 각각의 오브젝트에 대해, 상기 오브젝트와 연관되는 라벨을 포함하고,

상기 이미지 추출을 위한 장치는 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 상기 라벨들을 검색하는 수단을 포함하는, 이미지 추출을 위한 장치.

청구항 13

제 11 항에 있어서,

상기 이미지 추출을 위한 장치는 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 상기 후보 기하학적 배열과 연관되는 라벨을 검색하는 수단을 포함하는, 이미지 추출을 위한 장치.

청구항 14

제 11 항에 있어서,

상기 적어도 하나의 디지털 이미지는 제 1 비디오 파일로부터의 복수의 프레임들 및 제 2 비디오 파일로부터의 복수의 프레임들을 포함하는, 이미지 취출을 위한 장치.

청구항 15

제 11 항에 있어서,

상기 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 수단은 (A) 스피치 인식 및 (B) 터치스크린 중 적어도 하나를 이용하여 상기 정보를 획득하도록 구성되는, 이미지 취출을 위한 장치.

청구항 16

이미지 취출을 위한 장치로서,

(A) 복수의 선택된 오브젝트들을 식별하는 정보 및 (B) 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열의 선택을 획득하도록 구성된 판별기; 및

상기 선택에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 상기 후보 기하학적 배열에 따라 배열된 상기 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하도록 구성된 검색 엔진

을 포함하는, 이미지 취출을 위한 장치.

청구항 17

제 16 항에 있어서,

획득된 상기 정보는 상기 복수의 선택된 오브젝트들의 각각의 오브젝트에 대해, 상기 오브젝트와 연관되는 라벨을 포함하고,

상기 검색 엔진은 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 상기 라벨들을 검색하도록 구성되는, 이미지 취출을 위한 장치.

청구항 18

제 16 항에 있어서,

상기 검색 엔진은 상기 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 상기 후보 기하학적 배열과 연관되는 라벨을 검색하도록 구성되는, 이미지 취출을 위한 장치.

청구항 19

제 16 항에 있어서,

상기 적어도 하나의 디지털 이미지는 제 1 비디오 파일로부터의 복수의 프레임들 및 제 2 비디오 파일로부터의 복수의 프레임들을 포함하는, 이미지 취출을 위한 장치.

청구항 20

제 16 항에 있어서,

상기 판별기는 (A) 스피치 인식 및 (B) 터치스크린 중 적어도 하나를 이용하여 상기 복수의 선택된 오브젝트들을 식별하는 정보를 획득하도록 구성되는, 이미지 취출을 위한 장치.

발명의 설명

기술 분야

관련 출원들에 대한 상호 참조

[0001]

본 출원은 2013년 12월 20일자로 출원된 공동 소유된 미국 가특허출원 제61/919,627호, 2014년 8월 12일자로 출

[0002]

원된 미국 가특허출원 제62/036,502호, 2014년 10월 14일자로 출원된 미국 가특허출원 제62/063,498호, 및 2014년 12월 18일자로 출원된 미국 정규특허출원 제14/576,006호로부터의 우선권을 주장하고, 이들 내용들은 본 명세서에 참조로 그 전체가 명백히 포함된다.

[0003] 분야

[0004] 본 개시물은 일반적으로 이미지 취출에 관한 것이다.

배경 기술

[0005] 기술의 진보는 보다 소형이고 더 강력한 컴퓨팅 디바이스들을 발생시켰다. 예를 들어, 소형이고, 경량이며, 사용자들에 의해 쉽게 운반되는 휴대용 무선 전화기들, 개인 휴대 정보 단말기 (PDA) 들, 및 페이지 디바이스들과 같은, 무선 컴퓨팅 디바이스들을 포함하는, 다양한 휴대용 퍼스널 컴퓨팅 디바이스들이 현재 존재한다. 더 구체적으로는, 셀룰러 전화기들 및 인터넷 프로토콜 (IP) 전화기들과 같은 휴대용 무선 전화기들은 무선 네트워크들을 통해 보이스 및 데이터 패킷들을 통신할 수 있다. 게다가, 많은 이러한 무선 전화기들은 내부에 포함되는 다른 타입들의 디바이스들을 포함한다. 예를 들어, 무선 전화기는 또한 디지털 스틸 카메라, 디지털 비디오 카메라, 디지털 레코더, 및 오디오 파일 플레이어들을 포함할 수 있다. 또한, 이러한 무선 전화기들은 인터넷에 액세스하는데 이용될 수 있는 웹 브라우저 애플리케이션과 같은 소프트웨어 애플리케이션들을 포함하여, 실행가능한 명령들을 프로세싱할 수 있다. 이와 같이, 이들 무선 전화기들은 상당한 컴퓨팅 능력들을 포함할 수 있다.

발명의 내용

과제의 해결 수단

[0006] 일반적 구성에 따른 이미지 취출의 방법은 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 단계 및 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택하는 단계를 포함한다. 이 방법은 또한, 적어도 하나의 프로세서에 의해, 그리고 상기 선택하는 단계에 응답하여, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 단계를 포함한다. 명령들을 실행하는 머신으로 하여금 이러한 방법을 수행하게 하는 명령들을 갖는 컴퓨터 판독가능 저장 매체들 (예를 들어, 비-순시적 매체들) 이 또한 개시된다.

[0007] 일반적 구성에 따른 이미지 취출을 위한 장치는 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 수단 및 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택하는 수단을 포함한다. 이 장치는 또한, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 수단을 포함한다.

[0008] 다른 일반적 구성에 따른 이미지 취출을 위한 장치는 (A) 복수의 선택된 오브젝트들을 식별하는 정보 및 (B) 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열의 선택을 획득하도록 구성된 판별기를 포함한다. 장치는 또한, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하도록 구성된 검색 엔진을 포함한다.

[0009] 개시된 실시형태들 중 적어도 하나에 의해 제공된 하나의 특정 이점은 모바일 디바이스들 상에서의 이미지 추적 및 프레임 취출을 이용한 개선된 사용자 경험이다. 본 개시물의 다른 양태들, 이점들 및 피쳐들 (features) 은 다음 섹션들: 도면의 간단한 설명, 상세한 설명, 및 청구항들을 포함하는 전체 출원의 검토 후에 명백해질 것이다.

도면의 간단한 설명

- [0010] 도 1 은 전자 디바이스를 나타내는 블록 다이어그램이다.
- 도 2a 는 오브젝트 및 검출 모듈의 특정 예시적 실시형태를 나타내는 블록 다이어그램이다.
- 도 2b 는 도 2 의 오브젝트 및 검출 모듈 내의 컴포넌트들을 구현하는 프로세서의 특정 예시적 실시형태를 나타내는 블록 다이어그램이다.
- 도 3 은 모션-기반 추적 및 오브젝트 검출을 수행하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.

- 도 4 는 모션-기반 추적을 수행하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.
- 도 5 는 순방향-역방향 에러 (forward-backward error) 에 기초하여 모션-기반 추적에서 추적 에러를 추정하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.
- 도 6 은 오브젝트 검출을 수행하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.
- 도 7 은 상이한 윈도우 사이즈들을 갖는 이미지 윈도우의 특정 예시적 실시형태이다.
- 도 8 은 오브젝트 추적 및 검출 모듈의 특정 예시적 실시형태를 나타내는 블록 다이어그램이다.
- 도 9 는 평활화 모듈 (smoothing module) 의 특정 예시적 실시형태를 나타내는 블록 다이어그램이다.
- 도 10 은 모션 추적에서 지터 (jitter) 를 평활화하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.
- 도 11a 는 일반적 구성에 따라 메타데이터를 발생시키는 방법 (M100) 의 플로우차트를 도시한다.
- 도 11b 는 방법 (M100) 의 구현 (M110) 의 플로우차트를 도시한다.
- 도 11c 는 방법 (M100) 의 구현 (M120) 의 플로우차트를 도시한다.
- 도 11d 는 방법들 (M110 및 M120) 의 구현 (M130) 의 플로우차트를 도시한다.
- 도 12 는 3 개의 오브젝트들을 선택하기 위한 사용자 액션들의 시퀀스의 예를 도시한다.
- 도 13 은 3 개의 오브젝트들을 선택하기 위한 사용자 액션들의 시퀀스의 다른 예를 도시한다.
- 도 14 는 오브젝트들의 프레임들 및 향상된 디스플레이 내의 오브젝트들의 예들을 도시한다.
- 도 15a 내지 도 15c 는 선택 포인트를 조작하기 위해 배향-감응 디바이스를 이용하는 예들을 도시한다.
- 도 16 은 실제 거리들과 비교하여 오브젝트들의 이미지들 간의 거리들 사이의 불일치 (discrepancy) 의 예를 도시한다.
- 도 17a 는 대응하는 결정된 포지션을 획득하기 위해 오브젝트의 하부 바운딩 라인으로 오브젝트의 질량 중심을 프로젝팅하는 예를 도시한다.
- 도 17b 는 그라운드 평면에서의 결정된 포지션에 대한 링크된 포인트의 프로젝션의 예를 도시한다.
- 도 18a 및 도 18b 는 그라운드 평면에서의 거리들과 픽셀 좌표 공간에서의 거리들 사이의 관련성을 도시한다.
- 도 19a 는 농구 코트의 상부도를 도시하고, 도 19b 내지 도 19d 는 그 코트의 로케이션 공간의 불균일한 분할의 여러 예들을 도시한다.
- 도 20 은 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 다른 특정 실시형태이다.
- 도 21 은 클러스터링에 기초하여 프레임들을 취출하는데 이용되는 모바일 디바이스의 스크린의 특정 실시형태들을 나타낸다.
- 도 22 는 오브젝트 추적을 이용하는 비디오 프로세싱 방법의 특정 예시적 실시형태의 다른 플로우차트이다.
- 도 23 은 오브젝트 추적을 이용하는 비디오 프로세싱 방법들의 특정 예시적 실시형태들의 플로우차트들을 나타낸다.
- 도 24 는 직교 좌표들을 이용하여 오브젝트 로케이션에 대한 메타데이터를 인코딩하는 특정 실시형태를 나타낸다.
- 도 25 는 극 좌표들을 이용하여 오브젝트 로케이션에 대한 메타데이터를 인코딩하는 특정 실시형태를 나타낸다.
- 도 26 은 메타데이터를 인코딩하기 위한 방법의 특정 예시적 실시형태의 플로우차트이다.
- 도 27 은 오브젝트 추적을 이용하는 비디오 프로세싱 방법의 특정 예시적 실시형태의 다른 플로우차트이다.
- 도 28 은 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 다른 특정 실시형태이다.
- 도 29 는 오브젝트 추적을 이용하는 비디오 프로세싱 방법의 특정 예시적 실시형태의 다른 플로우차트이다.
- 도 30a 는 일반적 구성에 따른 장치 (A100) 의 블록 다이어그램을 도시한다.

- 도 30b 는 장치 (A100) 의 구현 (A110) 의 블록 다이어그램을 도시한다.
- 도 30c 는 장치 (A100) 의 구현 (A120) 의 블록 다이어그램을 도시한다.
- 도 30d 는 장치 (A110 및 A120) 의 구현 (A130) 의 블록 다이어그램을 도시한다.
- 도 31 은 비디오 프로세싱 기법들을 수행하도록 동작가능한 컴포넌트들을 포함하는 무선 디바이스의 블록 다이어그램이다.
- 도 32a 는 일반적 구성에 따른 장치 (MF100) 의 블록 다이어그램을 도시한다.
- 도 32b 는 장치 (MF100) 의 구현 (MF110) 의 블록 다이어그램을 도시한다.
- 도 32c 는 장치 (MF100) 의 구현 (MF120) 의 블록 다이어그램을 도시한다.
- 도 32d 는 장치 (MF110 및 MF120) 의 구현 (MF130) 의 블록 다이어그램을 도시한다.
- 도 33a 는 템플릿들의 세트의 예를 도시하고, 도 33b 는 템플릿에 대한 수정의 예들을 도시한다.
- 도 34a 는 선택된 오브젝트들의 상이한 형성들의 예들을 도시한다.
- 도 34b 는 불균일한 맵핑들의 예들을 도시한다.
- 도 34c 는 선택된 오브젝트들의 균일한 형성들을 도시한다.
- 도 35 는 판정 메트릭들의 세트의 예를 도시한다.
- 도 36 은 판정 메트릭들의 세트의 다른 예를 도시한다.
- 도 37a 는 방법 (M200) 의 구현 (M200) 의 플로우차트를 도시한다.
- 도 37b 내지 도 37d 는 방법들 (M110, M120, 및 M130) 각각의 구현들 (M210, M220, 및 M230) 의 플로우차트들을 도시한다.
- 도 38a 는 일반적 구성에 따라 메타데이터를 발생시키기 위한 장치 (A100) 의 블록 다이어그램을 도시한다.
- 도 38b 는 장치 (A100) 의 구현 (A110) 의 블록 다이어그램을 도시한다.
- 도 38c 는 장치 (A100) 의 구현 (A120) 의 블록 다이어그램을 도시한다.
- 도 38d 는 장치 (A110 및 A120) 의 구현 (A130) 의 블록 다이어그램을 도시한다.
- 도 38e 는 장치 (A100) 의 구현 (A200) 의 블록 다이어그램을 도시한다.
- 도 38f, 도 38g, 및 도 39a 는 장치 (A110, A120, 및 130) 각각의 구현들 (A210, A220, 및 A230) 의 블록 다이어그램들을 도시한다.
- 도 39b 는 장치 (A100) 의 구현 (A300) 의 블록 다이어그램을 도시한다.
- 도 39c 및 도 39d 는 장치 (A200 및 A230) 의 구현들 (A310 및 A330) 의 블록 다이어그램들을 도시한다.
- 도 40a 는 일반적 구성에 따라 메타데이터를 발생시키기 위한 장치 (MF100) 의 블록 다이어그램을 도시한다.
- 도 40b 는 장치 (MF100) 의 구현 (MF110) 의 블록 다이어그램을 도시한다.
- 도 40c 는 장치 (MF100) 의 구현 (MF120) 의 블록 다이어그램을 도시한다.
- 도 40e 는 장치 (MF100) 의 구현 (MF200) 의 블록 다이어그램을 도시한다.
- 도 40f, 도 40g, 및 도 41a 는 장치 (MF110, MF120, 및 MF130) 각각의 구현들 (MF210, MF220, 및 MF230) 의 블록 다이어그램들을 도시한다.
- 도 41b 는 장치 (MF100) 의 구현 (MF400) 의 블록 다이어그램을 도시한다.
- 도 41c 는 장치 (MF100) 의 구현 (MF300) 의 블록 다이어그램을 도시한다.
- 도 41d 및 도 41e 는 장치 (MF200 및 MF230) 의 구현들 (MF310 및 MF330) 의 블록 다이어그램들을 도시한다.
- 도 42 는 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 특정 실시형태를 도시한다.

도 43 은 오브젝트 로케이션에 대한 메타데이터를 인코딩하는 특정 실시형태를 도시한다.

도 44 는 비디오 프로세싱 방법의 특정 실시형태를 예시하는 플로우차트를 도시한다.

도 45a 는 일반적 구성에 따른 방법 (M500) 의 플로우차트를 도시한다.

도 45b 는 일반적 구성에 따른 장치 (A500) 의 블록 다이어그램을 도시한다.

도 45c 는 일반적 구성에 따른 장치 (MF500) 의 블록 다이어그램을 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0011] 무선 전화기 또는 다른 모바일 디바이스는 카메라로 비디오 스트림들을 캡처하거나 및/또는 다른 디바이스로부터 및/또는 네트워크를 통해 비디오 스트림들을 수신할 수도 있다. 비디오 스트림들 내의 오브젝트들을 추적하기 위한 새로운 및/또는 개선된 피쳐들 (features) 을 원할 수도 있다.
- [0012] 문맥상으로 명백히 제한되지 않는 한, 용어 "신호" 는 와이어, 버스, 또는 다른 송신 매체에 대해 표현되는 메모리 로케이션 (또는 메모리 로케이션들의 세트) 의 상태를 포함하는, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 본 명세서에서 사용된다. 문맥상으로 명백히 제한되지 않는 한, 용어 "발생시키는" 은 컴퓨팅하는 또는 그렇지 않으면 생성하는 과 같은, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 본 명세서에서 사용된다. 문맥상으로 명백히 제한되지 않는 한, 용어 "계산하는" 은 복수의 값들로부터 컴퓨팅하는, 평가하는, 추정하는, 및/또는 선택하는 과 같은, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 본 명세서에서 사용된다. 문맥상으로 명백히 제한되지 않는 한, 용어 "획득하는" 은 계산하는, 유도하는, (예를 들어, 외부 디바이스로부터) 수신하는, 및/또는 (예를 들어, 저장 엘리먼트들의 어레이로부터) 취출하는 과 같은, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 사용된다. 문맥상으로 명백히 제한되지 않는 한, 용어 "선택하는" 은 2 개 이상의 것들의 세트 중 적어도 하나, 그리고 그 전부보다는 더 적은 것을 식별하는, 나타내는, 적용하는, 및/또는 사용하는 과 같은, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 사용된다. 용어 "포함하는" 이 본 설명 및 청구항들에 사용되는 경우, 이것은 다른 엘리먼트들 또는 동작들을 배제시키지 않는다. ("A 가 B 에 기초한다" 에서처럼) 용어 "에 기초" 한다는 것은 (i) "로부터 유도된" 다는 것 (예를 들어, "B 가 A 의 프리커서 (precursor) 이다"), (ii) "적어도 ~ 에 기초" 한다는 것 (예를 들어, "A 는 적어도 B 에 기초한다") 그리고, 특정 문맥에서 적절하다면, (iii) "와 동일" 하다는 것 (예를 들어, "A 는 B 와 동일하다") 의 경우들을 포함하는, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 사용된다. 이와 유사하게, 용어 "에 응답하여" 는 "적어도 ~ 에 응답하여" 를 포함하는, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 사용된다.
- [0013] 달리 나타내지 않는 한, 특정 피쳐를 갖는 장치의 동작의 임의의 개시는 또한 유사한 피쳐를 갖는 방법을 개시하도록 명백히 의도되고 (그의 역도 같음), 특정 구성에 따른 장치의 동작의 임의의 개시는 또한 유사한 구성에 따른 방법을 개시하도록 명백히 의도된다 (그의 역도 같음). 용어 "구성 (configuration)" 은 그의 특정 문맥상으로 나타낸 바와 같이 방법, 장치, 및/또는 시스템과 관련하여 사용될 수도 있다. 특정 문맥상으로 달리 나타내지 않는 한, 용어들 "방법", "프로세스", "프로시저", 및 "기법" 은 일반적으로 그리고 상호교환가능하게 사용된다. 특정 문맥상으로 달리 나타내지 않는 한, 용어들 "장치" 및 "디바이스" 는 또한 일반적으로 그리고 상호교환가능하게 사용된다. 용어들 "엘리먼트" 및 "모듈" 은 통상적으로 보다 큰 구성의 일 부분을 나타내기 위해 사용된다. 문맥상으로 명백히 제한되지 않는 한, 용어 "시스템" 은 "공통된 목적에 기여하기 위해 상호작용하는 엘리먼트들의 그룹" 을 포함하는, 그의 통상의 의미들 중 임의의 의미를 나타내기 위해 본 명세서에서 사용된다.
- [0014] 달리 나타내지 않는 한, 용어 "시리즈" 는 2 개 이상의 아이템들의 시퀀스를 나타내기 위해 사용된다. 초기에 정관사에 의해 도입되지 않는 한, 청구항 엘리먼트를 수정하기 위해 사용된 서수 용어 (예를 들어, "제 1", "제 2", "제 3" 등) 는 그것만으로는 다른 것에 대한 청구항 엘리먼트의 임의의 우선순위 또는 순서를 나타내지 않으며, 오히려 청구항 엘리먼트를 동일한 명칭을 갖는 (그러나 서수 용어의 사용을 위한) 다른 청구항 엘리먼트로부터 단지 구별할 뿐이다. 문맥상으로 명백히 제한되지 않는 한, 용어들 "복수" 및 "세트" 각각은 하나보다 더 큰 정수의 양을 나타내기 위해 본 명세서에서 사용된다.
- [0015] 도 1 을 참조하면, 전자 디바이스 (102) 를 예시하는 블록 다이어그램이 도시된다. 전자 디바이스 (102) 는 또한 무선 통신 디바이스, 모바일 디바이스, 이동국, 가입자국, 클라이언트, 클라이언트 스테이션, 사용자 장비 (UE), 원격 스테이션, 액세스 단말기, 모바일 단말기, 단말기, 사용자 단말기, 가입자 유닛 등으로 지칭될 수도

있다. 전자 디바이스들의 예들로는 랩톱 또는 데스크톱 컴퓨터들, 셀룰러 폰들, 스마트 폰들, 무선 모뎀들, 전자 판독기들 (e-readers), 태블릿 디바이스들, 게이밍 시스템들 등을 포함한다. 이들 디바이스들 중 일부는 하나 이상의 산업 표준들에 따라 동작할 수도 있다.

[0016] 스마트폰 또는 태블릿 컴퓨터와 같은 전자 디바이스 (102) 는 카메라를 포함할 수도 있다. 카메라는 이미지 센서 (114) 및 광학 시스템 (118) (예를 들어, 렌즈들) 을 포함할 수도 있어서 광학 시스템 (118) 의 시야 (field of view) 내에 로케이팅되는 오브젝트들의 이미지들을 이미지 센서 (114) 상으로 포커싱한다. 전자 디바이스 (102) 는 또한 카메라 소프트웨어 애플리케이션 및 디스플레이 스크린을 포함할 수도 있다. 카메라 애플리케이션이 작동하고 있을 때, 광학 시스템 (118) 의 시야 내에 로케이팅되는 오브젝트들의 이미지들은 이미지 센서 (114) 에 의해 레코딩될 수도 있다. 이미지 센서 (114) 에 의해 레코딩되고 있는 이미지들은 디스플레이 스크린 상에 디스플레이될 수도 있다. 이들 이미지들은 비교적 높은 프레임 레이트에서 신속히 연속적으로 디스플레이될 수도 있어서, 임의의 주어진 시간의 순간에, 광학 시스템 (118) 의 시야 내에 로케이팅되는 오브젝트들은 디스플레이 스크린 상에 디스플레이된다. 실시형태들이 캡처된 프레임들 (예를 들어, 비디오 프레임들) 의 관점에서 설명되지만, 본 명세서에서 논의되는 기법들은 임의의 디지털 이미지에 대해 이용될 수도 있다. 그에 따라, 용어들 "프레임" 및 "디지털 이미지" 는 본 명세서에서 상호교환가능하게 사용될 수도 있다.

[0017] 카메라 애플리케이션의 사용자 인터페이스 (120) 는 디스플레이 스크린 상에 디스플레이되고 있는 하나 이상의 오브젝트들이 추적되는 것을 허용할 수도 있다. 전자 디바이스 (102) 의 사용자는 추적되어야 하는 오브젝트(들) 를 선택하도록 허용될 수도 있다. 게다가, 선택된 오브젝트(들) 는 오브젝트를 추후에 검출하기 위한 레퍼런스 (reference) 로서 이용될 수도 있다.

[0018] 하나의 구성에서, 디스플레이는, 예를 들어, 손가락, 스타일러스 또는 다른 툴에 의한, 물리적 터치로부터 입력을 수신하는 터치스크린 (116) 이다. 터치스크린 (116) 은 추적될 타깃 오브젝트를 정의하는 터치 입력을 수신할 수도 있다. 예를 들어, 전자 디바이스 (102) 가 관심있는 동물을 포함하는 자연의 장면을 캡처하고 있는 경우, 사용자는, 필요하다면, 동물이 추적되거나, 또는 검출되는 것의 요망을 나타내는 바운딩 박스 (bounding box) 를 동물 주위로 드로잉 (drawing) 할 수도 있다. 타깃 오브젝트들은 임의의 적합한 방법으로 선택될 수도 있다. 예를 들어, 추적되거나, 검출되거나, 또는 이 양쪽 모두가 행해져야 하는 타깃 오브젝트를 선택하기 위해 얼굴 인식, 보행자 인식 등이 이용될 수도 있다. 하나의 구성에서, 다수의 오브젝트들이 추적될 수도 있다. 사용자 인터페이스 (120) 는 사용자로 하여금 오브젝트 추적 및 검출 모듈 (104) 과 상호작용하게 하여, 예를 들어, 하나 이상의 타깃 오브젝트들을 선택 (즉, 정의) 할 수도 있다. 터치스크린 (116) 은 뷰파인더 (131) 를 포함할 수도 있다. 뷰파인더 (131) 는 비디오 스트림 또는 라이브 피드 (live feed) 를 디스플레이하는 터치스크린 (116) 의 부분을 지칭할 수도 있다. 예를 들어, 뷰파인더 (131) 는 전자 디바이스 (102) 상의 카메라에 의해 획득된 뷰를 디스플레이할 수도 있다.

[0019] 전자 디바이스 (102) 는 선택된 오브젝트를 추적하거나 및/또는 비디오 프레임에서 오브젝트를 검출하기 위한 오브젝트 추적 및 검출 모듈 (104) 을 포함할 수도 있다. 오브젝트 추적 및 검출 모듈 (104) 은 하나 이상의 오브젝트들을 추적하기 위한 모션 추적기 (106) 를 포함할 수도 있다. 모션 추적기 (106) 는 프레임 간에서 이미지 (예를 들어, 비디오 프레임) 상의 포인트들의 모션을 추적하여 이전 비디오 프레임과 현재 비디오 프레임 사이의 타깃 오브젝트의 로케이션 및/또는 로케이션의 변화를 추정하기 위해 모션-기반일 수도 있다.

[0020] 오브젝트 추적 및 검출 모듈은 또한 비디오 프레임에서 오브젝트를 검출하기 위한 오브젝트 검출기 (108) 를 포함할 수도 있다. 오브젝트 검출기 (108) 는, 현재 비디오 프레임의 전부 또는 일 부분을, (예를 들어, 비디오 프레임들의 시퀀스에서) 캡처된 이전 비디오 프레임 (112) 의 선택된 오브젝트 또는 부분과 비교하는 것에 의해 오브젝트를 검출하기 위해, 모션-기반 모델보다는 오히려, 오브젝트 모델을 이용할 수도 있다. 오브젝트 검출기 (108) 는 비디오 프레임 내의 다수의 오브젝트들을 검출하기 위해 이용될 수도 있다.

[0021] 오브젝트 추적 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. 하나의 예에서, 메모리 버퍼 (110) 는 캡처된 이전 비디오 프레임 (112) 을 저장할 수도 있다. 오브젝트 추적 및 검출 모듈 (104) 은 모션-기반 추적 및/또는 오브젝트 검출을 수행함에 있어서 캡처된 이전 비디오 프레임 (112) 에 관해 메모리 버퍼 (110) 로부터 제공된 데이터를 이용할 수도 있다. 모션-기반 추적 및 오브젝트 검출에 맞춰서 타깃 오브젝트를 더 정확히 추적 및/또는 검출하기 위해 메모리 버퍼 (110) 로부터의 피드백을 통해 모션 추적기 (106) 또는 오브젝트 검출기 (108) 에 데이터가 제공될 수도 있다. 예를 들어, 메모리 버퍼

(110) 는 모션 추적기 (106) 및 오브젝트 검출기 (108) 에 로케이션 및 윈도우 사이즈 데이터를 제공하여, 오브젝트를 추적 또는 검출할 때 오브젝트의 로케이션 및 사이즈를 더 정확히 집어내기 위해 이용될 수도 있는 하나 이상의 파라미터들을 모션 추적기 (106) 및 오브젝트 검출기 (108) 에 제공할 수도 있다.

[0022] 위에 명시된 바와 같이, 전자 디바이스 (102) 는 모션-기반 추적을 수행할 수도 있다. 모션-기반 추적은 다양한 방법들을 이용하여 수행될 수도 있다. 하나의 예에서, 모션 추적기 (106) 가 바운딩 박스 β_t 및 이미지들 I_t, I_{t+1} (예를 들어, 비디오 프레임들) 의 쌍을 수용하고 바운딩 박스 β_{t+1} 를 출력하는 메디안 플로우 방법 (median flow method) 에 의해 추적이 수행된다. 바운딩 박스 β_t 내의 직사각형 그리드 상에서 포인트들의 세트가 초기화될 수도 있고, 이 포인트들이 추적되어 I_t 와 I_{t+1} 사이의 스페스 (sparse) 모션 플로우를 발생시킬 수도 있다. 포인트 예측의 품질이 추정될 수도 있고 각각의 포인트에는 에러가 할당된다. 최악의 예측들 중 일 부분 (예를 들어, 50%) 이 필터링될 수도 있지만 나머지 예측들은 전체 바운딩 박스의 변위를 추정하는데 이용된다. 모션 추적기 (106) 는 전자 디바이스 (102) 에 의해 캡처된 각각의 비디오 프레임 상에서 모션-기반 추적을 수행할 수도 있다. 유사한 방법으로, 하나 이상의 그라디언트들 (gradients) (예를 들어, x 및 y 그라디언트들) 을 계산하고 프레임들의 쌍 사이의 차이를 이용하여 시간 그라디언트를 계산하고 다수의 그라디언트 값들을 이용하여 현재 비디오 프레임 내의 타깃 오브젝트를 정확히 추적하는 것에 의해 모션-기반 추적이 수행될 수도 있다. 모션-기반 추적에 관한 추가의 상세들이 아래에 제공된다.

[0023] 모션-기반 추적을 수행할 때, 모션 추적기 (106) 는 모션-추적 방법의 계산된 또는 추정된 정확도에 기초하여 추적 신뢰도 값을 결정할 수도 있다. 일부 구성들에서, 추적 신뢰도 값은, 타깃 오브젝트가 현재 비디오 프레임 또는 그 비디오 프레임의 정의된 윈도우 내에 있을 가능성 또는 확률에 대응하는, 0 과 1 사이의 실수일 수도 있다. 추적 신뢰도 값은 추적 임계치와 비교될 수도 있다. 추적 신뢰도 값이 추적 임계치보다 더 큰 경우, 타깃 오브젝트가 현재 비디오 프레임 내에서 발견될 가능성이 높을 수도 있다. 대안적으로, 추적 신뢰도 값이 추적 임계치 이하인 경우, 타깃 오브젝트가 현재 비디오 프레임 내에서 발견될지 여부의 가능성이 낮거나 또는 불확실할 수도 있다. 추적 신뢰도 값을 결정하기 위한 다양한 방법들이 이용될 수도 있다. 하나의 구성에서, 이전에 캡처된 비디오 프레임들로부터 이전에 저장된 이미지 패치들과 현재 비디오 프레임에서의 추적된 윈도우 (예를 들어, 추적 패치 윈도우) 사이의 정규화된 상호 상관 (normalized cross correlation; NCC) 을 계산하는 것에 의해 추적 신뢰도 값이 결정된다. 추적 신뢰도 값을 계산하는 것에 관한 추가의 상세들이 아래에 제공된다.

[0024] 전자 디바이스 (102) 는 또한 오브젝트 검출을 수행할 수도 있다. 오브젝트 검출은 다양한 방법들을 이용하여 수행될 수도 있다. 하나의 구성에서, 타깃 오브젝트가 현재 비디오 프레임에서 발견되는지 또는 현재 비디오 프레임의 특정 윈도우 또는 윈도우들의 서브세트 내에서 발견되는지 여부를 결정하기 위해 비디오 프레임 내의 윈도우들의 다수의 서브세트들의 콘텐츠가 뷰잉되는 슬라이딩 윈도우 방법을 이용하여 오브젝트 검출이 수행된다. 모든 가능한 윈도우 로케이션들 및 사이즈들의 서브세트 또는 그 전부는 비디오 프레임에서 검색될 수도 있다. 예를 들어, 각각의 윈도우는 데이터의 픽셀들에 대응할 수도 있고, 오브젝트 검출기 (108) 는 타깃 오브젝트가 특정 윈도우 또는 서브-윈도우 내에 있다는 신뢰도 레벨 (예를 들어, 2진 표시자 (binary indicator)) 을 결정하기 위해 데이터의 픽셀들을 이용하여 하나 이상의 연산들을 수행할 수도 있다. 하나 이상의 윈도우들과 연관된 신뢰도 레벨에 기초하여, 검출기 신뢰도 값이 현재 비디오 프레임에 대해 획득될 수도 있다. 게다가, 오브젝트 검출의 정확도 또는 효율을 증가시키기 위한 부가적인 기법들이 이용될 수도 있다. 이들 기법들 중 일부가 아래에 설명된다.

[0025] 일부 구성들에서, 모션 추적기 (106) 및 오브젝트 검출기 (108) 는 병행하기보다는 오히려 순차적으로 동작할 수도 있다. 예를 들어, 전자 디바이스 (102) 는 선택된 오브젝트 (예를 들어, 타깃 오브젝트) 의 모션-기반 추적을 수행하고, 추적된 파라미터에 기초하여 선택된 오브젝트의 오브젝트 검출을 순차적으로 수행할 수도 있다. 하나의 구성에서, 전자 디바이스 (102) 는 현재 비디오 프레임 상에서 모션-기반 추적을 수행할 수도 있다. 전자 디바이스 (102) 는 그 후에, 추적된 파라미터에 기초하여 현재 프레임 상에서 오브젝트 검출을 수행할 수도 있다. 하나의 구성에서, 추적된 파라미터는 신뢰도 값과 임계치 사이의 비교에 기초할 수도 있다. 예를 들어, 추적 신뢰도 값이 추적 임계치를 하회하는 경우, 전자 디바이스 (102) 는 오브젝트 검출을 수행할 수도 있다. 대안적으로, 추적 신뢰도 값이 추적 임계치를 상회하는 경우, 전자 디바이스 (102) 는 현재 비디오 프레임에 대한 오브젝트 검출을 스킵하고, 현재 비디오 프레임의 모션 추적 결과들에 기초하여 다음 비디오 프레임 상에서 모션-기반 추적을 계속 수행할 수도 있다. 다시 말해, 모션-기반 추적이 매우 좋지 않을 때에만, 예를 들어, 추적 신뢰도 값이 추적 임계치를 하회할 때에만, 오브젝트 검출이 수행될 수도 있다.

다. 오브젝트 검출이 수행될지 여부 및/또는 수행되는 방법을 고려할 때 다른 추적된 파라미터들이 이용될 수도 있다. 추적된 파라미터들의 예들로는 타깃 오브젝트의 구역, 윈도우 로케이션, 윈도우 사이즈, 스케일 레벨, 타깃 사이즈, 추적 및/또는 검출 신뢰도 값 또는 타깃 오브젝트의 효율적인 추적 및/또는 검출을 용이하게 하는데 이용될 수도 있는 다른 파라미터들을 포함할 수도 있다.

[0026] 추적된 파라미터에 기초하여 모션-기반 추적 및 오브젝트 검출을 순차적으로 수행하는 것은, 전자 디바이스 (102) 가 대규모의 연산들을 수행하는 일 없이 비디오 프레임 내에서 타깃 오브젝트를 추적 및/또는 검출하는 것을 가능하게 할 수도 있다. 구체적으로, 모션-기반 추적이 오브젝트 검출보다 덜 연산 집약적일 수도 있기 때문에, 전자 디바이스 (102) 는 현재 비디오 프레임 내에서 타깃 오브젝트를 정확히 추적하기 위해 모션-기반 추적이 이용될 수도 있는 경우 오브젝트 검출을 수행하는 것을 스킵할 수도 있다. 예를 들어, 추적 신뢰도 값이 특정 타깃 임계치를 초과한다고 전자 디바이스 (102) 가 결정하는 경우, 전자 디바이스 (102) 는 현재 비디오 프레임 내의 타깃 오브젝트의 로케이션 또는 존재를 정확히 결정하기 위해 현재 비디오 프레임 상에서 오브젝트 검출이 필요하지 않다고 결정할 수도 있다. 게다가, 많은 경우들에서 오브젝트 검출이 이로울 수도 있기 때문에, 전자 디바이스 (102) 는 추적 임계치 값과의 비교에 기초하여 모션-기반 추적이 부적절한 경우들에서 타깃 오브젝트를 더 정확히 검출하거나 또는 오브젝트 검출을 수행하기 위해 오브젝트 검출이 이용될 수도 있는 경우들을 결정할 수도 있다.

[0027] 일부 구성들에서, 현재 비디오 프레임 상에서 오브젝트 검출을 스킵하는 것보다는 오히려, 오브젝트 검출을 수행하는 프로세스를 좁히거나 또는 그 프로세스에 맞추기 위해 모션-기반 추적의 결과들 및/또는 메모리 버퍼 (110) 에 의해 제공된 부가적인 정보가 이용될 수도 있다. 예를 들어, 타깃 오브젝트가 모션-기반 추적 방법을 이용하여 정확히 추적될 수 없는 경우, 전자 디바이스 (102) 는 모션-기반 추적을 통해 제공된 파라미터들이 없는 것보다 더 적은 연산력 (computational power) 을 이용하여 오브젝트를 더 정확히 검출하기 위해 오브젝트 검출 동안 이용될 수도 있는 타깃 오브젝트와 연관된 로케이션, 윈도우 스케일 또는 다른 추적된 파라미터에 관한 정보를 여전히 추정 또는 획득할 수도 있다. 그에 따라, 모션-기반 추적이 추적 임계치를 초과하는 추적 신뢰도 값을 제공하지 않는 경우들에서조차도, 오브젝트 검출을 후속하여 수행할 때 모션-기반 추적의 결과들이 이용될 수도 있다.

[0028] 전자 디바이스 (102) 상의 뷰파인더 (131) 는 제 1 추적 영역 (133) 및 제 2 추적 영역 (135) 을 포함할 수도 있다. 터치스크린 (116) 을 이용하여 사용자에게 의해 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 양쪽이 특정될 수도 있다. 예를 들어, 사용자는 터치스크린 (116) 상의 포커스 링 (focus ring) 을 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 의 원하는 로케이션들로 드래그할 수도 있다. 요구되지 않았지만, 추적 영역들 중 하나는 고정될 수도 있다. 예를 들어, 제 1 추적 영역 (133) 은 오브젝트 (예를 들어, 걷고 있는 사람) 를 추적할 수도 있고 제 2 추적 영역 (135) 은 고정된 나무를 커버할 수도 있다. 하나의 구성에서, 제 2 추적 영역 (135) 은 전자 디바이스 (102) 상의 전체 터치스크린 (116) 을 커버할 수도 있다.

[0029] 전자 디바이스 (102) 는 비디오 프로세싱 모듈 (137) 을 포함할 수도 있다. 비디오 프로세싱 모듈 (137) 은 오버랩 (143) 을 포함할 수도 있다. 오버랩 (143) 은 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 사이의 오버랩의 양을 반영할 수도 있다. 예를 들어, 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 이 서로 전혀 오버랩되지 않는 경우, 오버랩 (143) 은 0% 일 수도 있다. 이와 마찬가지로, 제 1 추적 영역 (133) 이 제 2 추적 영역 (135) 과 완전히 오버랩되는 경우 (또는, 어떤 추적 영역이 더 큰지에 따라, 제 2 추적 영역 (135) 이 제 1 추적 영역 (133) 과 완전히 오버랩되는 경우), 오버랩 (143) 은 100% 일 수도 있다. 비디오 프로세싱 모듈 (137) 은 임계치 (145) 를 포함할 수도 있다. 도 13 에 대해 설명되는 바와 같이, 오버랩 (143) 은 임계치 (145) 와 비교되어 비디오 프로세싱이 수행되어야 하는지 여부를 결정할 수도 있다.

[0030] 비디오 프로세싱 모듈 (137) 은 또한 스크린 파티션 (147) 평선 (function) 을 포함할 수도 있다. 예를 들어, 도 11, 도 13, 도 15, 및 도 21 에 대해 설명되는 바와 같이, 스크린 파티션 (147) 은 뷰파인더 (132) 를 다수의 윈도우들로 파티셔닝하여 제 1 추적 영역 (133) 및 제 2 추적 영역 (135) 과 연관된 개개의 비디오 스트림들을 디스플레이할 수도 있다. 비디오 프로세싱 모듈 (137) 은 또한 지오메트리 (geometry) 추적 (149) 평선을 포함할 수도 있다. 예를 들어, 도 17 및 도 18 에 대해 설명되는 바와 같이, 지오메트리 추적 (149) 평선은 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 사이의 지오메트리를 추적할 수도 있다. 지오메트리는 뷰파인더 (131) 상에 디스플레이될 수도 있다. 비디오 프로세싱 모듈 (137) 은 또한 클러스터 형성 (151) 평선을 포함할 수도 있다. 클러스터 형성 (151) 평선은 지오메트리와 연관된 적어도 하나의 파라미터 (예를 들어, 질량 중심) 에 기초하여 클러스터들을 발생시킬 수도 있다. 예를 들어, 도 17 및 도 18 에 대해 설명되는 바와 같이, 각각의 클러스터는 제 1 추적 영역 (133) 과 제 2 추적 영역 (135) 사이에서 실질적으로

로 유사한 지오메트리를 갖는 비디오 프레임들을 포함할 수도 있다.

[0031] 도 2a 를 참조하면, 오브젝트 추적 및 검출 모듈 (204) 을 예시하는 블록 다이어그램이 도시된다. 오브젝트 추적 및 검출 모듈 (204) 은 전자 또는 무선 디바이스 내에서 구현될 수도 있다. 오브젝트 추적 및 검출 모듈 (204) 은 광학 플로우 모듈 (226) 및 추적 신뢰도 값 (228) 을 갖는 모션 추적기 (206) 를 포함할 수도 있다. 오브젝트 추적 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 추적기 (206) 및 오브젝트 검출기 (208) 에 제공될 수도 있는 캡처된 이전 비디오 프레임 (212) 과 연관된 데이터를 저장할 수도 있다. 오브젝트 추적 및 검출 모듈 (204), 모션 추적기 (206), 오브젝트 검출기 (208) 및 메모리 버퍼 (210) 는 도 1 과 관련되어 상술된 오브젝트 추적 및 검출 모듈 (104), 모션 추적기 (106), 오브젝트 검출기 (108) 및 메모리 버퍼 (110) 의 구성들일 수도 있다.

[0032] 모션 추적기 (206) 는 현재 비디오 프레임 (N) (224) 상에서 모션-기반 추적을 수행하는데 이용될 수도 있다. 예를 들어, 이전 비디오 프레임 (N-1) (222) 및 현재 비디오 프레임 (N) (224) 은 (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. 부가적인 비디오 프레임들이 오브젝트 추적 및 검출 모듈 (204) 에 의해 획득되고 프로세싱될 수도 있다. 이전 비디오 프레임 (N-1) (222) 이 모션 추적기 (206) 에 제공될 수도 있다. 게다가, 메모리 버퍼 (210) 는, 본 명세서에서 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 추적 및/또는 검출된 장소를 특정할 수도 있는 이전 비디오 프레임 (N-1) (222) 에 관한 추적 결과들을 퓨전 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.

[0033] 모션 추적기 (206) 는 비디오 프레임들의 시퀀스에서 현재 비디오 프레임 (N) (224) 을 후속하여 수신할 수도 있다. 모션 추적기 (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 추적기 (206) 는 광학 플로우 모듈 (226) 을 이용하여 현재 비디오 프레임 (N) (224) 상에서 오브젝트의 모션을 추적할 수도 있다. 광학 플로우 모듈 (226) 은 현재 비디오 프레임 (N) (224) 상에서 오브젝트의 모션-기반 추적을 수행하기 위한 하드웨어 및/또는 소프트웨어를 포함할 수도 있다. 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 을 비교하는 것에 의해, 모션 추적기 (206) 는 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 에 있을 가능성과 연관된 추적 신뢰도 값 (228) 을 결정할 수도 있다. 하나의 예에서, 추적 신뢰도 값 (228) 은 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 또는 현재 비디오 프레임 (N) (224) 내의 윈도우 내에 있다는 확실성의 퍼센티지에 기초한 실수 (예를 들어, 0 과 1 사이) 이다.

[0034] 오브젝트 검출기 (208) 는 현재 비디오 프레임 (N) (224) 상에서 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 추적된 파라미터에 기초하여 현재 비디오 프레임 (N) (224) 상에서 오브젝트 검출을 수행할 수도 있다. 추적된 파라미터는 타깃 오브젝트가 정확히 추적되고 있을 가능성에 대응하는 추적 신뢰도 값 (228) 을 포함할 수도 있다. 더 구체적으로는, 추적된 파라미터는 추적 신뢰도 값 (228) 과 추적 임계치 (250) 의 비교를 포함할 수도 있다. 추적된 파라미터는 또한 메모리 버퍼 (210) 로부터 제공된 정보를 포함할 수도 있다. 오브젝트를 검출할 때 이용될 수도 있는 추적된 파라미터들의 일부 예 들로는 구역, 윈도우 로케이션, 윈도우 사이즈, 또는 오브젝트 검출을 수행할 때 파라미터로서 오브젝트 검출기 (208) 에 의해 이용될 수도 있는 다른 정보를 포함한다.

[0035] 오브젝트 검출기 (208) 는 스캐너 로케이터 (230) 를 포함할 수도 있다. 스캐너 로케이터 (230) 는 윈도우 로케이션 선택기 (232) 및 랜더마이저 (randomizer; 234) 를 포함할 수도 있다. 윈도우 로케이션 선택기 (232) 는 비디오 프레임 내의 다수의 윈도우들을 선택할 수도 있다. 예를 들어, 비디오 프레임은 다수의 윈도우들을 포함할 수도 있는데, 이 윈도우들 각각은 연관된 로케이션 및 사이즈를 갖는다. 하나의 구성에서, 각각의 비디오 프레임은 다수 (예를 들어, 대략 10,000) 개의 오버래핑 윈도우들로 분할되고, 이 윈도우들 각각은 비디오 프레임에서의 총 픽셀들의 프랙션 (fraction) 을 포함한다. 대안적으로, 임의의 적합한 개수의 윈도우들이 존재할 수도 있고 이들은 오버랩되지 않을 수도 있다. 스캐너 로케이터 (230) 내의 윈도우 로케이션 선택기 (232) 는 타깃 오브젝트를 식별하려고 시도하는 윈도우의 로케이션을 선택할 수도 있다. 랜더

마이저 (234) 는 오브젝트를 검출하기 위한 가변 사이즈들 및 로케이션들의 윈도우들을 랜덤하게 선택할 수도 있다. 일부 구성들에서, 랜더마이저 (234) 는 비디오 프레임 내의 윈도우들을 랜덤하게 선택한다. 대안적으로, 랜더마이저 (234) 는 하나 이상의 팩터들 (factors) 에 기초하여 윈도우들을 더 정밀하게 선택할 수도 있다. 예를 들어, 랜더마이저 (234) 는 오브젝트가 로케이팅될 가능성이 가장 많은 장소의 구역, 사이즈 또는 일반적 로케이션에 기초하여 윈도우들의 선택을 제한할 수도 있다. 이 정보는 메모리 버퍼 (210) 를 통해 획득될 수도 있고, 또는 모션-기반 추적을 통해 획득될 수도 있는데, 이 모션-기반 추적은 전적으로 의존되 기에는 충분히 정확하지 않지만, 오브젝트 검출을 수행할 때 도움이 되는 정보를 제공할 수도 있다. 그에 따라, 랜더마이저 (234) 가 검색하기 위해 다수의 윈도우들을 랜덤하게 선택할 수도 있지만, 오브젝트 검출기 (208) 에 제공된 정보에 기초하여, 윈도우들의 선택이 좁혀질 수도 있고, 그에 따라 완전히 랜덤하지 않을 수도 있다.

[0036] 오브젝트 검출기 (208) 는 또한 특정 사이즈의 윈도우를 드로잉 또는 선택하기 위해 이용될 수도 있는 스캐너 스케일러 (236) 를 포함할 수도 있다. 이미지가 특정 윈도우 내에 있는지 여부를 검출하기 위해 오브젝트를 검출하거나 또는 윈도우들의 선택을 오리진얼 이미지와 비교할 때 윈도우들의 사이즈를 좁히기 위해 스캐너 로케이터 (230) 에 의해 윈도우 사이즈가 이용될 수도 있다. 스캐너 스케일러 (236) 는 오브젝트를 정의할 때 초기에는 특정 사이즈들 또는 스케일 레벨들의 하나 이상의 윈도우들을 선택할 수도 있고 또는, 대안적으로, 메모리 버퍼 (210) 로부터 제공된 정보에 기초하여 특정 사이즈들 또는 스케일 레벨들의 하나 이상의 윈도우들을 드로잉할 수도 있다.

[0037] 분류기 (238) 는 타깃 오브젝트 중 일부 또는 전부가 특정 윈도우에서 발견되는지 여부를 결정하는데 이용될 수도 있다. 일부 구성들에서, 분류기 (238) 는 타깃 오브젝트가 특정 윈도우 또는 서브-윈도우 내에서 검출되는지 여부를 나타내기 위해 각각의 윈도우에 대한 2진 값을 생성할 수도 있다. 이러한 분류 (예를 들어, 2진 분류) 는 오브젝트 검출기 (208) 에 의해 검색된 각각의 윈도우에 대해 수행될 수도 있다. 구체적으로, 분류기 (238) 는 오브젝트가 검출되는 각각의 윈도우에 대해서는 2진수 1 을, 그리고 오브젝트가 검출되지 않는 각각의 윈도우에 대해서는 2진수 0 을 발생시킬 수도 있다. 1들 및 0들의 개수 또는 조합에 기초하여, 오브젝트 검출기 (208) 는 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 내에 존재할 가능성을 나타내는 검출 신뢰도 값 (240) 을 결정할 수도 있다. 일부 구성들에서, 검출 신뢰도 값 (240) 은, 오브젝트가 정확히 검출되었을 퍼센티지 또는 확률을 나타내는, 0 과 1 사이의 실수이다.

[0038] 오브젝트 검출기 (208) 는, 구역, 타깃 사이즈, 윈도우 사이즈, 스케일 레벨, 윈도우 로케이션 및 하나 이상의 신뢰도 값들을 포함하는, 다양한 추적된 파라미터들에 따라 오브젝트 검출을 수행할 수도 있다. 일단 비디오 프레임의 윈도우들 또는 윈도우들의 서브세트가 검색되고 오브젝트 검출기 (208) 가 검색된 윈도우 각각에 대해 2진 값을 획득한다면, 오브젝트 검출기 (208) 는 가장 높은 신뢰도를 갖는 현재 비디오 프레임 상의 로케이션 또는 구역뿐만 아니라 윈도우 사이즈를 결정할 수도 있다. 이 로케이션 및 윈도우 사이즈는 후속 추적 및 검출에 이용되어 타깃 오브젝트를 더 정확히 추적 및/또는 검출할 수도 있다.

[0039] 타깃 오브젝트를 검출함에 있어서 다양한 기법들이 오브젝트 검출기 (208) 에 의해 이용될 수도 있다. 하나의 구성에서, 타깃 오브젝트를 검출하는 것은, 모든 가능한 윈도우 로케이션에서 그리고 모든 가능한 윈도우 사이즈에서 윈도우들에 대한 2진 분류를 수행하는 것을 포함할 수도 있다. 그러나, 모든 가능한 윈도우를 검색하는 것은 리소스 집약적이다. 따라서, 다른 구성에서, 오브젝트 검출기는 비디오 프레임에서의 모든 가능한 윈도우들보다는 오히려, 윈도우 로케이션들 및 사이즈들의 서브세트를 검색할 수도 있다. 예를 들어, 오브젝트 검출기 (208) 는 모든 가능한 윈도우들의 1% 를 검색할 수도 있다. 그 후에, 검출이 성공적이지 못한 경우 (예를 들어, 검출 신뢰도 값 (240) 이 검출 임계치 (252) 미만인 경우), 후속하는 캡처된 프레임에서 보다 높은 퍼센티지, 예를 들어, 2% 의 윈도우 로케이션들이 검색될 수도 있다. 검색된 윈도우 로케이션들의 퍼센티지에 있어서의 스텝 (step) 은 균일하거나, 불균일하거나, 느리거나 또는 빠를 수도 있다, 즉, 연속 프레임들은 1%, 2%, 3%, 4% 또는 1%, 2%, 4%, 8% 를 가질 수도 있다. 하나의 구성에서, 높은 검출 신뢰도 값에 응답하여, 즉, 타깃 오브젝트가 다음 비디오 프레임이라는 것을 보장하기 위해, 검색된 프레임들의 퍼센티지가 매우 높게 설정될 수도 있다 (예를 들어, 80%, 90%, 100%). 예를 들어, 검출 및 추적 임계치 값 (256) 을 초과한 검출 및 추적 신뢰도 값에 응답하여, 검색된 프레임들의 퍼센티지는 적어도 80% 로 점프할 수도 있다. 대안적으로, 퍼센티지는 60%, 70%, 90% 등으로 점프할 수도 있다. 부가적으로, 검출 및 추적 임계치 값에 대한 임의의 적합한 값, 예를 들어, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85 등이 이용될 수도 있다. 더욱이, 랜더마이저 (234) (난수 발생기) 에 기초하여, 검색된 윈도우들의 퍼센티지가 랜덤하게 결정될 수도 있다, 예를 들어, 1% 와 15% 사이의 윈도우들의 랜덤한 퍼센티지가 캡처된 프레임에서 검색될 수도 있다. 윈

도우 로케이션들 모두의 서브세트를 검색하는 것에 의해, 오브젝트 검출은 전자 디바이스 (102) 에서 보다 적은 리소스들을 이용할 수도 있다.

[0040] 본 명세서에서 설명되는 기법들은 각각의 로케이션에 대한 윈도우 사이즈들의 서브세트를 검색할 수도 있다. 각각의 윈도우 사이즈는 본 명세서에서 스케일 레벨이라고 지칭될 수도 있고, 각각의 스케일 레벨은 특정 윈도우 사이즈에 대응한다. 예를 들어, 20 개의 가능한 스케일 레벨들이 존재할 수도 있다. 모든 20 개의 스케일 레벨들을 검색하는 것보다는 오히려, 스케일 레벨들 또는 윈도우 사이즈들의 서브셋이 각각의 윈도우 로케이션에서 검색될 수도 있다.

[0041] 본 명세서에서 설명되는 기법들은 또한 검색된 윈도우 로케이션들 및 사이즈들에 맞추기 위해 메모리 버퍼 (210) 로부터의 피드백을 이용할 수도 있다. 다시 말해, 타깃 오브젝트가 성공적으로 검출 및/또는 추적되었던 마지막으로 캡처된 비디오 프레임의 로케이션 및 사이즈는 현재 비디오 프레임 (N) (224) 을 검색하기 위한 시작 포인트로서 이용될 수도 있다. 예를 들어, 최근 비디오 프레임에서 타깃 오브젝트가 검출되고 추적된 경우 (즉, 최근에 캡처된 비디오 프레임에 대한 검출 및 추적 신뢰도 값 (256) 이 검출 및 추적 임계치를 상회하는 경우), 스캐너 로케이터는 최근 프레임과 연관된 로케이션 및 사이즈에서 현재 캡처된 프레임을 검색하는 것을 시작할 수도 있다. 예를 들어, 타깃 오브젝트가 광학 시스템의 시야 밖으로 이동하거나 또는 멀리 떨어져 사라지는 경우, 타깃 오브젝트가 광학 시스템의 시야를 떠났거나 또는 멀리 떨어져 사라졌을 때와 동일한 사이즈에서 타깃 오브젝트가 다시 나타날 가능성이 더 많을 수도 있다. 따라서, 오브젝트 검출을 수행할 때 후속 비디오 프레임들에서 타깃 오브젝트를 검출하기 위해 사이즈 또는 사이즈들의 범위가 예측될 수도 있다.

[0042] 캡처된 비디오 프레임 (N) (224) 에서 검색된 윈도우 로케이션들 및 윈도우 사이즈들의 검색 범위는, 최근 비디오 프레임 (예를 들어, 이전 비디오 프레임 (N-1) (222)) 에서의 타깃 오브젝트와 연관된 윈도우 로케이션 및 윈도우 사이즈와 유사한 것들로 제한될 수도 있다. 본 명세서에서 사용되는 바와 같이, 용어 "검색 범위" 는 비디오 프레임에서 타깃 오브젝트를 검출 및/또는 추적할 때 활용될 수도 있는 후보 윈도우 로케이션들 또는 후보 윈도우 사이즈들 (또는 이들 양쪽) 의 세트를 지칭한다. 예를 들어, 검색된 윈도우 로케이션들의 서브셋은 타깃 오브젝트가 최근 비디오 프레임에서 발견되었던 장소에 기초하여 현재 비디오 프레임 (N) (224) 의 일 부분, 예를 들어, 현재 비디오 프레임 (N) (224) 의 4분면들 또는 절반들 중 하나 내에서부터 선택될 수도 있다. 다시 말해, 검색 공간은 타깃 오브젝트가 마지막으로 추적 또는 검출되었던 장소 인근으로 제한될 수도 있다. 이와 유사하게, 각각의 윈도우 로케이션에 대해 검색된 프레임들의 사이즈들은 타깃화된 오브젝트가 최근 비디오 프레임에서 발견되었던 윈도우의 사이즈에 기초하여 제한될 수도 있다. 예를 들어, 오브젝트가 8 의 스케일 레벨을 갖는 윈도우를 이용하여 최근 프레임에서 검출된 경우, 스캐너 스케일러 (236) 는 8 플러스 또는 마이너스 3, 즉, 스케일 레벨들 5 내지 11 의 현재 비디오 프레임 (N) (224) 에 대한 윈도우 스케일 레벨들만을 단지 선택할 수도 있다. 이것은 낮은 확률 검색을 더욱 없애고 오브젝트 검출의 효율을 증가시킬 수도 있다. 대안적으로, 최근 (현재가 아님) 비디오 프레임이 타깃 오브젝트를 검출하지 못한 경우 (즉, 최근 비디오 프레임에 대한 검출 및 추적 신뢰도 값 (256) 이 검출 및 추적 임계치를 하회하는 경우), 오브젝트 검출기 (208) 는 검색되는 검색 공간 (윈도우 로케이션들) 을 확장시킬 수도 있다, 예를 들어, 보다 넓은 범위의 이미지 또는 전체 이미지가 검색 대상이 될 수도 있다.

[0043] 오브젝트 추적 및 검출 모듈 (204) 은 단일 윈도우를 형성하기 위해 다수의 윈도우들을 병합시키기 위한 퓨전 모듈 (260) 을 포함할 수도 있다. 초기에는 2 개의 신뢰도 값들이 존재한다: 오브젝트 검출기 (208) 로부터의 검출 신뢰도 값 (240) 및 모션 추적기 (206) 로부터의 추적 신뢰도 값 (225). 퓨전 모듈 (260) 은 2 개의 신뢰도 값들을 검출 및 추적 신뢰도 값 (256) 으로 조합할 수도 있다 (예를 들어, 더 큰 신뢰도 값을 고를 수도 있다). 검출 및 추적 신뢰도 값 (256) 은 타깃 오브젝트가 비디오 프레임 상에서 식별되었는지 여부를 나타낼 수도 있다. 하나의 구성에서, 검출 및 추적 신뢰도 값 (256) 은 0 과 1 사이의 실수일 수도 있는데, 여기서 0 은 타깃 오브젝트가 특정 비디오 프레임에서 식별되었다는 가장 낮은 가능한 신뢰도를 나타내고 1 은 타깃 오브젝트가 특정 비디오 프레임에서 식별되었다는 가장 높은 가능한 신뢰도를 나타낸다. 다시 말해, 검출 및 추적 신뢰도 값 (256) 은 타깃 오브젝트가 발견되었을 가능성의 전체 표시로서 기능할 수도 있다. 게다가, 검출 및 추적 신뢰도 값 (256) 은 다음 비디오 프레임에서 검색하기 위한 윈도우들의 윈도우 로케이션, 윈도우 사이즈 또는 퍼센티지를 결정하기 위해 이용되는 파라미터일 수도 있다. 퓨전 모듈 (260) 은 현재 비디오 프레임 (N) (224) 에 관한 정보를 메모리 버퍼 (210) 에 제공하는데 이용될 수도 있다. 하나의 예에서, 퓨전 모듈 (260) 은 추적된 윈도우 (242) (예를 들어, 윈도우 로케이션 (244), 윈도우 사이즈 (246) 등) 및 검출 및 추적 신뢰도 값 (256) 에 관한 정보를 메모리 버퍼 (210) 에 제공할 수도 있다. 퓨전 모듈 (260)

은 모션 추적기 (206) 및 오브젝트 검출기 (208) 로부터의 추적 결과들 (예를 들어, 바운딩 박스들) 을 이용하여 조합된 추적 결과 (예를 들어, 바운딩 박스) 를 형성하고 검출 및 추적 신뢰도 값 (256) 을 계산할 수도 있다.

[0044] 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. 하나의 구성에서, 메모리 버퍼 (210) 는 이전 비디오 프레임 (N-1) (222) 에 대응하는 정보를 포함할 수도 있는 캡처된 이전 비디오 프레임 (212) 을 저장한다. 캡처된 이전 비디오 프레임 (212) 은 각각의 윈도우 (242) 에 대한 로케이션 (244), 윈도우 사이즈 (246) 및 (예를 들어, 분류기 (238) 로부터의) 2진 판정 (248) 을 포함하는, 하나 이상의 윈도우들 (242) 에 관한 정보를 포함할 수도 있다. 캡처된 이전 비디오 프레임 (212) 은 또한 추적 임계치 (250), 검출 임계치 (252) 및 검출 및 추적 임계치 (254) 를 포함할 수도 있다. 추적 임계치 (250) 는 추적 신뢰도 레벨이 추적 임계치 (250) 보다 더 큰지 여부를 결정하기 위해 (258) 오브젝트 추적 및 검출 모듈 (204) 상의 회로부 (예를 들어, 신뢰도 레벨 비교기) 또는 모션 추적기 (206) 에 제공될 수도 있다. 검출 임계치 (252) 는 검출 신뢰도 값 (240) 이 검출 임계치 (252) 보다 더 큰지 여부를 결정하기 위해 오브젝트 추적 및 검출 모듈 (204) 상의 다른 회로부 또는 오브젝트 검출기 (208) 에 제공될 수도 있다. 검출 및 추적 임계치 (254) 는 추적 임계치 (250) 및 검출 임계치 (252) 에 기초하여 조합된 값일 수도 있다. 검출 및 추적 임계치 (254) 는 검출 및 추적 신뢰도 값 (256) 과 비교되어 모션-기반 추적 및 오브젝트 검출을 위한 조합된 신뢰도 값을 결정할 수도 있다. 임계치들 각각은 타깃 오브젝트가 비디오 프레임 내에 로케이팅될 가능성에 기초할 수도 있다. 오브젝트 추적 및 검출 모듈 (204) 은 특정 검출 및 추적 신뢰도 값 (256) 이 획득될 때까지 현재 비디오 프레임 (N) (224) 상에서 모션-기반 추적 및/또는 검출을 수행할 수도 있다. 게다가, 모션-기반 추적 및 오브젝트 검출은 다수의 비디오 프레임들의 시퀀스에서 각각의 비디오 프레임 상에서 수행될 수도 있다.

[0045] 모션-기반 추적 및 오브젝트 검출을 수행하는 것은, 추적된 파라미터에 기초하여 모션-기반 추적 다음에 오브젝트 검출을 순차적으로 수행하는 것을 포함할 수도 있다. 특히, 본 시스템들 및 방법들은 2-단계 추적 및 검출 접근법을 구현할 수도 있다. 모션-기반 추적이 이용되는 오브젝트 검출로서 실제 오브젝트 식별보다는 오히려, 장면의 상대 모션에 기초하기 때문에, 모션-기반 추적이 오브젝트 검출을 수행하는 것보다 전자 디바이스에서 덜 리소스 집약적일 수도 있다. 이에 따라, 오브젝트 검출기 (208) 대신에 모션 추적기 (206) 를 이용하는 것이 더 효율적일 수도 있는데, 여기서 타깃 오브젝트는 오브젝트 검출을 또한 수행하는 일 없이 정확히 추적될 수도 있다.

[0046] 그에 따라, 오브젝트 검출기 (208) 와 병행하여 모션 추적기 (206) 를 이용하기보다는 오히려, 오브젝트 추적 및 검출 모듈 (204) 은 단지 모션 추적기 (206) 가 불충분한 경우 오브젝트 검출기 (208) 를 이용한다, 즉, 모션 추적 및 오브젝트 검출 (조금이나마 수행된다면) 은 병행하는 것 대신에 순차적으로 수행된다. 추적이 수행되는 각각의 비디오 프레임에 대해, 모션 추적기 (206) 는 추적 신뢰도 값 (228) 을 생성할 수도 있는데, 이 추적 신뢰도 값은, 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 에 있을 가능성을 나타내는, 0 과 1 사이의 실수일 수도 있다.

[0047] 2-단계 추적 및 검출 접근법의 하나의 구성에서, 모션 추적기 (206) 는 우선 현재 비디오 프레임 (N) (224) 상에서 모션-기반 추적을 수행할 수도 있다. 모션 추적기 (206) 는 모션-기반 추적 프로세스에 기초하여 추적 신뢰도 값 (228) 을 결정할 수도 있다. 메모리 버퍼 (210) 에 의해 제공된 추적 임계치 (250) 및 추적 신뢰도 값 (228) 을 이용하여, 오브젝트 추적 및 검출 모듈 (204) 내의 회로부 (예를 들어, 신뢰도 레벨 비교기) 는 추적 신뢰도 값 (228) 이 추적 임계치 (250) 를 초과하는지 여부를 결정할 수도 있다 (258). 추적 신뢰도 값 (228) 이 추적 임계치 (250) 보다 더 큰 경우, 오브젝트 추적 및 검출 모듈 (204) 은 오브젝트 검출을 수행하는 것을 스킵하고 추적 결과를 퓨전 모듈 (260) 에 제공하여 출력 (262) 을 생성할 수도 있다. 출력 (262) 은 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 내에 있다는 표시를 포함할 수도 있다. 게다가, 출력 (262) 은 타깃 오브젝트에 관한 추가적인 정보를 포함할 수도 있다.

[0048] 추적 신뢰도 값 (228) 이 추적 임계치 (250) 를 초과하지 않는 경우, 오브젝트 검출기 (208) 는 현재 비디오 프레임 (N) (224) 상에서 오브젝트 검출을 후속하여 수행할 수도 있다. 오브젝트 검출은 현재 비디오 프레임 (N) (224) 내의 윈도우들 전부 또는 그의 서브세트 상에서 수행될 수도 있다. 오브젝트 검출기 (208) 는 또한 메모리 버퍼 (210) 로부터 제공된 정보 및/또는 모션-기반 추적의 결과들에 기초하여 윈도우들, 윈도우 사이즈들 또는 다른 검출 기준들의 서브세트를 선택할 수도 있다. 오브젝트 검출기 (208) 에 제공된 하나 이상의 추적된 파라미터들에 기초하여 더 또는 덜 강건한 (robust) 프로세스를 이용하여 오브젝트 검출이 수행될 수도 있다. 오브젝트 검출기 (208) 는 검출 신뢰도 값 (240) 을 결정하고 그 검출 신뢰도 값 (240) 을 검출

임계치 (252) 와 비교할 수도 있다. 검출 신뢰도 값 (240) 이 검출 임계치 (252) 를 상회하는 경우, 오브젝트 검출기 (208) 는 검출 결과를 퓨전 모듈 (260) 에 제공하여 출력 (262) 을 생성할 수도 있다. 출력 (262) 은 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 내에 있다는 표시를 포함하거나 및/또는 검출된 오브젝트에 관한 부가적인 정보를 포함할 수도 있다.

[0049] 대안적으로, 검출 신뢰도 값 (240) 이 검출 임계치 (252) 이하인 경우, 오브젝트 검출기 (208) 는 현재 비디오 프레임 (N) (224) 내의 더 많은 개수의 윈도우들을 검색하는 것과 같은, 더 강건한 방법을 이용하여 오브젝트 검출을 다시 수행할 수도 있다. 오브젝트 검출기 (208) 는 만족스러운 검출 신뢰도 값 (240) 이 획득될 때까지 오브젝트 검출의 프로세스를 반복할 수도 있다. 일단 만족스러운 검출 신뢰도 값 (240) 이 획득되어 현재 비디오 프레임 내의 타깃 오브젝트가 식별된다면, 오브젝트 추적 및 검출 모듈 (204) 은 다음 비디오 프레임 상에서 추적 및 검출을 수행하는데 이용될 수도 있다.

[0050] 도 2b 를 참조하면, 오브젝트 추적 및 검출 모듈 (204) 내의 컴포넌트들을 구현하는 프로세서 (264) 의 특정 예시적 실시형태가 도시된다. 도 2a 에 도시된 바와 같이, 오브젝트 추적 및 검출 모듈 (204) 은 프로세서 (264) 에 의해 구현될 수도 있다. 상이한 컴포넌트들을 구현하기 위해 상이한 프로세서들이 이용될 수도 있다 (예를 들어, 하나의 프로세서는 모션 추적기 (206) 를 구현할 수도 있고, 다른 프로세서는 오브젝트 검출기 (208) 를 구현하는데 이용될 수도 있으며 또 다른 프로세서는 메모리 버퍼 (210) 를 구현하는데 이용될 수도 있다).

[0051] 도 3 을 참조하면, 모션-기반 추적 및 오브젝트 검출을 수행하기 위한 방법 (300) 의 특정 예시적 실시형태의 플로우차트가 도시된다. 방법 (300) 은 전자 디바이스 (102), 예를 들어, 오브젝트 추적 및 검출 모듈 (104) 에 의해 구현될 수도 있다. 전자 디바이스 (102) 는 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 을 비교하는 것에 의해 현재 비디오 프레임 (N) (224) 에 대한 모션-기반 추적을 수행할 수도 있다 (302). 오브젝트를 추적하는 것은 이미지들의 쌍들 사이의 포인트들을 추적하는 것에 의해 메디안 플로우 방법을 이용하여 수행될 수도 있다. 모션-기반 추적의 다른 방법들이 또한 이용될 수도 있다. 부가적으로, 모션-기반 추적은 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.

[0052] 전자 디바이스 (102) 는 추적 신뢰도 값 (228) 을 결정할 수도 있다 (304). 추적 신뢰도 값 (228) 은 타깃 오브젝트가 정확히 추적되었을 가능성 또는 확실성을 나타낼 수도 있다. 전자 디바이스 (102) 는 추적 신뢰도 값 (228) 이 추적 임계치 (250) 보다 더 큰지 여부를 결정할 수도 있다 (306). 추적 신뢰도 값 (228) 이 추적 임계치 (250) 보다 더 큰 경우, 전자 디바이스 (102) 는 다음 비디오 프레임에 대해 모션-기반 추적을 수행할 수도 있다 (308). 게다가, 전자 디바이스 (102) 는 모션-기반 추적의 결과에 기초하여 현재 비디오 프레임 (N) (224) 상에서 오브젝트 검출을 수행하는 것을 스킵할 수도 있다. 다시 말해, 모션 추적이 매우 좋지 않을 때에만, 즉, 추적 신뢰도 값 (228) 이 추적 임계치 (250) 보다 더 크지 않은 경우, 오브젝트 검출이 현재 비디오 프레임 (N) (224) 에 대해 수행될 수도 있다. 그러나, 추적 신뢰도 값 (228) 이 추적 임계치 (250) 보다 더 크지 않은 경우, 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 에 대해 오브젝트 검출을 수행할 수도 있다 (310). 전자 디바이스 (102) 는 모션-기반 추적에 순차적으로 오브젝트 검출을 수행할 수도 있다. 일부 구성들에서, 오브젝트 검출은 가변 강건성으로 다수회 수행되어 보다 높은 검출 신뢰도 값 (240) 을 획득할 수도 있다.

[0053] 도 4 를 참조하면, 모션-기반 추적을 수행하기 위한 방법 (400) 의 특정 예시적 실시형태의 플로우차트가 도시된다. 방법 (400) 은 전자 디바이스 (102), 예를 들어, 오브젝트 추적 및 검출 모듈 (104) 에 의해 구현될 수도 있다. 전자 디바이스 (102) 는 바운딩 박스를 이용하여 타깃 오브젝트를 식별할 수도 있다 (402). 오브젝트를 식별하는 것 (402) 은 관심있는 오브젝트가 선택되는 터치스크린 (116) 또는 다른 입력 방법을 이용하여 수동으로 수행될 수도 있다. 다수의 오브젝트들이 유사한 방법으로 식별될 수도 있다. 게다가, 추적될 오브젝트를 식별하기 위해 다른 입력 방법들이 이용될 수도 있다. 하나의 예에서, 바운딩 박스를 타깃 오브젝트 주위로 수동으로 드로잉하는 것에 의해 오브젝트가 식별된다.

[0054] 전자 디바이스 (102) 는 바운딩 박스 내의 그리드 상의 포인트들을 초기화할 수도 있다 (404). 그리드 상의 포인트들은 바운딩 박스 전반에 걸쳐 균일하게 이격될 수도 있다. 게다가, 포인트들은 2 개의 이미지들 (예를 들어, 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224)) 사이의 그리드 상에서 추적될 수도 있다 (406). 하나의 예에서, 포인트들은 이미지들 사이에서 스퀘스 모션 플로우를 발생시키는 루카스-카나데 (Lucas-Kanade) 추적기에 의해 추적된다. 전자 디바이스 (102) 는 2 개의 이미지들 (예를 들어, 이전

비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 사이의 추적 에러를 추정할 수도 있다 (408).

추적 에러를 추정하는 것 (408) 은 추적된 포인트들의 각각의 포인트에 에러 값을 할당하는 것을 포함할 수도 있다. 게다가, 추적 에러를 추정하는 것 (408) 은, 예를 들어, 순방향-역방향 에러 (forward-backward error), 정규화된 상호 상관 (NCC) 및 차분 제곱합 (sum-of-square differences) 을 포함하는, 다양한 방법들을 이용하여 수행될 수도 있다. 추적 신뢰도 값 (228) 을 획득하기 위해 그리고 궁극적으로 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 에 있을 가능성을 결정하는 것을 위해 추정된 추적 에러가 이용될 수도 있다.

하나의 구성에서, 추적 신뢰도 값 (228) 은 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 에서의 추적된 윈도우 사이의 정규화된 상호 상관 (NCC) 을 계산하는 것에 의해 획득될 수도 있다. 추적 에러는 또한, 도 5 와 관련하여 아래에 더 상세히 설명되는 순방향-역방향 에러 추정을 포함하는, 부가적인 기법들을 이용하여 추정될 수도 있다. 게다가, 전자 디바이스 (102) 는 바깥쪽 (outlying) 포인트 예측들을 필터링할 수도 있다 (410). 예를 들어, 전자 디바이스는 최악의 예측들의 50% 를 필터링할 수도 있다.

나머지 예측들은 바운딩 박스의 변위를 추정하는데 이용될 수도 있다.

[0055] 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 추적 프로세스는 그 후에 다음 비디오 프레임에 대해 반복될 수도 있거나 또는, 추적 신뢰도 값 (228) 이 추적 임계치 (250) 이하인 경우에는, 모션-기반 추적 프로세스는 타깃 오브젝트가 정확히 추적될 수도 있을 때까지 다음 비디오 프레임에 대해 중단될 수도 있다. 일부 구성들에서, 현재 비디오 프레임 (N) (224) 에 대한 모션-기반 추적이 만족스러운 결과를 제공하지 못한 경우, 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 상에서 오브젝트 검출을 수행하여 타깃 오브젝트를 로케이팅함에 있어서 보다 높은 레벨의 신뢰도를 획득할 수도 있다. 일부 구성들에서, 모션-기반 추적이 만족스러운 결과들을 생성할 수 없는 경우 (예를 들어, 타깃 오브젝트가 비디오 프레임의 범위 밖으로 이동할 때), 타깃 오브젝트가 검출될 때까지 오브젝트 검출이 임의의 후속 비디오 프레임들 상에서 수행될 수도 있다.

[0056] 도 5 를 참조하면, 순방향-역방향 에러에 기초하여 모션-기반 추적에서 추적 에러를 추정하기 위한 방법 (500) 의 특정 예시적 실시형태의 플로우차트가 도시된다. 방법 (500) 은 전자 디바이스 (102) (예를 들어, 오브젝트 추적 및 검출 모듈 (104)) 에 의해 구현될 수도 있다. 일부 구성들에서, 전자 디바이스 (102) 는 추적된 윈도우들 사이의 정규화된 상호 상관 (NCC) 을 계산할 수도 있다. 정규화된 상호 상관 (NCC) 은 추적 신뢰도 값 (228) 을 결정하는데 이용될 수도 있다. 전자 디바이스 (102) 는 또한 정규화된 상호 상관 (NCC) 과 상보적인 다양한 추적 에러 추정 기법들 (예를 들어, 순방향-역방향 에러, 차분 제곱합) 을 이용할 수도 있다. 순방향-역방향 에러 추정을 이용하는 예에서, 전자 디바이스 (102) 는 이전 비디오 프레임 (N-1) (222) 과 현재 비디오 프레임 (N) (224) 사이의 순방향 추적을 수행하여 순방향 궤적을 결정할 수도 있다 (502). 순방향 추적은 k 개의 단계들에 대해 순방향인 이미지를 추적하는 것을 포함할 수도 있다. 결과적인 순방향 궤적은 $(x_t, x_{t+1}, \dots, x_{t+k})$ 와 동일할 수도 있는데, 여기서 x_t 는 시간에 있어서의 포인트 로케이션이고 k 는 이미지들의 시퀀스의 길이를 나타낸다. 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 과 이전 비디오 프레임 (N-1) (222) 사이의 역방향 추적을 수행하여 역방향 궤적을 결정할 수도 있다 (504). 결과적인 역방향 궤적은 $(\hat{x}_t, \hat{x}_{t+1}, \dots, \hat{x}_{t+k})$ 와 동일할 수도 있는데, 여기서 $\hat{x}_{t+k} = x_{t+k}$ 이다.

[0057] 전자 디바이스 (102) 는 순방향 궤적과 역방향 궤적 사이의 순방향-역방향 에러를 결정할 수도 있다 (506). 순방향-역방향 에러는 순방향 궤적과 역방향 궤적 사이의 거리로서 정의될 수도 있다. 게다가, 다양한 거리들이 궤적 비교를 위해 정의될 수도 있다. 하나의 구성에서, 순방향-역방향 에러를 결정할 때 확인 (validation) 궤적의 초기 포인트와 종단 포인트 사이의 유클리디안 (Euclidean) 거리가 이용될 수도 있다. 하나의 구성에서, 순방향-역방향 에러는 추적 에러로서 이용될 수도 있고, 이 추적 에러는 추적 신뢰도 값 (228) 을 결정하는데 이용될 수도 있다.

[0058] 도 6 을 참조하면, 오브젝트 검출을 수행하기 위한 방법 (600) 의 특정 예시적 실시형태의 플로우차트가 도시된다. 방법 (600) 은 전자 디바이스 (102) (예를 들어, 오브젝트 추적 및 검출 모듈 (104)) 에 의해 구현될 수도 있다. 전자 디바이스 (102) 는 현재 비디오 프레임 (N) (224) 에서의 윈도우 로케이션들 및 사이즈들의 서브셋을 검색하는 것에 의해 현재 비디오 프레임 (N) (224) 상에서 오브젝트 검출 및 모션-기반 추적을 수행할 수도 있다 (602).

[0059] 전자 디바이스 (102) 는 검출 및 추적 신뢰도 값 (256) 을 결정할 수도 있다 (604). 검출 및 추적 신뢰도

값 (256) 은 타깃 오브젝트가 현재 비디오 프레임 (N) (224) 에서 발견되는지 또는 특정 윈도우 내에서 발견되는지 여부의 신뢰도의 레벨을 제공할 수도 있다. 전자 디바이스 (102) 는 또한 검출 및 신뢰도 값 (256) 이 검출 및 추적 임계치 (254) 보다 더 큰지 여부를 결정할 수도 있다 (606). 검출 및 신뢰도 값 (256) 이 검출 및 추적 임계치 (254) 보다 더 큰 경우, 전자 디바이스 (102) 는 다음 비디오 프레임에서의 윈도우들 및 사이즈들의 서브세트 (예를 들어, 동일한 서브세트) 를 이용하여 다음 비디오 프레임 상에서 오브젝트 검출을 수행할 수도 있다 (608). 대안적으로, 검출 및 신뢰도 값 (256) 이 검출 및 추적 임계치 (254) 보다 더 작은 경우, 전자 디바이스 (102) 는 다음 비디오 프레임에서의 윈도우 로케이션들 및 사이즈들의 보다 큰 서브세트를 이용하여 다음 비디오 프레임 상에서 오브젝트 검출을 수행할 수도 있다 (610). 일부 구성들에서, 신뢰도 값 (256) 이 검출 및 추적 임계치 (254) 보다 더 작은 경우, 전자 디바이스 (102) 는 다음 비디오 프레임의 모든 윈도우들 및/또는 전체 검색 공간을 이용하여 다음 비디오 프레임 상에서 오브젝트 검출을 수행할 수도 있다 (610).

[0060] 도 7 을 참조하면, 상이한 윈도우 사이즈들 (766) 을 갖는 이미지 윈도우 (700) 의 특정 실시형태가 도시된다. 구체적으로, 도 7 은 10 개의 가능한 윈도우 사이즈들 (766a 내지 766j) 의 세트를 예시한다. 각각의 윈도우 사이즈 (766) 는 스케일 레벨 (예를 들어, 1 내지 10) 에 대응할 수도 있다. 본 명세서에서 직사각형으로서 도시되었지만, 검색되는 윈도우들은 임의의 형상, 예를 들어, 정사각형, 직사각형, 원형, 타원형, 사용자-정의된 것 등일 수도 있다. 더욱이, 임의의 개수의 윈도우 사이즈들 (766) 또는 스케일 레벨들, 예를 들어, 5, 15, 20, 30 등이 이용가능할 수도 있다.

[0061] 검색 범위는 특정 로케이션에 대해 이용된 윈도우 사이즈들의 서브세트에 의해 표시될 수도 있다, 예를 들어, 현재 비디오 프레임 (N) (224) 에서 검색되는 윈도우 사이즈들은 최근 프레임에서의 타깃 오브젝트와 연관된 윈도우 로케이션 및 윈도우 사이즈와 유사한 것들로 제한될 수도 있다. 예를 들어, 피드백 없이, 오브젝트 검출기 (208) 는 각각의 선택된 윈도우 로케이션에 대해 10 개의 윈도우 사이즈들 (766a 내지 766j) 모두를 검색할 수도 있다. 그러나, 오브젝트가 제 5 윈도우 사이즈 (766e) 를 갖는 윈도우를 이용하여 최근 (현재가 아님) 비디오 프레임에서 검출된 경우, 스캐너 스케일러 (236) 는 5 플러스 또는 마이너스 3, 즉, 윈도우 사이즈들 2 내지 8 의 현재 캡처된 프레임에 대한 윈도우 사이즈들만을 단지 선택할 수도 있다. 다시 말해, 제 1 윈도우 사이즈 (766a), 제 9 윈도우 사이즈 (766i) 및 제 10 윈도우 사이즈 (766j) 를 갖는 윈도우들은 최근 또는 이전 비디오 프레임 (N-1) (222) 으로부터의 피드백에 기초하여 검색되지 않을 수도 있다. 이것은 낮은 확률 검색을 더욱 없애고 오브젝트 검출의 효율을 증가시킬 수도 있다. 다시 말해, 최근 비디오 프레임으로부터의 피드백을 이용하면 수행된 연산들을 감소시키는 것을 도울 수도 있다. 대안적으로, 최근 비디오 프레임이 타깃 오브젝트를 검출하지 못한 경우 (즉, 최근 캡처된 프레임에 대한 검출 및 추적 신뢰도 값 (256) 이 검출 및 추적 임계치 (254) 보다 작은 경우), 오브젝트 검출기 (208) 는 사이즈 레벨들의 서브세트를 이용함으로써 검색 범위를 제한하지 않을 수도 있다.

[0062] 도 8 을 참조하면, 오브젝트 추적 및 검출 모듈 (804) 의 특정 예시적 실시형태가 도시된다. 도 8 에 예시된 오브젝트 추적 및 검출 모듈 (804) 은 도 2 에 예시된 오브젝트 추적 및 검출 모듈 (204) 과 유사한 모듈들을 포함하고 유사한 기능성을 수행할 수도 있다. 구체적으로, 도 8 에 예시된 오브젝트 검출기 (808), 모션 추적기 (806), 스캐너 로케이터 (830), 윈도우 로케이션 선택기 (832), 랜더마이저 (834), 스캐너 스케일러 (836), 분류기 (838), 퓨전 모듈 (860), 메모리 버퍼 (810), 캡처된 이전 비디오 프레임 (812), 윈도우 (842), 로케이션 (844), 사이즈 (846), 2진 판정 (848), 추적 임계치 (850), 검출 임계치 (852), 검출 및 추적 임계치 (854), 검출 신뢰도 값 (840), 추적 신뢰도 값 (828) 및 검출 및 추적 신뢰도 값 (856) 은 도 2 에 예시된 오브젝트 검출기 (208), 모션 추적기 (206), 스캐너 로케이터 (230), 윈도우 로케이션 선택기 (232), 랜더마이저 (234), 스캐너 스케일러 (236), 분류기 (238), 퓨전 모듈 (260), 메모리 버퍼 (210), 캡처된 이전 비디오 프레임 (212), 윈도우 (242), 로케이션 (244), 사이즈 (246), 2진 판정 (248), 추적 임계치 (250), 검출 임계치 (252), 검출 및 추적 임계치 (254), 검출 신뢰도 값 (240), 추적 신뢰도 값 (228) 및 검출 및 추적 신뢰도 값 (256) 에 대응하고 이들과 유사한 기능성을 가질 수도 있다.

[0063] 오브젝트 추적 및 검출 모듈 (804) 은 타깃 모션 및 추적 에러로 인한 지터링 영향 (jittering effect) 을 감소시키는데 이용되는 평활화 모듈 (smoothing module; 861) 을 포함할 수도 있다. 다시 말해, 평활화 모듈 (861) 은 추적 결과들을 평활화하여, 검색 윈도우가 로케이션 (x, y) (844) 과 사이즈 (폭, 높이) (846) 양쪽에서 보다 평활화된 궤적을 갖게 한다. 평활화 모듈 (861) 은 단순한 이동 평균 (moving average; MA) 필터들 또는 자기 회귀 (auto regression; AR) 필터들일 수 있다. 로케이션 (844) 및 사이즈 (846) 에 대한 평활화 정도는 상이할 수 있다. 칼만 (Kalman) 필터와 같은 예측 필터들이 또한 로케이션 (844) 평활화에 적합

할 수도 있다. 그에 따라, 평활화 모듈 (861) 은 입력으로서 비평활화된 로케이션 (863) 및 비평활화된 사이즈 (865) 를 수신하고 평활화된 로케이션 (867) 및 평활화된 사이즈 (869) 를 출력할 수도 있다.

[0064] 도 9 를 참조하면, 평활화 모듈 (961) 의 특정 예시적 실시형태가 도시된다. 평활화 모듈 (961) 은 타깃 모션 및 추적 에러로 인한 지터링 영향을 감소시키는데 이용될 수도 있어서, 즉, 추적 결과들 (바운딩 박스) 이 로케이션(x, y) 과 사이즈(폭, 높이) 양쪽에서 보다 평활화된 궤적을 갖도록 한다. 하나의 구성에서, 로케이션 평활화 필터 (971) 및 사이즈 평활화 필터 (973) 는 입력으로서 비평활화된 로케이션 (963) 및 비평활화된 사이즈 (965) 를 수신하고 평활화된 로케이션 (967) 및 평활화된 사이즈 (969) 를 출력하기 위한 자기 회귀 (AR) 모델을 이용하여 구현된다.

[0065] 자기 회귀 (AR) 모델에서, X 는 평활화될 변수, 로케이션 또는 사이즈 중 어느 하나인 것으로 가정한다. 더욱이, X' 는 오브젝트 추적기에 의한 X 의 출력이라고 하자. 이 구성에서, 시간 t 에서 X 의 평활화 필터링, X_t 는 식 (1) 에 따라 설명될 수 있다:

[0066]
$$X_t = W * X'_t + (1 - W) * X_{t-1} \tag{1}$$

[0067] 여기서 X'_t 는 시간 t 에서의 X 의 추적기 출력이고, X_{t-1} 은 시간 t-1 에서의 X 의 평활화된 결과이며, W (0 ≤ W ≤ 1) 는 평활화 효과를 제어하는 평활화 가중치이다. 예를 들어, X'_t 는 현재 비디오 프레임 (N) (224) 에 대해 선택된 윈도우 로케이션 또는 윈도우 사이즈일 수도 있고, X_{t-1} 은 이전 비디오 프레임 (N-1) (222) 에 대해 이용된 윈도우 로케이션 또는 윈도우 사이즈일 수도 있다.

[0068] 상이한 평활화 가중치 W 는 로케이션 평활화 필터 (971) 및 사이즈 평활화 필터 (973) 에 대해 이용될 수 있다. 예를 들어, 하나의 구현에서, W_{location} = 0.8 이고 W_{size} = 0.4 이어서, 윈도우 로케이션에 대해서는 평활화 효과가 더 작지만 윈도우 사이즈에 대해서는 평활화 효과가 더 강하다. 평활화 가중치들의 이러한 선택은 더 적은 추적 딜레이와 더 적은 지터링 양쪽을 생성할 것이다.

[0069] 평활화 가중치의 선택은 또한 검출 및 추적 신뢰도 값 (856) 이 소정 임계치 (예를 들어, 검출 및 추적 임계치 (854)) 를 하회할 때 감소될 수도 있다. 이것은 잠재적인 추적 또는 검출 에러들이 높을 때 보다 강한 필터링을 유발할 수도 있다. 예를 들어, 낮은 추적 신뢰도 (예를 들어, 검출 및 추적 신뢰도 값 (856) 이 검출 및 추적 임계치 (854) 를 하회함) 에 응답하여, 로케이션 및 사이즈에 대한 평활화 가중치들은 W_{location} = 0.65 및 W_{size} = 0.2 로 각각 설정될 수도 있다. 다시 말해, 가중치들 중 하나 또는 양쪽이 감소될 수도 있는데, 이는 윈도우 로케이션 및 사이즈 선택이 현재 비디오 프레임의 것들보다 이전 비디오 프레임들의 윈도우 로케이션들 및 사이즈들에 더 많이 의지하게 할 수도 있다.

[0070] 가중시키는 것은 검출 및 추적 신뢰도 값 (856) 보다는 오히려 추적 신뢰도 값 (828) 또는 검출 신뢰도 값 (840) 에 기초할 수도 있다. 예를 들어, 평활화 가중치들 W_{location} 및 W_{size} 는 추적 신뢰도 값 (828) 이 추적 임계치 (850) 를 하회하는 것에 응답하여 감소될 수도 있다, 즉, 열악한 모션 추적에 응답하여 보다 강한 필터링이 이용될 수도 있다. 대안적으로, 평활화 가중치들은 검출 신뢰도 값 (840) 이 검출 임계치 (852) 를 하회하는 것에 응답하여 감소될 수도 있다, 즉, 열악한 오브젝트 검출에 응답하여 보다 강한 필터링이 이용될 수도 있다.

[0071] 다른 구성에서, 칼만 필터링이 이용되어 윈도우 로케이션을 평활화할 수도 있다. 이러한 구성에서, 필터링 은 식 (2) 내지 식 (7) 에 따라 정의될 수도 있다:

[0072]
$$x_k = F_k x_{k-1} + w_k \tag{2}$$

[0072]
$$z_k = H x_{k-1} + v_k \tag{3}$$

[0073] 여기서 x_{k-1} 은 시간 k-1 에서의 이전 상태이고, x_k 는 x_k = [x, y, \dot{x} , \dot{y}] 에 의해 정의된 현재 상태이며, 여기서 (x, y) 는 바운딩 박스 중심 로케이션이고, (\dot{x} , \dot{y}) 는 각각의 방향에서의 속도이다. 더욱이, 상태 천이 모델 F_k 및 관측 모델 H 는 식 (4) 및 식 (5) 에 의해 각각 정의될 수도 있다:

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

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

[0074]

[0075] 여기서 Δt 는 튜닝가능 파라미터이다. 부가적으로, w_k 는 식 (6) 에 따라 공분산 Q 에 의한 제로 평균 다변수 정규 분포로부터 드로잉되는 것으로 가정되는 프로세스 노이즈이다 (즉, $w_k \sim N(0, Q)$):

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

[0076]

[0077] 여기서 σ_1 은 튜닝가능 파라미터이다. 이와 유사하게, v_k 는 식 (7) 에 따라 공분산 R 에 의한 제로 평균 가우스 화이트 노이즈인 것으로 가정되는 관측 노이즈이다 (즉, $v_k \sim N(0, R)$):

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

[0078]

[0079] 여기서 σ_2 은 튜닝가능 파라미터이다.

[0080]

도 10 을 참조하면, 모션 추적 결과들에서의 지터를 평활화하기 위한 방법 (1000) 의 특정 예시적 실시형태의 플로우차트가 도시된다. 방법 (1000) 은 전자 디바이스 (102), 예를 들어, 전자 디바이스 (102) 에서의 오브젝트 추적 및 검출 모듈 (804) 에 의해 수행될 수도 있다. 전자 디바이스 (102) 는 현재 비디오 프레임 (224) 과 연관된 하나 이상의 윈도우 로케이션들 및 하나 이상의 윈도우 사이즈들, 예를 들어, 비평활화된 로케이션 (863) 및 비평활화된 사이즈 (865) 를 결정할 수도 있다 (1002). 전자 디바이스 (102) 는 또한 하나 이상의 윈도우 로케이션들 및 하나 이상의 윈도우 사이즈들을 필터링하여 하나 이상의 평활화된 윈도우 로케이션들 (867) 및 하나 이상의 평활화된 윈도우 사이즈들 (869) 을 생성할 수도 있다 (1004). 예를 들어, 이것은 이동 평균 필터, 자기 회귀 필터 또는 칼만 필터를 이용하는 것을 포함할 수도 있다. 하나의 구성에서, 낮은 추적 신뢰도 (예를 들어, 검출 및 추적 신뢰도 값 (856) 이 검출 및 추적 임계치 (854) 를 하회함) 에 응답하여, 로케이션 및 사이즈에 대한 평활화 가중치들이 감소될 수도 있다. 대안적으로, 평활화 가중치들은 검출 신뢰도 값 (840) 또는 추적 신뢰도 값 (828) 에 기초하여 감소될 수도 있다. 전자 디바이스는 또한 하나 이상의 평활화된 윈도우 로케이션들 (867) 및 하나 이상의 평활화된 사이즈들 (869) 에 의해 정의된 하나 이상의 윈도우들을 이용하여 현재 비디오 프레임 (224) 내에서 타깃 오브젝트를 검출할 수도 있다 (1006).

[0081]

장면이 이미징된다는 맥락에서, 용어 "오브젝트" 는 장면 내의 물리적 오브젝트를 지칭한다. 비디오 스트림의 맥락에서, 용어 "오브젝트" 는 오브젝트의, 비디오 스트림 내의, 표현 (예를 들어, 비디오 스트림의 프레임들에서의 오브젝트의 이미지들) 을 지칭한다. 본 명세서에서 사용되는 바와 같이 용어 "모바일 디바이스" 는 다음의 품 액터들 중 임의의 것에 있어서의 디바이스들을 포함한다: 홀더블들 (holdables) (예를 들어, 스마트폰들), 드라이버블들 (drivables) (예를 들어, 운반체들 또는 로봇들), 웨어러블들 (wearables) (예를 들어, 의류 또는 액세서리들), 및 플라이어블들 (flyables) (예를 들어, 드론들). 모바일 디바이스는 하나 이상의 스크린들 (예를 들어, 터치스크린) 및/또는 하나 이상의 이미지-캡처링 디바이스들 (예를 들어, 카메라) 을 포함할 수도 있다.

[0082]

디지털 이미지 (예를 들어, 비디오 스트림의 프레임) 는 하나 이상의 오브젝트들을 포함할 수도 있는 장면을 나타낸다. 통상 오브젝트들은 물리적 및 유형의 오브젝트들 (예를 들어, 사람들) 이다. 도 12 에서, 예를 들어, 장면에서의 오브젝트들은 3 명의 사람들, 별, 및 나무를 포함한다.

[0083]

하나 이상의 디지털 이미지들 (예를 들어, 비디오 스트림의 프레임들) 로부터, 하나 이상의 이미지들 내에서 시간의 경과에 따른 선택된 오브젝트들의 로케이션들을 표현하는 정보를 갖는 연관된 스트림 또는 파일을 생성하

는 것이 바람직할 수도 있다. 하나의 예에서, 이러한 정보는 관심 포인트의 절대적 로케이션 (예를 들어, 선택된 오브젝트들의 질량 중심) 을 포함하는데, 이는 시간의 경과에 따라 변화할 수도 있다. 다른 예에서, 이러한 정보는 관심 포인트의 로케이션에 대한 선택된 오브젝트의 로케이션들 (예를 들어, 선택된 오브젝트들의 질량 중심) 을 포함하는데, 이는 시간의 경과에 따라 변화할 수도 있다. 관심 포인트의 로케이션 (예를 들어, 선택된 오브젝트의, 또는 다수의 오브젝트들의 질량 중심의 로케이션) 은 메타데이터 (예를 들어, 하나 이상의 각각의 로케이션 코드부들의, 하나 이상의 인덱스들, 이를 테면 코드워드들) 로서 인코딩될 수도 있다.

[0084] 도 11a 는 태스크들 (T100 및 T300) 을 포함하는 제 1 구성에 따라 메타데이터를 발생시키는 방법 (M100) 의 플로우차트를 도시한다. 태스크 T100 은 복수의 오브젝트들을 선택하기 위해 적어도 하나의 디지털 이미지를 이용하며, 여기서 비디오 스트림은 오브젝트들을 물리적 공간에 관하여 나타낸다. 로케이션 공간에서의 선택된 오브젝트의 포지션들을 나타내는 정보에 기초하여, 태스크 T300 은 선택된 오브젝트들의 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 식별하는 메타데이터를 생성한다. 태스크들 (T100 및 T300) 은 예를 들어, 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 또는 이들의 임의의 조합에 의해 수행될 수도 있다.

[0085] 태스크 T100 은 복수의 오브젝트들을 선택하기 위해 적어도 하나의 디지털 이미지를 이용하며, 여기서 적어도 하나의 디지털 이미지는 오브젝트들을 물리적 공간에 관하여 나타낸다. 선택은 사용자에 의한 직접적인 액션에 응답하여 수행될 수도 있다. 이러한 경우, 스크린은 사용자에게 적어도 하나의 디지털 이미지를 디스플레이하는데 이용될 수도 있고, 사용자는 디스플레이에서 보이는 오브젝트들 중에서 특정의 원하는 오브젝트들을 선택하기 위한 표시를 입력할 수도 있다. 도 11b 는 적어도 하나의 디지털 이미지를 디스플레이하는 태스크 T50 을 포함하는 방법 (M100) 의 구현 (M110) 의 플로우차트를 도시한다. 디스플레이가 터치스크린 상에서 수행되는 경우, 사용자는 원하는 오브젝트 내에서 포인트를 선택하기 위해 스크린을 터치하는 것에 의해 선택을 나타낼 수도 있다.

[0086] 적어도 하나의 디지털 이미지는 비디오 스트림의 하나 이상의 프레임들일 수도 있다. 비디오 스트림은 프레임들의 시리즈를 설명하고, 여기서 각각의 프레임은 픽셀 좌표 공간에서의 이미지를 표현한다. 비디오 스트림은 통상 스트림으로부터 프레임들을 복구하는데 이용될 수도 있는 다른 정보 (예를 들어, 각각의 프레임에 대한 대응하는 프레임 시작 코드 또는 패킷 및 프레임 종단 코드 또는 패킷) 를 포함한다. 비디오 스트림은 또한, 특정 프레임과 연관될 수도 있는 임베딩된 데이터 (예를 들어, 메타데이터) 를 포함할 수도 있다. 비디오 스트림은 (가시 및/또는 다른 파장들에 민감할 수도 있는) 카메라 또는 다른 이미징 디바이스에 의해 생성되거나, 다른 디바이스로부터 스트리밍되거나, 또는 (예를 들어, 자기 또는 광학 매체 상에 저장된 정보로부터) 디코더에 의해 생성될 수도 있고, 압축된 또는 비압축된 형태로 있을 수도 있다. 비디오 스트림은 구조화된 광 이미지 또는 다른 심도 카메라 (예를 들어, 마이크로소프트 키넥트 (Microsoft Kinect)) 에 의해 캡처된 이미지들에 기초하여 비디오 스트림과 같은 심도 정보를 포함할 수도 있다. 이러한 비디오 스트림은, 예를 들어, 각각의 픽셀의 심도 값을 대응하는 컬러에 맵핑시키는 것에 의해 터치스크린 상에 디스플레이될 수도 있다. 비디오 스트림은 라이브 (live) 이거나, 딜레이되거나, 또는 (예를 들어, 미리 레코딩된) 스토리지로부터 추출될 수도 있다.

[0087] 비디오 스트림은 디지털 이미지들 ("프레임들") 의 시간 시퀀스이다. 하나의 예에서, 이미지들의 시퀀스는 특정 프레임 레이트에서 (예를 들어, 동일한 프레임 레이트에서의 제시를 위해) 캡처된다. 비디오 스트림은 각각의 프레임의 시작과 종단을 나타내는 정보를 포함할 수도 있다. 하나의 예에서, 비디오 스트림은 각각의 프레임의 제 1 픽셀 전의 프레임 시작 코드 및 각각의 프레임의 마지막 픽셀 후의 프레임 종단 코드를 포함한다. 다른 예에서, 비디오 스트림은 비디오 스트림의 프레임들의 픽셀 정보를 반송하는 하나 이상의 데이터 신호들, 및 비디오 스트림의 각각의 프레임의 시작과 종단을 나타내는 하나 이상의 동기화 신호들을 포함한다. 1 차 비디오 스트림의 특정 포맷에 따라, 1 차 비디오 스트림은 부가적인 정보 (예를 들어, 헤더들, 메타데이터) 를 포함할 수도 있다. 1 차 비디오 스트림이 프레임 시작과 종단 코드들 사이에 프레임 데이터 이외의 데이터 (즉, 픽셀 값들) 를 포함하는 것도 또한 가능하다. 이러한 다른 데이터는, 예를 들어, 센서 구성 상세들 및/또는 이미지 통계 값들을 포함할 수도 있다.

[0088] 비디오 스트림의 파싱은 (예를 들어, 프레임 시작과 종단 정보를 이용하여) 비디오 스트림의 프레임들을 추출하는 것 및 프로세싱을 위해 이들을 이용가능하게 만드는 것을 포함한다. 예를 들어, 파싱은 각각의 프레임을, 버퍼와 같은, 대응하는 특정한 로케이션에 저장하는 것을 포함할 수도 있다. 도 11c 는 적어도 하나의 디지털 이미지를 생성하기 위해 비디오 스트림을 파싱하는 태스크 T30 을 포함하는 방법 (M100) 의 구현

(M120)의 플로우차트를 도시한다. 대안적으로, 비디오 스트림의 하나 이상의 프레임들이 임의의 시간에 (예를 들어, 메모리에서의 버퍼에서) 방법 (M100)에 이용가능하게 되도록, 다른 프로세스가 (예를 들어, 디스플레이를 위해) 비디오 스트림을 파싱하도록 실행할 수도 있다. 도 11d는 방법들 (M110 및 M120)의 구현 (M130)의 플로우차트를 도시한다.

[0089] 이러한 버퍼 (예를 들어, 메모리 버퍼 (110))는 통상적으로 한 번에 적어도 하나의 풀 (full) 프레임을 저장하는 것이 가능하다. 1 차 비디오 스트림의 파싱은 다수의 프레임 버퍼들 간에서 교호하도록 구성될 수도 있다. 예를 들어, 하나의 이러한 프레임 버퍼는 프로세싱을 위해 풀 프레임을 저장할 수도 있는 한편, 시퀀스에서 후속하는 프레임은 추출되어 다른 프레임 버퍼에 저장된다. 대안적으로, 태스크 T30 또는 파싱 프로세스는 파싱된 프레임들을 원형 버퍼에 저장하도록 구성될 수도 있어서, 새로운 프레임들이 추출됨에 따라 버퍼가 업데이트된다 (그리고 구 (old) 프레임들이 오버라이팅된다).

[0090] 1 차 비디오 스트림의 파싱은 1 차 비디오 스트림을 하나 이상의 연관된 오디오 스트림들 및/또는 메타데이터 스트림들과 분리시키기 위한 디멀티플렉싱 동작을 포함할 수도 있다. 대안적으로, 이러한 동작은 파싱 프로세스 또는 태스크 T30의 상류에서 수행될 수도 있다 (예를 들어, 1 차 비디오 스트림은 디멀티플렉서에 의해 제공될 수도 있다).

[0091] 도 12는 3개의 오브젝트들을 선택하기 위한 사용자 액션들의 시퀀스의 예를 도시한다. 패널 A는 디바이스의 터치스크린 상에서 디스플레이된 장면을 도시한다. 디바이스는 선택 모드를 개시하기 위해 사용자에게 의해 조작될 수도 있다. 예를 들어, 터치스크린은 사용자가 적절한 아이콘을 터치할 때 선택 모드에 진입하도록 구성될 수도 있다. 패널 A는 프레임의 상부 우측 코너에서의 선택 툴 아이콘의 예를 도시한다. 패널 B에서, 사용자는 선택 모드를 개시하기 위해 선택 툴 아이콘을 터치한다. 패널 C에서, 툴 아이콘은 터치 액션에 응답하여 하이라이팅되어 디바이스가 선택 모드에 있음을 나타내고, 사용자는 제 1 오브젝트를 터치하여 그것을 선택한다. 패널 D에서, 제 1 오브젝트는 터치 액션에 응답하여 하이라이팅되어 (예를 들어, 도시된 바와 같이 실루엣팅되거나 (silhouetted), 또는 아웃라이닝되어 (outlined)) 오브젝트가 선택됨을 나타내고, 사용자는 제 2 오브젝트를 터치하여 그것을 선택한다. 패널 E에서, 제 2 오브젝트는 터치 액션에 응답하여 하이라이팅되어 오브젝트가 선택됨을 나타내고, 사용자는 제 3 오브젝트를 터치하여 그것을 선택한다. 패널 F에서, 제 3 오브젝트는 터치 액션에 응답하여 하이라이팅되어 오브젝트가 선택됨을 나타내고, 사용자는 선택 툴 아이콘을 다시 터치하여 선택 모드를 종료한다. 디바이스는 사용자가 선택 모드에서 선택된 오브젝트를 다시 터치하는 경우 그 선택된 오브젝트를 선택해제하도록 구성될 수도 있다.

[0092] 도 13은 3개의 오브젝트들을 선택하기 위한 사용자 액션들의 시퀀스의 다른 예를 도시한다. 이 경우, 디바이스는 도 12에 도시된 것과는 상이하게 선택 모드에서 동작한다. 패널 A에서, 사용자는 선택 모드를 개시하기 위해 선택 툴 아이콘을 터치한다. 패널 B에서, 툴 아이콘은 터치 액션에 응답하여 하이라이팅되어 디바이스가 선택 모드에 있음을 나타내고, 사용자는 제 1 오브젝트를 터치하여 그것을 선택한다. 패널 C에서, 제 1 오브젝트는 터치 액션에 응답하여 하이라이팅되어 오브젝트가 선택됨을 나타내고, 그 선택에 응답하여, 선택 모드가 종료되고 하이라이팅이 툴 아이콘으로부터 제거된다. 사용자가 선택 툴 아이콘을 다시 터치하여 선택 모드를 개시하고, 패널 D에서, 툴 아이콘은 터치 액션에 응답하여 하이라이팅되어 디바이스가 선택 모드에 있음을 나타내고, 사용자는 제 2 오브젝트를 터치하여 그것을 선택한다. 패널 E에서, 제 2 오브젝트는 터치 액션에 응답하여 하이라이팅되어 오브젝트가 선택됨을 나타내고, 그 선택에 응답하여, 선택 모드가 종료되고 하이라이팅이 툴 아이콘으로부터 제거된다. 사용자가 선택 툴 아이콘을 다시 터치하여 선택 모드를 개시하고, 패널 F에서, 툴 아이콘은 터치 액션에 응답하여 하이라이팅되어 디바이스가 선택 모드에 있음을 나타내고, 사용자는 제 3 오브젝트를 터치하여 그것을 선택한다.

[0093] 다른 예에서, 터치스크린은 2-액션 선택에 응답하여 선택 모드에 진입하도록 구성될 수도 있다: 제 1 액션은 메뉴 모드를 선택하고 (예를 들어, 사용자는 스크린의 지정된 영역을 터치하여 메뉴를 디스플레이함) 제 2 액션은 메뉴로부터 선택 모드를 선택한다. 추가 예에서, 터치스크린은 선택 툴 아이콘을 터치 및 홀딩하고, 아이콘을 스크린 상의 원하는 포인트로 드래그하고, 그 후에 아이콘을 릴리스하여 포인트를 선택하는 것에 의해 사용자가 오브젝트를 선택하도록 구성될 수도 있다. 대안적으로, 사용자는 원하는 오브젝트의 적어도 일 부분을 포함하는 이미지의 영역 (예를 들어, 바운딩 박스 또는 타원 또는 라쏘 (lasso))을 선택하는 것에 의해 선택을 나타내도록 터치스크린을 조작할 수도 있다. 이러한 경우, 사용자는 원하는 오브젝트들을 개별적으로 및/또는 그룹으로서 선택할 수도 있다.

[0094] 사용자로 하여금 선택 모드 동안 디스플레이되는 시야를 조작하게 하는 (예를 들어, 핀치 (pinch) 액션을 수행

하여 시야를 넓히고, 줌 액션을 수행하여 시야를 좁히는) 것이 바람직할 수도 있다. 터치스크린의 조작 외에도, 직접적인 사용자 선택을 위한 다른 가능한 양상들은 하나 이상의 버튼들 또는 다른 스위치들을 작동시키는 것 및/또는 제스처 인식을 포함한다. 이러한 사용자에게 의한 직접적인 액션에 대한 대안으로서, 태스크 T100 은 적어도 하나의 디지털 이미지 (예를 들어, 비디오 스트림의 하나 이상의 프레임들)로부터의 정보에 하나 이상의 미리 결정된 기준들을 적용하는 것에 의해 선택을 자동적으로 수행할 수도 있다. 이러한 기준들은, 예를 들어, 선수의 등 번호 (jersey number) 의 인식, 얼굴 인식, 및/또는 (예를 들어, 체육 행사에서 공과 같은, 빨리 이동하는 오브젝트를 식별하기 위한) 속도 검출을 포함할 수도 있다.

[0095] 방법 (M100) 은, 오브젝트를, 태스크 T100 에서의 그의 선택에 응답하여 추적하도록 구현될 수도 있다. 예를 들어, 이하 설명되는 포지션-결정 태스크 T200 은 (예를 들어, 오브젝트 추적 및 검출 모듈 (204) 및/또는 오브젝트 추적 및 검출 모듈 (804) 에 대하여 상술한 바와 같이) 비디오 스트림 내에서 하나 이상의 선택된 오브젝트들을 또한 추적하도록 구현될 수도 있다. 대안적으로, 오브젝트 추적은 태스크 T100 에 앞서 시작될 수도 있다. 이러한 경우, 태스크 T100 은 추적된 오브젝트들의 세트 중에서의 선택을 위해 구현될 수도 있다. 예를 들어, 사용자는 추적된 오브젝트들의 디스플레이 중에서 선택할 수도 있다. (예를 들어, 태스크 T200 에 의한) 이러한 오브젝트 추적을 위해 이용될 수도 있는 알고리즘들의 예들로는 루카스-카나데 방법 (그리고 카나데-루카스-토마시 (Kanade-Lucas-Tomasi) 추적기와 같은 변형들) 및 평균-시프트 (mean-shift) 추적을 포함한다. 오브젝트의 추적은 1 차 비디오 스트림의 프레임의 잔여분 (rest) 으로부터 오브젝트를 세그먼트화하는 것을 포함할 수도 있지만, 오브젝트 추적은 또한 이러한 세그먼트화의 부존재시에 수행될 수도 있다.

[0096] 디지털 이미지 (예를 들어, 비디오 스트림의 프레임) 의 잔여분으로부터의 오브젝트의 세그먼트화는 프레임의 잔여분에 대한 오브젝트의 향상된 디스플레이 (예를 들어, 실루엣팅 또는 아웃라이닝) 를 지원하기 위해 이용될 수도 있다. 태스크 T100 에서의 오브젝트의 선택에 응답하여 (예를 들어, 디스플레이 태스크 T50 에 의해) 세그먼트화가 수행될 수도 있다. 이러한 세그먼트화는 오브젝트의 선택을 확인함으로써 사용자에게 시각적 피드백을 제공하기 위해 이용될 수도 있다. 세그먼트화는 또한 선택이 시작되기 전에 수행될 수도 있다. 이러한 경우, 태스크 T100 에서의 선택은 이미 세그먼트화된 오브젝트들의 세트 (예를 들어, 세그먼트화된 오브젝트들이 향상된 디스플레이) 중에서 행해질 수도 있다.

[0097] 도 14 의 좌측 패널들은 오브젝트들을 포함하는 장면들을 나타내는 비디오 스트림들로부터의 프레임들의 2 개의 예들을 도시한다. 오브젝트의 향상된 디스플레이는 태스크 T100 에서의 선택 전에 또는 이러한 선택에 응답하여 (예를 들어, 디스플레이 태스크 T50 에 의해) 수행될 수도 있다. 향상된 디스플레이의 하나의 예는, 배경과 대비되는 컬러 (예를 들어, 도 14 의 중심 패널들에 도시된 바와 같이, 블랙 또는 옐로우) 로 디스플레이될 수도 있는, 오브젝트의 실루엣이다. 이 대비되는 컬러는 모든 오브젝트들에 대해 동일할 수도 있고 또는 (예를 들어, 각각의 오브젝트의 로컬 배경과 대비되기 위해) 오브젝트들 간에 상이할 수도 있다. (대비되는 컬러로 마찬가지로 행해질 수도 있는) 향상된 디스플레이의 다른 예들로는 오브젝트의 아웃라이닝 그리고 오브젝트에, 그 내부에, 또는 그 주위에 그래픽 엘리먼트 (예를 들어, 오브젝트의 정점 또는 중심에서의 삼각형, 도 14 의 우측 패널들에 도시된 오브젝트 주위의 타원 등) 를 디스플레이하는 것을 포함한다.

[0098] 사용자에게 의한 오브젝트들의 직접적인 선택을 위한 맥락으로서 비디오 스트림의 디스플레이를 이용하는 것은 느리게 이동하는 오브젝트들에 대해 적절할 수도 있다. 그러나, 원하는 오브젝트들 중 임의의 오브젝트가 디스플레이 내에서 빨리 이동하고 있는 경우, 이러한 배열은 사용자에게 불만스러울 수도 있다. 이러한 빠른 이동의 경우들에서는, 사용자에게 의한 오브젝트들의 직접적인 선택을 위한 맥락으로서 비디오 스트림의 단일 프레임을 그 대신에 디스플레이하는 것이 바람직할 수도 있다. 상술된 바와 같이, 프레임에서의 오브젝트들이 세그먼트화되고 향상되는 디스플레이로부터 선택이 행해질 수도 있다. 스트림-디스플레이 선택과 프레임-디스플레이 선택 양쪽이 지원되는 경우, 이들 디스플레이 모드들 중에서의 선택이 사용자에게 의해 및/또는 자동적으로 수행될 수도 있다. 자동적인 선택의 경우, 예를 들어, 광학 플로우, 모션 벡터들, 프레임들 간의 픽셀 차이들 등과 같은 척도를 이용하여 스트림에서의 모션의 정도가 정량화될 수도 있다. 이러한 경우, 프레임-디스플레이 모드는 척도의 값이 임계치를 상회할 때 선택될 수도 있고, 스트림-디스플레이 모드는 척도의 값이 임계치를 하회할 때 선택될 수도 있다.

[0099] 로케이션 공간에서의 선택된 오브젝트들의 포지션들을 나타내는 정보에 기초하여, 태스크 T300 은 선택된 오브젝트들의 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 식별하는 메타데이터를 생성한다. 로케이션 공간은 디지털 이미지 (예를 들어, 비디오 스트림의 프레임) 의 (x,y) 데카르트 좌표 공간과 같은, (예를 들어, 비디오 스트림의) 적어도 하나의 디지털 이미지의 픽셀 좌표 공간일 수도 있다. 디지털 이미지

(예를 들어, 비디오 프레임)의 픽셀 좌표 공간의 원점은 통상 이미지의 상부 좌측 또는 하부 좌측 코너로서 지정된다. 선택된 오브젝트의 포지션은 예를 들어, 사용자가 스크린을 터치하여 오브젝트를 선택한 로케이션일 수도 있다. 대안적으로, 선택된 오브젝트들의 포지션들은 본 명세서에서 설명한 바와 같이 오브젝트 추적 및 검출 모듈 (204 또는 804)과 같은 오브젝트 추적기에 의해 제공될 수도 있다.

[0100] 기하학적 배열은 선택된 오브젝트들의 공간에서의 배열이다. 예를 들어, 기하학적 배열은 (예를 들어, 다각형의 정점들로서) 선택된 오브젝트들의 포지션들에 의해 설명되는 형상일 수도 있다. 태스크 T300은 오브젝트 포지션들에 하나 이상의 메트릭들을 적용하는 것에 의해 후보 기하학적 배열들 중 적절한 하나의 후보 기하학적 배열을 식별하도록 구현될 수도 있다. 이러한 경우에, 태스크 T300은 메트릭 값들에 따라 후보를 선택하기 위해 판정 트리를 이용할 수도 있다. 생성된 메타데이터는 또한 다음 중에서 임의의 것과 같은 부가적인 정보를 포함할 수도 있다: 후보에 적용될 스케일 팩터 및/또는 애스펙트 팩터, 후보에 대한 배향 방향, 및 후보가 축 둘레로 플립핑될지 여부를 나타내는 미리 플래그.

[0101] 태스크 T300은 복수의 후보 기하학적 배열들 중에서 어느 것이 선택된 오브젝트들의 포지션들에 의해 설명된 형상에 가장 유사한지를 결정하도록 구현될 수도 있다. 도 35는 태스크 T300이 3개의 포지션들 (A, B, C)에 의해 설명된 형상을 "등거리", "근거리-근거리-원거리 (near-near-far)" 및 "근거리-원거리-원거리 (near-far-far)"로 라벨링된 3개의 후보 기하학적 배열들의 세트의 하나에 매칭시키도록 구현될 수도 있는 하나의 예를 도시한다.

[0102] 이 예에서, 포인트들 (즉, AB, BC, 및 AC) 사이의 3개의 거리들은 그들 중에서의 최소, 중간, 및 최대를 결정하기 위해 크기에 의해 계산 및 소팅된다. 최소 거리는 최대 거리의 2/3인 값과 비교된다. 최소 거리가 더 크면, 태스크 T300은 도 35의 상부 좌측에 박스로 나타낸 바와 같이, 후보 "등거리"를 선택한다. 최소 거리가 더 작으면, 중간 거리는 최소와 최대 거리들의 합의 절반인 값과 비교된다. 중간 거리가 더 크면, 태스크 T300은 도 35의 하부에 박스로 나타낸 바와 같이 후보 "근거리-원거리-원거리"를 선택한다. 중간 거리가 더 작으면, 태스크 T300은 도 35의 상부 우측에 박스로 나타낸 바와 같이, 후보 "근거리-근거리-원거리"를 선택한다.

[0103] 도 35는 원한다면 인접한 후보들 중 어느 하나에 태스크 T300에 의해 할당될 수도 있는, 판정 경계들에 있는 3개의 경우들을 도시한다. 이 경우에, 적절한 후보를 식별하는 메타데이터는 3개의 가능한 값들 중 하나를 갖는 파라미터일 수도 있으며, 이들 각각은 3개의 후보 기하학적 배열들 중 대응하는 하나의 후보 기하학적 배열을 나타낸다. 메타데이터는 또한, (예를 들어, 회전으로서) 후보에 적용될 배향을 나타낼 수도 있다. 하나의 예에서, 배향은 최소 거리에 반대인 포인트를 통하여 기하학적 배열의 중심으로부터 라인의 방향으로 나타내진다. 다른 예에서, 배향은 선택된 오브젝트들의 특정 오브젝트에 대응하는 포인트를 통하여 기하학적 배열의 중심으로부터 라인의 방향으로 나타내진다.

[0104] 도 37a는 로케이션 공간에서 선택된 오브젝트들의 포지션들을 나타내는 정보를 획득하는 태스크 T200을 포함하는 방법 (M200)의 구현 (M200)의 플로우차트를 도시한다. 태스크 T200은 태스크 T100에서의 선택에 응답하여 실행하도록 구현될 수도 있다. 로케이션 공간은 디지털 이미지 (예를 들어, 비디오 스트림의 프레임)의 (x,y) 데카르트 좌표 공간과 같은, (예를 들어, 비디오 스트림의) 적어도 하나의 디지털 이미지의 픽셀 좌표 공간일 수도 있다. 디지털 이미지 (예를 들어, 비디오 프레임)의 픽셀 좌표 공간의 원점은 통상 이미지의 상부 좌측 또는 하부 좌측 코너로서 지정된다. 도 37b 내지 도 37d는 태스크 T200을 포함하는 방법들 (M110, M120, 및 M130) 각각의 구현들 (M210, M220, 및 M230)의 플로우차트들을 도시한다.

[0105] 디스플레이의 픽셀 좌표 공간은 로케이션 공간과 동일할 수도 있고, 또는 로케이션 공간은 디스플레이의 픽셀 좌표 공간을 포함할 수도 있다. 예를 들어, 디스플레이의 프레임은 (예를 들어, 시야를 좁히는 줌 액션에 응답하여) 비디오 스트림의 대응하는 프레임의 일 부분만을 포함할 수도 있다. 대안적으로, 디스플레이의 픽셀 좌표 공간은 로케이션 공간을 포함할 수도 있다. 예를 들어, 비디오 스트림은 디스플레이의 프레임보다 더 작은 디스플레이의 윈도우에 디스플레이될 수도 있으며, 여기서 다른 콘텐츠가 디스플레이의 다른 윈도우에 디스플레이된다.

[0106] 태스크 T200은 선택된 오브젝트의 포지션을 픽셀 좌표 공간에서의 선택된 오브젝트의 이미지의 포지션으로서 결정하도록 구현될 수도 있다. 예를 들어, 결정된 포지션은 오브젝트의 질량 중심일 수도 있다. 질량 중심은 일정한 밀도가 가정되면 중심과 동일하지만, 밀도의 차이들은 또한 예를 들어 컬러 및/또는 투명도의 차이들로 나타내질 수도 있다. 결정된 포지션의 다른 예들은 오브젝트의 상부 또는 하부 픽셀, 오브젝트의 질량 중심의 오브젝트가 상부 또는 하부에의 수직 프로젝션, 또는 오브젝트의 특정 피처의 포지션을 포함한다.

도 17a 는 대응하는 결정된 포지션 (DP10) 을 획득하기 위해 오브젝트의 하부 바운딩 라인으로 오브젝트의 질량 중심 (C10) 을 프로젝팅하는 예를 도시한다.

[0107] 대안적으로 또는 부가적으로, 태스크 T200 은 다수의 오브젝트들의 형성의 레퍼런스 포지션을 결정하도록 구현될 수도 있다. 예를 들어, 태스크 T200 은 선택된 오브젝트들의 형성의 포지션으로서 레퍼런스 포지션을 계산하도록 구현될 수도 있다. 이러한 경우에, 결정된 레퍼런스 포지션은 형성의 질량 중심 (여기서 상이한 오브젝트들은 동일하거나 상이한 밀도들을 갖는다), 오브젝트들 간의 상부 또는 하부 픽셀, 그 형성의 질량 중심의 그 형성의 상부 또는 하부에의 수직 프로젝션, 또는 오브젝트들 중 특정 하나의 오브젝트의 포지션일 수도 있다. 이러한 경우에, 태스크 T200 은 (예를 들어, 레퍼런스 포지션을 원점으로서 이용하기 위해) 레퍼런스 포지션에 대한 포지션들로서 선택된 오브젝트들의 포지션들을 결정할 수도 있다.

[0108] 선택된 오브젝트들은 그들이 공통된 오브젝트의 부분들이 아니도록 서로 물리적으로 분리될 수도 있다. 예를 들어, 오브젝트들 중 적어도 하나는 다른 오브젝트들에 독립적으로 이동하는 것이 가능할 수도 있다. 특정 예에서, 오브젝트들 중 하나 이상은 상이한 사람들, 동물들, 및/또는 운반체들이다.

[0109] 비디오 스트림 또는 디스플레이의 픽셀 좌표 공간은 2 차원 또는 3 차원일 수도 있다. 3 차원 디스플레이들의 예들은 스테레오스코픽 (예를 들어, 좌측/우측) 및 볼륨메트릭 (volumetric) 디스플레이들을 포함한다. 이러한 로케이션 공간에서의 오브젝트의 사용자 선택은 예를 들어, 글로브, 제스처 (예를 들어, 배향-감응 디바이스를 킬링하는 것), 또는 압력-감응 터치스크린을 이용하여 3 개의 차원들에서 선택 포인트를 조작하는 것을 포함할 수도 있다.

[0110] 대안적으로, 로케이션 공간은 물리적 공간 (예를 들어, 장면 공간) 일 수도 있다. 하나의 이러한 예에서, 결정된 레퍼런스 포지션은 구조화된 광 이미지 또는 다른 심도 카메라 (예를 들어, 마이크로소프트 키넥트) 로부터의 비디오 스트림과 같은, 심도 정보를 포함하는 비디오 스트림을 이용하여 선택된 적어도 하나의 포인트에 기초한다. 이러한 비디오 스트림은 예를 들어, 각각의 픽셀의 심도 값을 대응하는 컬러에 맵핑시키는 것에 의해 터치스크린 상에 디스플레이될 수도 있다. 결정된 레퍼런스 포지션은 물리적 공간에서의 각각의 선택된 오브젝트의 이미지의 포지션에 기초할 수도 있고, 또는 상기 예 (예를 들어, 질량 중심) 에서처럼, 선택된 오브젝트들의 형성의 포지션일 수도 있다. 이러한 로케이션 공간에서의 오브젝트의 사용자 선택은 예를 들어, 글로브, 제스처 (예를 들어, 배향-감응 디바이스를 킬링하는 것), 또는 압력-감응 터치스크린을 이용하여 3 개의 차원들에서 선택 포인트를 조작하는 것을 포함할 수도 있다.

[0111] 도 15a 내지 도 15c 는 디스플레이된 장면의 심도 차원에서 선택 포인트를 조작하기 위해 배향-감응 디바이스를 이용하는 예를 도시한다. 이러한 배향-감응 디바이스는 예를 들어, 중력 축 (예를 들어, 지구의 중심과 디바이스를 통과한 축) 에 대한 디바이스의 배향을 나타내는 하나 이상의 배향 센서들을 갖는 모바일 디바이스일 수도 있다. 이러한 하나 이상의 배향 센서들은 예를 들어, 하나 이상의 관성 센서들 (예를 들어, 자이로스코프들 및/또는 가속도계들), 및/또는 하나 이상의 자계 센서들 (예를 들어, 자력계들) 을 포함할 수도 있다.

[0112] 도 15a 내지 도 15c 각각에서, 좌측 패널은 배향-감응 디바이스의 스크린 상의 동일한 디지털 이미지의 디스플레이를 나타내고, 우측 패널은 중력 축 (g) 에 대한 디바이스의 대응하는 배향 (디스플레이의 수직 라인을 통과하여 단면으로 도시됨, 이는 굽게 나타내짐) 을 나타낸다. 이들 도면들 각각에 디스플레이된 바와 같은 장면은 전경의 (즉, 캡처링 디바이스에 가장 가까운) 오브젝트, 중간 그라운드 오브젝트, 및 배경의 (즉, 캡처링 디바이스로부터 가장 먼) 오브젝트를 포함한다. 이들 예들 각각에서, 톨트의 정도는 디스플레이의 좌측에 슬라이더로 나타내진다. 디바이스가 도 15a 에 도시한 바와 같이 그라운드에 수직일 때, 톨트의 정도는 낮고 전경 오브젝트가 (좌측 패널에 하이라이팅으로 나타낸 바와 같이) 선택된다. 디바이스가 도 15c 에 도시한 바와 같이 그라운드에 수평인 경우, 톨트의 정도는 높고 배경 오브젝트가 (좌측 패널에 하이라이팅으로 나타낸 바와 같이) 선택된다. 도 15b 에 도시한 바와 같이, 디바이스가 이들 배향들 사이에서 그라운드에 대해 45 도 각도로 있는 경우, 톨트의 정도는 중간에 있고, 중간 그라운드의 오브젝트가 (좌측 패널에 하이라이팅으로 나타낸 바와 같이) 선택된다.

[0113] 다른 예에서, 물리적 공간에서의 오브젝트의 포지션은 오브젝트 내 또는 상의 하나 이상의 포지션 센서들 (예를 들어, GPS 센서들) 로부터의 정보에 의해 나타내질 수도 있다. 예를 들어, 오브젝트는 하나 이상의 이러한 센서들을 포함할 수도 있고 (예를 들어, 공) 또는 하나 이상의 이러한 센서들을 착용하고 있을 수도 있다 (예를 들어, 사람 또는 동물). 디스플레이의 픽셀 좌표 공간에서의 사용자-선택된 포지션의 물리적 공간에서의 오브젝트의 센싱된 포지션에의 링킹은 오브젝트 피쳐들의 인식 (예를 들어, 컬러, 등 번호) 및/또는 터치스크린 디바이스에서의 포지션 및/또는 배향 센서들을 이용하여 수행될 수도 있다.

- [0114] 부가적으로 또는 대안적으로, 물리적 공간에서의 오브젝트의 포지션은 픽셀 좌표 공간에서의 오브젝트의 이미지의 포지션으로부터 프로젝팅될 수도 있다. 디스플레이 평면에서의 오브젝트 이미지들의 포지션들 사이의 공간 관계는 물리적 공간에서의 대응하는 오브젝트들의 포지션들 사이의 공간 관계에 제대로 상관되지 않을 수도 있다. 이러한 불일치는 디스플레이 평면 (예를 들어, 비디오 스트림의 캡처 동안 초점 평면) 이 그라운드 평면에 수직인 경우 가장 높을 수도 있다.
- [0115] 도 16 은 장면 공간에서의 실제 오브젝트들 간의 거리들과 비교하여 2-D 픽셀 좌표 공간에서의 오브젝트들의 이미지들 간의 거리들 사이의 불일치의 예를 도시한다. 상부 패널 A 는 디지털 이미지 (예를 들어, 비디오 스트림의 프레임) 를 도시하고, 하부 패널 A 는 나타낸 바와 같은 장면의 좌측에서의 3 명의 사람들의 중심들 사이의 거리들을 도시한다. 이들 거리들은 가장 왼쪽의 사람이 나머지 2 명의 사람들이 서로로부터 있는 것보다 그 나머지 2 명의 사람들 각각으로부터 더 멀리 있다는 것을 나타낸다. 상부 패널 B 는 동일한 장면을 바로 위에서 본 것으로서 도시하고, 하부 패널 B 는 중간 사람이 나머지 2 명이 서로로부터 있는 것보다 실제로는 그 나머지 2 명으로부터 훨씬 더 멀리 떨어져 있다는 것을 도시하여서, 하부 패널 A 에 나타낸 거리 관계가 장면 공간에서의 실제 거리 관계에 대하여 부정확하다.
- [0116] 물리적 공간은 장면 공간에서의 그라운드 평면과 같은 2 차원 공간일 수도 있다. 하나의 이러한 예에서, 로케이션 공간은 장면 공간에서의 그라운드 평면이고, 픽셀 좌표 공간에서의 포인트는 그라운드 평면에서의 대응하는 포인트에 프로젝팅된다. 그라운드 평면은 예를 들어, 도 16 의 패널 C 에 도시한 바와 같은 농구 코트와 같은 육상 경기장일 수도 있다. 상이한 시야들을 갖고 및/또는 상이한 뷰포인트들로부터 캡처되는 동일하거나 유사한 공간에서의 유사한 이벤트들의 비디오 스트림들 간에 오브젝트 포지션들의 일관성을 지원하기 위해 이러한 로케이션 공간을 이용하는 것이 바람직할 수도 있다.
- [0117] 디지털 이미지의 픽셀 좌표 공간에서의 포인트를 그라운드 평면에서의 대응하는 포인트에 연관시키는 것은 (예를 들어, 하나 이상의 센서들 및/또는 심도 카메라로부터의 데이터에 의해 나타낸 바와 같이) 장면 공간에서의 링크된 포인트를 그라운드 평면으로 하향으로 프로젝팅하는 것 (또는 적절하다면, 링크된 포인트를 그라운드 평면으로 상향으로 프로젝팅하는 것) 에 의해 수행될 수도 있다. 도 17b 는 링크된 포인트 (LP10) (이는 상기 논의한 바와 같이 오브젝트 상의 GPS 센서 및/또는 다른 포지션 센서에 의해 나타내질 수도 있다) 의 그라운드 평면에서의 결정된 포지션 (DP20) 에의 이러한 프로젝션의 예를 도시한다.
- [0118] 대안적으로, 디지털 이미지의 픽셀 좌표 공간에서의 포인트를 그라운드 평면에서의 대응하는 포인트에 연관시키는 것은 (그라운드 평면에 직교하는 평면에서 캡처링 디바이스의 초점축을 통과하는 단면들을 도시하는) 도 18a 및 도 18b 에 도시한 바와 같이 픽셀 좌표 공간으로부터 그라운드 평면으로 포인트를 프로젝팅하는 것에 의해 수행될 수도 있다. 이러한 프로젝션은 그라운드 평면에 대한 캡처링 디바이스의 배향각 (a) 이 (예를 들어, 상술한 바와 같이 하나 이상의 배향 센서들로부터) 알려져 있다면 바로 수행될 수도 있다. 대안적으로, 배향각 (a) 은 비디오 프레임에서의 기점 정보, 이를 테면 그라운드 평면 기점들 (예를 들어, 육상 경기장의 경계 라인들), 수평선 (예를 들어, 광의 차이로 나타내짐), 오브젝트들을 바운딩하는 수평 피쳐들 (예를 들어, 공연 무대의 전방 및 후방을 나타내는 라인들), 또는 적어도 수평 피쳐를 바운딩하는 후방으로부터 추론될 수도 있다. 도 18b 에 도시한 바와 같이, 그라운드 평면에서의 거리들 (r 및 s) 사이의 비율은 배향각 (a) 에 의해, 초점 평면에서의 대응하는 거리들 (p 및 q) 사이의 비율에 관련된다. 시야각 (b) 이 알려져 있고, 초점 평면에서의 거리들 (p 및 q) 사이의 비율이 디스플레이 공간에 이미징된 바와 같이 (즉, 디스플레이의 픽셀 좌표 공간에서의 수직 라인을 따라, 서로에 대해 그 거리들이 보여지는 바와 같이) 대응하는 거리들과 동일한 것으로 가정될 수도 있다.
- [0119] 하나의 이러한 예에서, 태스크 T200 은 (예를 들어, 상술한 바와 같이) 디스플레이 평면을 그라운드 평면에 맵핑시키기 위해 이미지 기점 정보 및/또는 캡처링 디바이스의 배향을 이용하도록 구현된다. 이 구현에서, 태스크 T200 은 (예를 들어, 도 17a 에 도시한 바와 같이) 디스플레이 평면에서의 하향의 오브젝트의 질량 중심을 오브젝트의 하부에 프로젝팅하는 것에 의해 디스플레이 평면에서 각각의 선택된 오브젝트의 포지션을 결정하고, 이 디스플레이 포지션을 그라운드 평면에서의 대응하는 포지션에 맵핑시키는 것에 의해 로케이션 공간에서 오브젝트 포지션을 결정한다. 이 경우에, 태스크 T200 은 상술된 예들 중 임의의 예에 따라, 결정된 레퍼런스 포지션을 그라운드 평면에서의 오브젝트 포지션들의 평균으로서, 또는 그라운드 평면에서의 오브젝트 포지션들의 형성의 포지션으로서 계산하도록 구현될 수도 있다.
- [0120] 태스크 T300 은 메타데이터를 적어도 하나의 디지털 이미지와 (예를 들어, 비디오 스트림과) 연관시키도록 구현될 수도 있다. 이러한 메타데이터는 또한, 선택된 오브젝트들의 라벨들 (예를 들어, 사람의 이름), 선택된

오브젝트들 간의 기하학적 배열의 식별물 (identification) 등과 같이, 추적된 오브젝트들에 관한 다른 정보를 포함할 수도 있다. 부가적으로 또는 대안적으로, 메타데이터는 디바이스의 다른 엘리먼트들로부터 및/또는 수신한 바와 같이 적어도 하나의 이미지 또는 비디오 스트림에서의 메타데이터로부터 획득될 수도 있는 정보 (예를 들어, 카메라 설정들, 카메라 배향, 센서 특성들, 캡처의 로케이션 및/또는 시간) 를 포함할 수도 있다.

[0121] 태스크 T300 의 이러한 구현은 XML (eXtensible Markup Language), KLV (Key-Length-Value), *.srt (SubRip file format), 및 *.vtt (Web Video Text Track format) 중에서 임의의 것과 같은 원하는 포맷으로 메타데이터를 인코딩할 수도 있다. 일부 포맷들에서, 메타데이터의 적어도 일부가 디지털 이미지 또는 비디오 스트림 내에 (예를 들어, 임베딩된 데이터에 대해 예비되는 프레임 데이터의 일 부분 내에) 포함될 수도 있다. 대안적으로, 태스크 T300 은 비디오 스트림과 동기화되는 메타데이터 스트림으로서 또는 별개의 파일로서 메타데이터를 패키징하도록 구현될 수도 있다.

[0122] 비디오 스트림 상에 태스크 T300 의 반복된 인스턴스들 (예를 들어, 태스크들 (T200 및 T300) 의 반복된 인스턴스들) 을 포함하기 위해 방법 (M100) 을 구현하는 것이 바람직할 수도 있다: 예를 들어, 비디오 스트림의 각각의 프레임에 대해, 또는 각각의 제 n 프레임에 대해 (여기서 n 은 1 보다 더 큰 정수이다), 또는 특정된 이벤트 (예를 들어, 선택된 오브젝트가 이동했다는 추적 태스크에 의한 검출) 시. 이러한 경우에, 메타데이터 파일 또는 스트림 내의 메타데이터의 아이템들은 이러한 아이템과 비디오 스트림의 대응하는 프레임 사이의 원하는 연관성을 나타내기 위해 타임스탬핑될 (timestamped) 수도 있다. 이러한 경우에, 태스크 T300 은 또한, 비디오 및 메타데이터 스트림들 (가능하게는 하나 이상의 오디오 스트림들을 물론 가짐) 을 멀티플렉싱하는 것을 포함할 수도 있다.

[0123] 메타데이터를 포함하는 스트림이 송신 및/또는 저장될 경우에, 다른 정보 및/또는 포맷팅은 특정 스트림 포맷을 준수하도록 수행될 수도 있다 (예를 들어, 메타데이터 및/또는 스트림은 헤더-플러스-페이로드 (header-plus-payload) 포맷으로 패키징될 수도 있다). 태스크 T300 은 메타데이터를 포함하는 스트림을 컴프레서 또는 다른 인코더에 제공하도록 구현될 수도 있는데, 이 다른 인코더는 스트림을 인코딩하도록 구성될 수도 있다. 이러한 스트림을 저장하는데 이용될 수도 있는 저장 파일 포맷들의 예들로는 다음의 비디오 컨테이너 포맷들 중 임의의 것을 포함한다: AVI, WMV, MOV, MKV, MPG, 및 MP4.

[0124] 일부 애플리케이션들에서, 선택된 오브젝트들의 기하학적 배열은 로케이션 공간에서의 다수의 선택된 오브젝트들의 분포일 수도 있다 (예를 들어, 공공 장소에서의 많은 사람들 무리, 많은 그룹의 동물들, 많은 그룹의 에어본 또는 워터본 운반체들 또는 다른 오브젝트들). 태스크 T300 은 각 개개의 오브젝트에 대한 포지션을 나타내는 메타데이터를 생성할 필요가 없을 수도 있고, 정확한 총 개수의 선택된 오브젝트들이 또한 불필요할 수도 있다.

[0125] 이러한 경우에, 태스크 T300 은 복수의 후보 기하학적 배열들 중에서 어느 것이 로케이션 공간에서의 선택된 오브젝트들의 분포에 가장 유사한지를 결정하도록 구현될 수도 있다. 예를 들어, (예를 들어, 군중에서의 화자의 포지션과 같은) 레퍼런스 포지션에 대한 선택된 오브젝트들의 주어진 (예를 들어, 관측된) 분포에 가장 밀접하게 근사하는 후보를 결정하기 위해 태스크 T300 을 구현하는 것이 바람직할 수도 있다.

[0126] 도 36 은 태스크 T300 이 도면의 우측에 10 개의 후보들의 각각의 중심에 나타내는, 레퍼런스 포지션에 대한 다수의 오브젝트들의 분포를 매칭시키도록 구현될 수도 있는 하나의 예를 도시한다. 이 경우에, 원형 로케이션 공간은 도면의 좌측의 도형들에 도시한 바와 같이 16 개의 구역들로 분할되고, 각각의 구역에서의 오브젝트들의 개수가 결정된다.

[0127] 도 36 의 좌측의 4 개의 큰 도형들은 이 선택을 위해 이용될 수도 있는 4 개의 판정 메트릭들의 세트의 하나의 예를 예시한다: (A) 내측 구역들에서의 오브젝트들의 개수와 외측 구역들에서의 오브젝트들의 개수 사이의 비율, (B) 가장 밀집된 절반의 오브젝트들의 개수와 나머지 절반의 오브젝트들의 개수 사이의 비율 (대안적으로는, 가장 밀집된 절반의 오브젝트들의 개수와 오브젝트들의 총 개수 사이의 비율), (C) 가장 밀집된 4분면의 오브젝트들의 개수와 반대편 4분면의 오브젝트들의 개수 사이의 비율, 및 (D) 가장 밀집된 대향하는 4분면들의 오브젝트들의 개수와 나머지 2 개의 4분면들의 오브젝트들의 개수 사이의 비율 (대안적으로는, 가장 밀집된 대향하는 4분면들의 오브젝트들의 개수와 오브젝트들의 총 개수 사이의 비율).

[0128] 선택된 오브젝트들의 개수가 큰 경우라도, 로케이션 공간에서의 오브젝트들의 특정 분포에 대한 이들 4 개의 메트릭들의 값들은 매우 낮은 연산 복잡도로 계산될 수도 있다. 판정 트리는 그 후, 예를 들어, 주어진 분포에 대한 이들 메트릭들의 값들에 따라 도 36 의 우측의 10 개의 후보 기하학적 배열들 중에서 선택하는데 이용

될 수도 있다. 이러한 경우에, 태스크 T300 은 선택된 후보를 식별하고 선택된 오브젝트들의 총 개수를 나타내는 메타데이터를 생성하도록 구현될 수도 있다. 메타데이터의 사이즈를 훨씬 더 감소시키기 위해, 태스크 T300 은 (예를 들어, 총 개수를 특정된 양자화 팩터로 나누고 그 결과를 가장 가까운 정수로 라운딩함으로써) 선택된 오브젝트들의 총 개수의 근사값을 나타내는 메타데이터를 생성하도록 구현될 수도 있다.

[0129] 태스크 T300 은 또한, 도 33a, 도 33b, 도 34a 내지 도 34c 에 대하여 이하 설명한 바와 같은 기하학적 배열들 중에서 선택하는 부가적인 예들 및 형성들, 형성 데이터, 및 형성 코드북들의 연관된 논의 중 임의의 것을 수행하도록 구현될 수도 있다. 본 명세서에서 설명한 바와 같은 방법 (M100) 의 구현들은 선택된 오브젝트들 중에서의 특정 기하학적 배열에 대한 매치를 나타내는 것으로서 인덱싱된 프레임들이 메타데이터를 검색하는 것에 의한 취출을 위해 쉽게 식별되는 것을 허용하기 위해 (예를 들어, 비디오 스트림의 캡처 동안) 메타데이터를 제공하는데 이용되어, 그에 의해 전체 비디오 스트림을 검토할 필요성을 회피할 수도 있다.

[0130] 도 20 을 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 특정 실시형태가 도시된다. 도 20 에 나타내진 실시형태에 대하여 설명된 비디오 프로세싱 기법들은 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 또는 이들의 임의의 조합에 의해 수행될 수도 있다.

[0131] 도 20 은 제 1 시간 인스턴스에서 캡처되는 장면을 나타내는 디지털 이미지 (1700) 를 도시한다. 예를 들어, 이미지 (1700) 는 모바일 디바이스의 스크린 상에 디스플레이될 비디오 스트림에 대응할 수도 있다. 모바일 디바이스는 (예를 들어, 모바일 디바이스의 카메라로) 스트림을 캡처하거나 또는 다른 디바이스로부터 스트림을 수신하도록 구성될 수도 있다. 프레임 (1700) 에 나타낸 장면은 제 1 오브젝트 (1702), 제 2 오브젝트 (1704), 제 3 오브젝트 (1706), 및 별을 포함한다. 예시된 실시형태에서, 제 1 오브젝트 (1702) 는 제 1 사람에 대응할 수도 있고, 제 2 오브젝트 (1704) 는 나무에 대응할 수도 있으며, 제 3 오브젝트 (1706) 는 제 2 사람에 대응할 수도 있다. 다른 실시형태에서, 제 1 오브젝트 (1702) 와 제 3 오브젝트 (1706) 중 적어도 하나는 모바일 디바이스를 통해 제어되는 로봇에 대응할 수도 있다.

[0132] 도 20 은 또한 모바일 디바이스의 스크린 상의 이미지 (1700) 의 디스플레이 (1710) 를 도시한다. 모바일 디바이스의 스크린 (예를 들어, 터치스크린) 은 뷰파인더에 대응할 수도 있다. 디스플레이 (1710) 는 사용자가 (예를 들어, 태스크 T200 에 의한) 추적을 위해 오브젝트들을 선택하는 것을 가능하게 하는 사용자 선택 메커니즘 (1711) 을 포함할 수도 있다. 예를 들어, 사용자는 스크린 상에서 사용자 선택 메커니즘 (1711) 을 터치하고 제 1 오브젝트 (1702) 주위로 제 1 포커스 링 (1712) 을 드래그하여 제 1 오브젝트 (1702) 의 추적을 선택 및 가능하게 할 수도 있다. 이러한 제 1 포커스 링 (1712) 의 배치에 응답하여, 모바일 디바이스는 추적하기 위한 제 1 오브젝트 (1702) 를 프레임 (1700) 에서 선택할 수도 있다. 유사한 방식으로, 사용자는 또한 스크린 상에서 사용자 선택 메커니즘 (1711) 을 터치하고 제 2 오브젝트 (1704) 및 제 3 오브젝트 (1706) 주위로 제 2 포커스 링 (1714) 및 제 3 포커스 링 (1716) 을 드래그하여, 제 2 오브젝트 (1704) 및 제 3 오브젝트 (1706) 각각의 추적을 선택 및 가능하게 할 수도 있다. 이러한 제 2 포커스 링 (1714) 및 제 3 포커스 링 (1716) 의 배치에 응답하여, 모바일 디바이스는 추적하기 위한 제 2 오브젝트 (1704) 및 제 3 오브젝트 (1706) 를 각각 선택할 수도 있다.

[0133] 도 20 은 또한 (예를 들어, 제 1 시간 인스턴스에 후속하는) 제 2 시간 인스턴스에서 캡처되는 장면을 나타내는 이미지 (1720) 를 도시한다. 이미지 (1720) 에서, 각각의 오브젝트 (1702 내지 1706) 의 포지션은 제 1 시간 인스턴스에서의 장면을 나타내는 이미지 (1700) 에서의 대응하는 포지션들에 대해 변화되었다. 예를 들어, 이미지 (1720) 에서 제 1 오브젝트 (1702) 및 제 3 오브젝트 (1706) 가 이동하였다. 제 2 오브젝트 (1704) (예를 들어, 나무) 가 고정된 오브젝트이지만, 제 2 오브젝트 (1704) 는 이동한 것으로 보여질 수도 있다. 예를 들어, 모바일 디바이스 (또는 다른 캡처링 디바이스) 의 포지션이 이동할 수도 있고, 이는, 차례로, 제 2 오브젝트 (1704) 가 이동하였다는 착각을 일으킬 수도 있다.

[0134] 모바일 디바이스는 선택된 오브젝트들 (1702 내지 1706) 중에서의 하나 이상의 공간 관계들 (또한 "지오메트리" 라고도 불림) 을 모니터링할 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702) 와 제 2 오브젝트 (1704) 사이의 제 1 거리, 제 2 오브젝트 (1704) 와 제 3 오브젝트 (1706) 사이의 제 2 거리, 및 제 3 오브젝트 (1706) 와 제 1 오브젝트 (1702) 사이의 제 3 거리를 추적하고 측정할 수도 있다. 특정 실시형태에서, (예를 들어, 태스크 T200 에 의해) 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적하

는 것은 트리거 이벤트시에 게시할 수도 있다. 비제한적 예로서, 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적하는 것은 제 1 오브젝트 (1702) 가 제 2 오브젝트 (1704) 를 교차할 때에 게시할 수도 있다. 다른 실시형태들에서, 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적하는 것은 사용자 입력시에 게시할 수도 있다. 비제한적 예로서, 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적하는 것은 오브젝트들 (1702 내지 1706) 중 2 개 이상의 오브젝트들의 사용자 선택시에 게시할 수도 있다. 다른 예로서, 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적하는 것은 사용자가 추적 모드를 게시하도록 선택할 때에 게시할 수도 있다.

[0135] 도 20 은 또한 (예를 들어, 디스플레이 태스크 T50 에 의한) 모바일 디바이스의 스크린 상의 이미지 (1720) 의 디스플레이 (1730) 를 도시한다. 모바일 디바이스는 디스플레이 (1730) 상에 지오메트리의 표시를 포함할 수도 있다. 예를 들어, 제 1 오브젝트 (1702) 와 제 2 오브젝트 (1704) 사이의 제 1 거리의 제 1 표시 (L_{12}) 가 디스플레이 (1730) 에 포함될 수도 있다. 제 1 거리는 이미지 (1720) 에서의 제 1 오브젝트 (1702) 의 로케이션 및 이미지 (1720) 에서의 제 2 오브젝트 (1704) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 오브젝트 (1702) 의 로케이션 및 제 2 오브젝트 (1704) 의 로케이션을 추적하여 제 1 표시 (L_{12}) 를 발생시킬 수도 있다. 제 2 오브젝트 (1704) 와 제 3 오브젝트 (1706) 사이의 제 2 거리의 제 2 표시 (L_{23}) 가 또한 디스플레이 (1730) 에 포함될 수도 있다. 제 2 거리는 이미지 (1720) 에서의 제 2 오브젝트 (1704) 의 로케이션 및 이미지 (1720) 에서의 제 3 오브젝트 (1706) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 2 오브젝트 (1704) 의 로케이션 및 제 3 오브젝트 (1706) 의 로케이션을 추적하여 제 2 표시 (L_{23}) 를 발생시킬 수도 있다. 제 1 오브젝트 (1702) 와 제 3 오브젝트 (1706) 사이의 제 3 거리의 제 3 표시 (L_{13}) 가 또한 디스플레이 (1730) 에 포함될 수도 있다. 제 3 거리는 프레임 (1720) 에서의 제 1 오브젝트 (1702) 의 로케이션 및 프레임 (1720) 에서의 제 3 오브젝트 (1706) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702) 의 로케이션 및 제 2 오브젝트 (1704) 의 로케이션을 추적하여 제 3 표시 (L_{13}) 를 발생시킬 수도 있다.

[0136] 모바일 디바이스는 또한 오브젝트들 (1702 내지 1706) 의 합성 지오메트리 (예를 들어, 선택된 오브젝트들의 기하학적 배열) 를 추적할 수도 있다. 예를 들어, 예시된 실시형태에서, 오브젝트들 (1702 내지 1706) 의 기하학적 배열은 각각의 표시 (L_{12} , L_{23} , L_{13}) 에 의해 형성된 삼각형으로서 특성화될 수도 있다. 각각의 오브젝트 사이의 지오메트리의 적어도 하나의 파라미터, 또는 합성 지오메트리의 적어도 하나의 파라미터는 이미지들 (예를 들어, 하나 이상의 비디오 스트림들의 프레임들) 을 클러스터링하는데 이용될 수도 있다. 비제한적 예로서, 특정 이미지들은 2 개의 특정 오브젝트들 (1702 내지 1706) 사이의 거리 (예를 들어, 제 1 거리, 제 2 거리, 및/또는 제 3 거리) 에 기초하여 클러스터링될 수도 있다. 다른 비제한적 예로서, 특정 이미지들은 다수의 오브젝트들 (1702 내지 1706) 의 질량 중심 (C_{M2}) 에 기초하여 클러스터링될 수도 있다. 예를 들어, 모바일 디바이스는 각각의 이미지에 대한 각각의 표시 (L_{12} , L_{23} , L_{13}) 에 의해 형성된 삼각형의 질량 중심 (C_{M2}) 을 결정할 수도 있다. 특정 실시형태에서, 질량 중심 (C_{M2}) 은 계산되고, 인덱싱되며, 스크린 (1730) 상에 디스플레이될 수도 있다. 다른 실시형태에서, 질량 중심 (C_{M2}) 은 단지 계산되고 인덱싱될 수도 있다. 실질적으로 유사한 질량 중심을 갖는 프레임들은 함께 클러스터링되고 메모리 내에 인덱싱될 수도 있다. 이러한 추적 동안, 디스플레이의 중심이 질량 중심과 일치하도록 디스플레이 병진 동작을 제어하는 것이 바람직할 수도 있다. 부가적으로 또는 대안적으로, 선택된 오브젝트들 모두가 디스플레이 내에 가시적으로 남아있도록 카메라의 광학 및/또는 디지털 줌 동작을 제어하는 것이 바람직할 수도 있다.

[0137] 특정 실시형태에서, 모바일 디바이스는 각각의 오브젝트 (1702 내지 1706) 사이의 지오메트리 및/또는 각각의 프레임에 대한 합성 지오메트리를 추적할 수도 있다 (예를 들어, 프레임-바이-프레임 (frame-by-frame) 기반으로 지오메트리들을 추적할 수도 있다). 다른 실시형태들에서, 모바일 디바이스는 각각의 오브젝트 (1702 내지 1706) 사이의 지오메트리 및/또는 합성 지오메트리를 비연속 간격들에서 추적할 수도 있다 (예를 들어, 비연속 프레임들에서 지오메트리들을 추적할 수도 있다). 비-연속 프레임들에서 지오메트리들을 추적하는 것은 모바일 디바이스에서 소비되는 전력량을 감소 (예를 들어, 배터리 수명을 연장) 시키고, 지오메트리들과 연관된 인덱싱 정보 및/또는 클러스터링 정보에 대해 이용되는 메모리의 양을 감소시킬 수도 있다.

[0138] 예를 들어, 특정 실시형태에서, 모바일 디바이스는 매 프레임마다 각각의 오브젝트 (1702 내지 1706) (예를 들어, 3 개의 오브젝트들) 를 추적할 수도 있다. 추적 동안, 각각의 오브젝트 (1702 내지 1706) 는 오브젝트

(1702 내지 1706)의 사이즈 및 오브젝트 (1702 내지 1706)의 포지션과 연관된 적어도 3개의 값들로 표현될 수도 있다. 예를 들어, 하나의 값은 오브젝트 (1702 내지 1706)의 x-좌표에 대응할 수도 있고, 하나의 값은 오브젝트 (1702 내지 1706)의 y-좌표에 대응할 수도 있으며, 다른 값은 오브젝트 (1702 내지 1706)의 사이즈 (예를 들어, 바운딩 박스)에 대응할 수도 있다. 각각의 값은 32-비트 (예를 들어, 4 바이트) 수에 대응할 수도 있다. 따라서, 데이터의 적어도 288 비트들 (3개의 오브젝트들 * 3개의 값들/오브젝트 * 32-비트들/값)이 추적 동안 각각의 프레임에 대해 메모리 내로 수집되고 로깅될 수도 있다. 한 시간 (예를 들어, 3600 초) 동안 프레임-바이-프레임 기반으로 3개의 오브젝트들 (1702 내지 1706)을 추적하고 인덱싱하기 위해, 모바일 디바이스가 초 당 30개의 프레임들을 캡처한다고 가정하면, 데이터의 적어도 3,888,000 바이트들이 메모리 내로 수집되고 로깅된다.

[0139] 그러나, 비-연속 프레임들에서 지오메트리들을 추적하는 것은 메모리 요건들을 완화시킬 수도 있다. 비제한적 예로서, 한 시간 동안 15개의 프레임들의 매 간격마다 각각의 오브젝트 (1702 내지 1706)를 추적하고 인덱싱하는 것은 메모리 공간의 93.33% 절약을 가져올 수도 있다. 비-연속 프레임들에서 지오메트리들을 추적하는 것은 오브젝트들 (1702 내지 1706)과 연관된 비교적 작은 양의 이동이 존재하는 시나리오들에서 특이 이룰 수도 있다. 모바일 디바이스는 사용자 입력 (예를 들어, 사용자 선호도)에 기초하여 프레임들을 추적하고 인덱싱하는 빈도를 결정할 수도 있거나 및/또는, 예를 들어, 하나 이상의 선택된 오브젝트들 각각의 포지션에서의 시간이 지남에 따른 변화의 정도 및/또는 사용자 표시에 기초하여 빈도를 다르게 할 수도 있다.

[0140] 각각의 오브젝트 (1702 내지 1706)는 또한 스크린 (1730) 상의 오브젝트 (1702 내지 1706)의 로케이션에 대응하는 개개의 질량 중심을 가질 수도 있다. 예를 들어, 제 1 오브젝트 (1702)는 제 1 오브젝트 (1702)의 로케이션에 질량 중심 (C_{01})을 가질 수도 있고, 제 2 오브젝트 (1704)는 제 2 오브젝트 (1704)의 로케이션에 질량 중심 (C_{02})을 가질 수도 있다는 것 등이다. 각각의 오브젝트 (1702 내지 1706)에 대한 질량 중심은 또한 모바일 디바이스가 추적하는 기하학적 파라미터일 수도 있다.

[0141] 도 20은 또한 (예를 들어, 제 2 시간 인스턴스에 후속하는) 제 3 시간 인스턴스에서 캡처되는 장면을 나타내는 이미지 (1740)를 도시한다. 이미지 (1740)에서, 각각의 오브젝트 (1702 내지 1706)의 포지션은 제 2 시간 인스턴스에서의 이미지 (1720)에 나타난 대응하는 포지션들에 대해 변화되었다.

[0142] 도 20은 또한 (예를 들어, 디스플레이 태스크 T50에 의한) 모바일 디바이스의 스크린 상의 이미지 (1740)의 디스플레이 (1750)를 도시한다. 모바일 디바이스는 디스플레이 (1750)에서의 지오메트리의 표시를 디스플레이할 수도 있다. 예를 들어, 제 1 오브젝트 (1702)와 제 2 오브젝트 (1704)사이의 제 1 거리의 제 1 표시 (L_{12})는 디스플레이 (1750)에 포함될 수도 있고, 제 2 오브젝트 (1704)와 제 3 오브젝트 (1706)사이의 제 2 거리의 제 2 표시 (L_{23})는 디스플레이 (1750)에 포함될 수도 있으며, 제 1 오브젝트 (1702)와 제 3 오브젝트 (1706)사이의 제 3 거리의 제 3 표시 (L_{13})는 디스플레이 (1750)에 포함될 수도 있다. 부가적으로, 모바일 디바이스는 또한 오브젝트들 (1702 내지 1706)의 합성 지오메트리를 추적할 수도 있다. 예를 들어, 예시된 실시형태에서, 오브젝트들 (1702 내지 1706)의 합성 지오메트리는 각각의 표시 (L_{12} , L_{23} , L_{13})에 의해 형성된 삼각형에 대응할 수도 있다.

[0143] 다른 실시형태에서, 모바일 디바이스는 각각의 오브젝트 (1702 내지 1706)의 이동을 추적하고, 시간이 지남에 따라 각각의 오브젝트 (1702 내지 1706)의 이동을 디스플레이하는 플롯들의 시퀀스를 발생시킬 수도 있다. 모바일 디바이스는 매 프레임마다 각각의 오브젝트 (1702 내지 1706)에 대한 플롯을 발생시킬 수도 있고, 또는 비-연속 프레임들에서 각각의 오브젝트 (1702 내지 1706)에 대한 플롯을 발생시켜서, 상술된 바와 같이, 메모리 요건들을 완화시키고 전력 절약들을 개선시킬 수도 있다.

[0144] 현재 디스플레이 상에서 가시적이지 않은 (예를 들어, 비디오 스트림에 의해 캡처된 장면 내에 현재 있지 않은) 오브젝트들에 대한 하나 이상의 선택된 오브젝트들의 공간 관계를 추적하도록 디바이스를 구성하는 것이 (예를 들어, 태스크 T200을 구현하는 것이) 또한 가능하다. 예를 들어, 현재 스크린 밖에 있는 고정된 (레퍼런스) 오브젝트들에 대한 이동하는 오브젝트들의 공간 관계들이 추적될 수도 있다. 특정 예에서, 디바이스는 선택된 선수가 20 피트의 골 (goal) 또는 바스켓 (basket) 내에 있거나, 및/또는, 골/바스켓이 현재 스크린 안에 있지 않은 경우에도, 골/바스켓을 향해 이동하고 있을 때마다 나타내도록 구성될 수도 있다. 이러한 경우, 디바이스는 현재 디스플레이된 장면보다 더 큰 맵을 생성하도록 구성될 수도 있다. (예를 들어, GPS 센서와 같은 로케이션 센서에 의해 나타낸 바와 같이) 카메라의 로케이션이 알려진 경우에 대해, 현재 뷰가 그 더 큰 장면 맵에 얼마나 관련되는지를 결정하기 위해 카메라의 하나 이상의 배향 센서들이 이용될 수도

있다. 이러한 하나 이상의 배향 센서들은 중력 축 (예를 들어, 지구의 중심과 디바이스를 통과한 축) 에 대한 카메라의 배향을 나타내기 위해, 하나 이상의 관성 센서들 (예를 들어, 자이로스코프들 및/또는 가속도계들), 및/또는 하나 이상의 자계 센서들 (예를 들어, 자력계들) 을 포함할 수도 있다. 부가적으로 또는 대안적으로, 스포츠 필드 또는 코트 상의 마킹들, 또는 공연 무대의 장면 고정물들 (scene fixtures) 과 같은, 고정된 레퍼런스 오브젝트들의 가시적 피쳐들이 기점들 (fiducials) 로서 이용될 수도 있다. 이러한 맵핑은 또한 이동하고 있는 카메라에 대해 행해질 수도 있지만, 레퍼런스 오브젝트의 외관은 그것이 뷰잉되게 되는 각도가 변화함에 따라 변화할 수도 있다.

[0145] 도 20 에 나타낸 실시형태는 모바일 디바이스가 이미지들 (1700, 1720, 1740) 에서 오브젝트들 (1702 내지 1706) 의 지오메트리들을 추적하는 것, 그리고 도 21 에 대해 후술되는 바와 같이, 실질적으로 유사한 지오메트리들을 갖는 추출을 위한 프레임들을 클러스터링하는 것을 가능하게 할 수도 있다. 모바일 디바이스는 단일 카메라를 이용하여 또는 다수의 카메라들을 이용하여 도 20 에 대해 설명되는 기법들을 수행할 수도 있다.

[0146] 본 명세서에서 설명되는 바와 같이 메타데이터를 발생시키기 위한 용도들 및 적용들은 다음 중 임의의 것을 포함할 수도 있다: 스포츠 이벤트들 (예를 들어, 축구, 농구, 풋볼, 하키), 소셜 이벤트들 (예를 들어, 웨딩, 파티, 댄싱), 예술 공연들 (예를 들어, 연극, 학교 연극, 콘서트, 이중주), 및 보안 또는 감시 모니터링. 모바일 디바이스는 또한, 예를 들어, 조이스틱, 제스처-인식 카메라 시스템, 또는 사용자 커맨드들을 검출하기 위한 터치 및/또는 배향 센서들 (예를 들어, 상술된 하나 이상의 배향 센서들) 이 구비된 글로브의 이용에 의해 사용자 선택이 수행되는 헤드-마운티드 (head-mounted) 디스플레이일 수도 있다. 비디오 스트림은 라이브일 수도 있고 (예를 들어, 디바이스에 의해 캡처되거나 또는 다른 캡처링 디바이스로부터, 예를 들어, 블루투스나 같은 단거리 연결을 통해, 또는 인터넷에 대한 Wi-Fi 연결과 같은 네트워크를 통해 스트리밍될 수도 있고) 또는 (다시, 디바이스에 의해 또는 다른 디바이스에 의해) 레코딩될 수도 있다.

[0147] 스트림 내의 선택된 오브젝트들의 상대 로케이션들 및/또는 그 선택된 오브젝트들 중에서의 공간 관계들에 따라 비디오 스트림의 프레임들을 분류하는 것이 바람직할 수도 있다. 도 21 을 참조하면, 클러스터링에 기초하여 프레임들을 추출하는데 이용되는 모바일 디바이스의 디스플레이 (1800) 의 특정 실시형태가 도시된다. 디스플레이 (1800) 는 수평 축과 수직 축을 갖는 그래프를 포함할 수도 있다. 수평 축은 모바일 디바이스에 의해 캡처되는 상이한 비디오 파일들에 대응할 수도 있다. 예를 들어, 예시된 실시형태에서, 모바일 디바이스는 제 1 비디오 파일 및 제 2 비디오 파일을 캡처하였다. 수직 축은 비디오 파일들의 시간 인덱스들에 대응할 수도 있다.

[0148] 각각의 비디오 파일은 비디오 파일에서 선택된 오브젝트들의 지오메트리와 연관된 적어도 하나의 파라미터에 기초하는 클러스터들을 포함할 수도 있다. 예를 들어, 비디오 파일들에서의 각각의 클러스터는 선택된 오브젝트들 사이에서 실질적으로 유사한 지오메트리를 갖는 프레임들의 그룹을 포함할 수도 있다. 도 21 에 나타낸 클러스터들은 오브젝트들 사이의 질량 중심에 기초할 수도 있다; 그러나, 다른 실시형태들에서, 클러스터들은 하나 이상의 상이한 파라미터들 (예를 들어, 형상, 거리, 공간 특성들, 컬러 스킴들 등) 에 기초할 수도 있다. 파라미터들은 사용자에게 의해 선택될 수도 있다.

[0149] 제 1 비디오의 시작에서 제 1 비디오의 대략 10:00 마크까지의 프레임들은 제 1 클러스터 (예를 들어, 클러스터 1) 로 그룹화될 수도 있다. 제 1 클러스터는 선택된 오브젝트들 사이의 스크린 (1800) 상의 질량 중심이 대략 150 에서의 x-좌표 및 대략 250 에서의 y-좌표를 갖는 프레임들에 대응할 수도 있다. 도 20 의 제 2 프레임을 참조하면, 선택된 오브젝트들 (1702 내지 1706) 의 질량 중심 (C_{M2}) (예를 들어, 삼각형의 질량 중심 (C_{M2})) 은 스크린 (1730) 상의 대략 150 에서의 x-좌표 및 스크린 (1730) 상의 대략 250 에서의 y-좌표를 가질 수도 있다. 따라서, 제 2 프레임 (및 오브젝트들 (1702 내지 1706) 사이에서 실질적으로 유사한 질량 중심을 갖는 다른 프레임들) 은 도 21 의 제 1 클러스터에 배치될 수도 있다. 따라서, 사용자가 제 1 클러스터를 선택하는 경우, 모바일 디바이스는 대응하는 비디오 스트림들 (예를 들어, 비디오의 부분들) 을 디스플레이할 수도 있는데 여기서 선택된 오브젝트들 (1702 내지 1706) 은 도 20 의 제 2 프레임과 실질적으로 유사한 구성을 갖는다.

[0150] 제 1 비디오의 대략 10:00 마크로부터 제 1 비디오의 대략 20:00 마크까지의 프레임들은 제 2 클러스터 (예를 들어, 클러스터 2) 로 그룹화될 수도 있다. 제 2 클러스터는 선택된 오브젝트들 사이의 스크린 (1800) 상의 질량 중심이 대략 100 에서의 x-좌표 및 대략 250 에서의 y-좌표를 갖는 프레임들에 대응할 수도 있다. 도 20 의 제 3 프레임을 참조하면, 선택된 오브젝트들 (1702 내지 1706) 의 질량 중심 (C_{M3}) 은 스크린 (1750) 상

의 대략 100 에서의 x-좌표 및 스크린 (1750) 상의 대략 250 에서의 y-좌표를 가질 수도 있다. 따라서, 제 3 프레임 (및 오브젝트들 (1702 내지 1706) 사이에서 실질적으로 유사한 질량 중심을 갖는 다른 프레임들) 은 도 21 의 제 2 클러스터에 배치될 수도 있다. 따라서, 사용자가 제 2 클러스터를 선택하는 경우, 모바일 디바이스는 대응하는 비디오 스트림들 (예를 들어, 비디오의 부분들) 을 디스플레이할 수도 있는데 여기서 선택된 오브젝트들 (1702 내지 1706) 은 도 21 의 제 3 프레임과 실질적으로 유사한 구성을 갖는다.

[0151] 부가적인 클러스터들 (예를 들어, 제 3 클러스터) 은 상이한 질량 중심들 (예를 들어, 175 에서의 x-좌표 및 325 에서의 y-좌표) 을 갖는 프레임들에 대응할 수도 있다. 사용자는 비디오의 특정 부분들과 연관된 구성 (예를 들어, 질량 중심) 을 갖는 클러스터들을 선택하는 것에 의해 비디오의 특정 부분들을 취출 (예를 들어, 로케이팅) 할 수도 있다. 따라서, 선택된 오브젝트들 (1702 내지 1706) 의 기하학적 파라미터들 (예를 들어, 각각의 및/또는 모든 선택된 오브젝트들의 질량 중심, 거리) 에 기초하는 클러스터들로 프레임들을 인덱싱하는 것에 의해, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기 (fast-forward), 되감기 등) 해야 하는 일 없이 비디오의 원하는 부분들을 쉽게 로케이팅할 수도 있다. 클러스터들로 프레임들을 인덱싱하는 것은 또한 어떤 구성 (예를 들어, 지오메트리) 이 최다 빈도로, 최소 빈도로 등으로 발생했는지를 뷰잉하는 것을 사용자에게 허용할 수도 있다. 이러한 기하학적 파라미터들의 이러한 인덱싱 및/또는 값들은 상술된 바와 같이 태스크 T300 에 의해 생성된 메타데이터 중에 포함될 수도 있다.

[0152] 도 21 은 또한 비디오 스트림의 프레임들의 클러스터들과 연관된 구역들의 디스플레이 (1802) 를 포함한다. 예를 들어, 클러스터들의 디스플레이 (1802) 는 도 20 의 장면의 비디오 스트림에 대응할 수도 있다. 디스플레이 (1802) 는 10 개의 구역들 (C1 내지 C10) 을 포함한다. 예시된 실시형태에서, 구역들 (C1 내지 C10) 은 특정 프레임들에서의 오브젝트들 (1702 내지 1706) 의 질량 중심에 기초한다. 다른 실시형태들에서, 구역들 (C1 내지 C10) 은 다른 기하학적 파라미터들 (예를 들어, 오브젝트들 (1702 내지 1706) 각각의 질량 중심, 형상들, 거리들, 공간 특성들, 컬러 스킴들 등) 에 기초할 수도 있다. 특정 실시형태에서, 각각의 구역 (C1 내지 C10) 은 동일한 사이즈를 가질 수도 있다 (예를 들어, 각각의 구역 (C1 내지 C10) 은 "보로노이 셀 (Voronoi Cell)" 일 수도 있다).

[0153] 각각의 구역 (C1 내지 C10) 은 특정 클러스터와 연관된 질량 중심에 대응하는 x-좌표 및 y-좌표를 포함할 수도 있다. 비제한적 예로서, 제 1 구역 (C1) 에 대응하는 질량 중심은 75 의 x-좌표 및 580 의 y-좌표를 가질 수도 있다. 부가적으로, 제 2 구역 (C2) 에 대응하는 질량 중심은 215 의 x-좌표 및 580 의 y-좌표를 가질 수도 있다. 도 20 의 장면의 비디오 스트림에서의 프레임들은 오브젝트들 (1702 내지 1706) 의 질량 중심에 의해 인덱싱될 수도 있다. 각각의 프레임은 질량 중심에 기초하여 구역들 (C1 내지 C10) 중 하나에 대응하는 클러스터에 배치될 수도 있다. 예를 들어, 제 1 구역 (C1) 의 질량 중심과 가장 가까운 질량 중심을 갖는 프레임은 대응하는 제 1 클러스터 내에 배치될 수도 있다. 이러한 클러스터링 정보는 태스크 T300 에 의해 생성되고 메타데이터로서 (예를 들어, 본 명세서에서 설명되는 바와 같이 비디오 스트림과 동기화되는 메타데이터 스트림에서) 대응하는 프레임들과 연관될 수도 있다. 이러한 경우, 특정 클러스터 내에서 인덱싱되는 프레임들이 메타데이터를 검색함으로써 취출을 위해 쉽게 식별되고, 그에 의해 전체 비디오 스트림을 검토할 필요성을 피하게 할 수도 있다.

[0154] 모바일 디바이스는 특정 클러스터의 사용자 선택에 응답하여 특정 지오메트리 (예를 들어, 질량 중심) 를 갖는 비디오 프레임들을 디스플레이할 수도 있다. 예를 들어, 사용자가 제 1 구역 (C1) 을 선택 (예를 들어, 터치) 하는 경우, 모바일 디바이스는 오브젝트들 (1702 내지 1706) 의 질량 중심이 제 1 (C1) 구역 내에 있는 비디오 프레임들의 클러스터를 디스플레이할 수도 있다. 특정 실시형태에서, 사용자는 제 1 구역 (C1) 에서 스크린을 터치하고, 임계치 (예를 들어, 3 초) 를 초과하는 시간의 주기 동안 제 1 구역 (C1) 을 홀딩할 수도 있다. 3 초 동안 제 1 구역을 홀딩한 후에, 제 1 구역 (C1) 은 디스플레이 (1802) 상에서 확대될 수도 있다. 예를 들어, 디스플레이 (1802) 는 1804 에서 도시된 바와 같이 제 1 구역 (C1) 을 나타낼 수도 있다. 이 모드에서, 제 1 구역 (1804) 은 특정 프레임들의 질량 중심 ($C_{M1023-1026}$) 에 기초하여 특정 프레임들 (예를 들어, 프레임 (1023) 내지 프레임 (1026)) 을 예시할 수도 있다. 특정 실시형태에서, 사용자는 특정 프레임 (예를 들어, 프레임 (1024)) 을 선택할 수도 있고, 모바일 디바이스는 프레임 (1024) 에서 비디오 스트림의 재생을 개시할 수도 있다.

[0155] 도 22 를 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법 (1900) 의 특정 실시형태를 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법 (1900) 은 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추

적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9의 평활화 모듈 (961), 또는 이들의 임의의 조합을 이용하여 수행될 수도 있다.

[0156] 방법 (1900)은, 1902에서, 모바일 디바이스에서, 장면에서의 다수의 오브젝트들의 선택을 받는 단계를 포함한다. 예를 들어, 도 20을 참조하면, 사용자는 스크린 (1710) 상에서 사용자 선택 메커니즘 (1711)을 터치하고 제 1 오브젝트 (1702) 주위로 제 1 포커스 링 (1712)을, 제 2 오브젝트 (1704) 주위로 제 2 포커스 링 (1714)을, 그리고 제 3 오브젝트 (1706) 주위로 제 3 포커스 링 (1716)을 드래그하여, 제 1, 제 2, 및 제 3 오브젝트들 (1702 내지 1706) 각각의 추적을 가능하게 할 수도 있다. 디스플레이된 장면 (1700)은 모바일 디바이스의 스크린 상에서 캡처되고 디스플레이되는 비디오 스트림에 대응할 수도 있다. 모바일 디바이스는 (예를 들어, 모바일 디바이스의 카메라로) 스트림을 캡처하거나 또는 다른 디바이스로부터 스트림을 수신하도록 구성될 수도 있다.

[0157] 1904에서, 각각의 오브젝트 사이의 지오메트리가 추적될 수도 있다. 예를 들어, 도 20을 참조하면, 모바일 디바이스는 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적할 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702)와 제 2 오브젝트 (1704) 사이의 제 1 거리, 제 2 오브젝트 (1704)와 제 3 오브젝트 (1706) 사이의 제 2 거리, 및 제 3 오브젝트 (1706)와 제 1 오브젝트 (1702) 사이의 제 3 거리를 추적하고 측정할 수도 있다.

[0158] 1906에서, 지오메트리의 표시가 스크린 상에 디스플레이될 수도 있다. 예를 들어, 도 20을 참조하면, 제 1 오브젝트 (1702)와 제 2 오브젝트 (1704) 사이의 제 1 거리의 제 1 표시 (L_{12})가 디스플레이 (1730)에 포함될 수도 있다. 제 1 거리는 이미지 (1720)에서의 제 1 오브젝트 (1702)의 로케이션 및 이미지 (1720)에서의 제 2 오브젝트 (1704)의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702)의 로케이션 및 제 2 오브젝트 (1704)의 로케이션을 추적하여 제 1 표시 (L_{12})를 발생시킬 수도 있다. 제 2 오브젝트 (1704)와 제 3 오브젝트 (1706) 사이의 제 2 거리의 제 2 표시 (L_{23})가 또한 디스플레이 (1730)에 포함될 수도 있다. 제 2 거리는 이미지 (1720)에서의 제 2 오브젝트 (1704)의 로케이션 및 이미지 (1720)에서의 제 3 오브젝트 (1706)의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 2 오브젝트 (1704)의 로케이션 및 제 3 오브젝트 (1706)의 로케이션을 추적하여 제 2 표시 (L_{23})를 발생시킬 수도 있다. 제 1 오브젝트 (1702)와 제 3 오브젝트 (1706) 사이의 제 3 거리의 제 3 표시 (L_{13})가 또한 디스플레이 (1730)에 포함될 수도 있다. 제 3 거리는 이미지 (1720)에서의 제 1 오브젝트 (1702)의 로케이션 및 이미지 (1720)에서의 제 3 오브젝트 (1706)의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702)의 로케이션 및 제 2 오브젝트 (1704)의 로케이션을 추적하여 제 3 표시 (L_{13})를 발생시킬 수도 있다.

[0159] 도 22의 방법 (1900)은 모바일 디바이스가 이미지들 (1700, 1720, 1740)에서 오브젝트들 (1702 내지 1706)의 지오메트리들을 추적하는 것, 그리고 실질적으로 유사한 지오메트리들을 갖는 쿼리를 위한 프레임들을 클러스터링하는 것을 가능하게 할 수도 있다. 모바일 디바이스는 단일 카메라를 이용하여 또는 다수의 카메라들을 이용하여 방법 (1900)을 수행할 수도 있다. 추가 예에서, 이 방법은 하나 이상의 오브젝트들 중에서의 특정된 공간 관계가 (예를 들어, 태스크 T200에 의해) 검출될 때, 예컨대, 선택된 제 1 오브젝트와 선택된 제 2 오브젝트 사이의 거리가 특정된 임계 거리보다 더 작아질 (대안적으로, 더 커질) 때 경보 조건을 나타내는 단계를 포함할 수도 있다.

[0160] 도 23을 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법들 (2000, 2010)의 특정 실시형태들을 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법들 (2000, 2010)은 도 1의 전자 디바이스 (102), 도 2b의 프로세서 (264)에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9의 평활화 모듈 (961), 또는 이들의 임의의 조합을 이용하여 수행될 수도 있다.

[0161] 방법 (2000)은, 2002에서, 모바일 디바이스에서, 장면에서의 다수의 오브젝트들의 선택을 받는 단계를 포함한다. 예를 들어, 도 20을 참조하면, 사용자는 스크린 (1710) 상에서 사용자 선택 메커니즘 (1711)을 터치하고 제 1 오브젝트 (1702) 주위로 제 1 포커스 링 (1712)을, 제 2 오브젝트 (1704) 주위로 제 2 포커스 링 (1714)을, 그리고 제 3 오브젝트 (1706) 주위로 제 3 포커스 링 (1716)을 드래그하여, 제 1, 제 2, 및 제 3 오브젝트들 (1702 내지 1706) 각각의 추적을 가능하게 할 수도 있다. 디스플레이된 장면 (1700)은 모바일 디바이스의 스크린 상에서 캡처되고 디스플레이되는 비디오 스트림에 대응할 수도 있다. 모바일 디바이스는

(예를 들어, 모바일 디바이스의 카메라로) 스트림을 캡처하거나 또는 다른 디바이스로부터 스트림을 수신하도록 구성될 수도 있다.

- [0162] 2004 에서, 각각의 오브젝트 사이의 지오메트리가 추적될 수도 있다. 예를 들어, 도 20 을 참조하면, 모바일 디바이스는 각각의 선택된 오브젝트 (1702 내지 1706) 사이의 지오메트리를 추적할 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (1702) 와 제 2 오브젝트 (1704) 사이의 제 1 거리, 제 2 오브젝트 (1704) 와 제 3 오브젝트 (1706) 사이의 제 2 거리, 및 제 3 오브젝트 (1706) 와 제 1 오브젝트 (1702) 사이의 제 3 거리를 추적하고 측정할 수도 있다. 부가적으로, 모바일 디바이스는 또한 오브젝트들 (1702 내지 1706) 의 합성 지오메트리를 추적할 수도 있다. 예를 들어, 예시된 실시형태에서, 오브젝트들 (1702 내지 1706) 의 합성 지오메트리는 3 개의 표시들 (L_{12} , L_{23} , L_{13}) 에 의해 형성될 수도 있는 삼각형에 대응할 수도 있다.
- [0163] 2006 에서, 비디오 스트림의 프레임들이 지오메트리와 연관된 적어도 하나의 파라미터에 기초하여 클러스터링될 수도 있다. 예를 들어, 도 21 을 참조하면, 비디오 파일들에서의 각각의 클러스터는 선택된 오브젝트들 (1706 내지 1708) 사이에 실질적으로 유사한 지오메트리를 갖는 프레임들의 그룹을 포함할 수도 있다. 도 21 에 나타난 클러스터들은 오브젝트들 (1706 내지 1708) 사이의 질량 중심 (C_{M2} , C_{M3}) (예를 들어, 삼각형의 질량 중심) 에 기초할 수도 있다.
- [0164] 특정 실시형태에서, 방법 (2000) 은 오브젝트들이 특정 지오메트리에 있을 때 프레임들을 추출하는 단계를 포함할 수도 있다. 예를 들어, 도 21 을 참조하면, 모바일 디바이스는 프레임들에서의 오브젝트들 (1702 내지 1706) 의 질량 중심 (C_{M2} , C_{M3}) 에 기초하여 프레임들을 인덱싱할 수도 있다. 특정 질량 중심 (예를 들어, 특정 지오메트리) 을 갖는 프레임들은 특정 질량 중심과 연관된 클러스터를 선택하는 것에 의해 쉽게 달성될 수도 있다. 예를 들어, 모바일 디바이스는 오브젝트들 (1702 내지 1706) 이 도 21 의 스크린 (1800) 상에서 제 1 클러스터 (예를 들어, 클러스터 1) 를 선택하는 것에 의해 도 20 의 제 2 프레임에서 오브젝트들 (1702 내지 1706) 과 실질적으로 유사한 지오메트리를 가질 때 프레임들을 추출할 수도 있다. 예를 들어, 사용자가 제 1 클러스터를 선택하는 것에 응답하여, 모바일 디바이스는 오브젝트들 (1702 내지 1706) 의 질량 중심이 대략 150 에서의 x-좌표 및 대략 250 에서의 y-좌표를 프레임들을 추출할 수도 있다.
- [0165] 방법 (2010) 은, 2012 에서, 모바일 디바이스에서, 비디오 스트림에서 선택된 오브젝트들의 특정 지오메트리의 표시를 수신하는 단계를 포함할 수도 있다. 비제한적 예로서, 도 21 을 참조하면, 모바일 디바이스는 오브젝트들 (1702 내지 1704) 의 질량 중심이 제 1 구역 ($C1$) 의 질량 중심에 가장 가까운 비디오 스트림들을 디스플레이하기 위해 표시 (예를 들어, 사용자가 디스플레이 (1802) 에 도시된 제 1 구역 ($C1$) 에서 스크린을 터치함) 를 수신할 수도 있다.
- [0166] 2014 에서, 비디오 스트림의 프레임들은 특정 지오메트리에 기초하여 추출될 수도 있다. 예를 들어, 도 21 을 참조하면, 모바일 디바이스는 오브젝트들 (1702 내지 1704) 의 질량 중심이 제 1 구역 ($C1$) 에 대응하는 영역에 있는 도 20 의 장면에서의 비디오 스트림의 프레임들을 추출할 수도 있다.
- [0167] 2016 에서, 추출된 프레임들은 모바일 디바이스의 스크린 상에 디스플레이될 수도 있다. 예를 들어, 도 21 을 참조하면, 모바일 디바이스는 오브젝트들 (1702 내지 1704) 의 질량 중심이 제 1 구역 ($C1$) 에 대응하는 영역에 있는 비디오 스트림의 프레임들을 디스플레이 (예를 들어, 플레이) 할 수도 있다. 특정 실시형태에서, 모바일 디바이스는 순차적인 순서로 비디오 스트림들을 플레이할 수도 있다.
- [0168] 도 23 의 방법들 (2000, 2010) 은 사용자가 비디오의 특정 부분들과 연관된 구성 (예를 들어, 질량 중심) 을 갖는 클러스터들을 선택하는 것에 의해 비디오의 특정 부분들을 추출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 따라서, 선택된 오브젝트들 (1702 내지 1706) 의 기하학적 파라미터들에 기초하는 클러스터들로 프레임들을 인덱싱하는 것 (예를 들어, 본 명세서에서 설명한 바와 같이 태스크 T300 에 의해 메타데이터를 생성하는 것) 에 의해, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기, 되감기 등) 해야 하는 일 없이 비디오의 원하는 부분들을 쉽게 로케이팅할 수도 있다.
- [0169] 태스크 T300 은 또한 결정된 레퍼런스 포지션에 기초하여, 로케이션 공간을 분할하는 복수의 별개의 구역들 중 하나를 식별하는 메타데이터를 생성하도록 구현될 수도 있고, 복수의 별개의 구역들은 동일하지 않은 사이즈의 구역들을 포함한다. 상술한 바와 같이, 로케이션 공간은 적어도 하나의 디지털 이미지의 픽셀 좌표 공간 (예를 들어, 비디오 스트림의 적어도 하나의 프레임의 픽셀 좌표 공간) 또는 2 개 이상의 차원들에서, 물리적 공간 (예를 들어, 장면 공간의 그라운드 평면) 일 수도 있고, 개개의 프레임에서 표현된 공간을 넘어 연장할

수도 있다. 이러한 메타데이터는 예를 들어, 레퍼런스 위치션을 포함하는 로케이션 공간의 구역을 나타내는 (예를 들어, 하나 이상의 양자화 코드북들로의) 하나 이상의 인덱스들로서 구현될 수도 있다.

[0170] 분할 스킴은 로케이션 공간을 불균일하게 분할하도록 (예를 들어, 동일하지 않은 사이즈 및/또는 형상의 구역들을 포함하도록) 구성될 수도 있다. 예를 들어, 관측된 데이터 포인트들 (예를 들어, 오브젝트 위치선들)의 세트는 각각의 별개의 구역이 m 개의 관측된 데이터 포인트들을 포함하도록 로케이션 공간을 분할하는 분할 스킴 (예를 들어, 발생 빈도에 따른 분할)을 생성하는데 이용될 수도 있다. 대안적으로 또는 부가적으로, 분할 스킴은 높은 관심있는 공간의 영역들이 낮은 관심있는 영역들보다 더 많은 별개의 구역들로 분할되도록 로케이션 공간을 분할하도록 구성될 수도 있다.

[0171] 분할 스킴에 의해 나타내진 다양한 구역들은 로케이션 공간의 각각의 영역들과 연관된 상이한 관심 정도들에 따라 집중될 수도 있다. 도 19b 내지 도 19d 는 (예를 들어, 도 19a 에 나타낸 바와 같이) 농구 코트에 적용되는 바와 같이 로케이션 공간의 불균일한 분할의 여러 예들을 도시한다. 이러한 분할 스킴은 (예를 들어, 도 19b 및 도 19d 에서와 같이) 코트 공간의 특정 바스켓에 가까운 더 작은 구역들로의 또는 (예를 들어, 도 19c 에서와 같이) 어느 하나의 바스켓에 가까운 더 작은 영역들로의 분할을 인코딩할 수도 있다. 방법 (M100) 이 디지털 이미지 또는 비디오 스트림의 캡처 동안 수행될 수도 있도록, 선험적으로 (예를 들어, 스트림이 캡처되기 전에) 로케이션 공간의 관심-기반 불균일한 분할을 나타내는 것이 가능할 수도 있다.

[0172] 유사한 방식으로, 분할 스킴에 의해 나타내진 다양한 구역들은 로케이션 공간의 각각의 영역들 내의 맵핑된 데이터 포인트 (예를 들어, 오브젝트의 위치선 또는 다수의 오브젝트들의 질량 중심)의 발생 빈도의 차이들에 따라 집중될 수도 있다. 이러한 경우에, 로케이션 공간의 불균일한 분할은 귀납적으로 (예를 들어, 캡처된 스트림에 대한 훈련의 결과로서) 획득될 수도 있다. 분할은 각각의 별개의 구역에서 동일한 수의 관측들이 발견되도록 구현될 수도 있다.

[0173] 하나의 예에서, 분할 스킴은 로케이션 공간에서의 오브젝트 위치선들 (예를 들어, 관측된 오브젝트 위치선들)의 k 개의 클러스터들로의 k -평균 클러스터링을 수행하는 것에 의해 획득된다. 예를 들어, 분할 스킴은 비디오 스트림의 복수의 프레임들에서 관측된 바와 같은 오브젝트 위치선들 (예를 들어, 선택된 오브젝트들 중 하나 이상의 위치선들)의 k 개의 클러스터들로의 k -평균 클러스터링을 수행하는 것에 의해 획득될 수도 있다. 이러한 클러스터링은 로케이션 공간을 보로노이 셀들로 파티셔닝하고, 여기서 결과의 k 평균은 셀들의 발생기들 (예를 들어, 중심들)이고, 이러한 방식으로 (예를 들어, 클러스터링을 이용하여) 획득된 분할 스킴은 양자화 코드북이라고 또한 불릴 수도 있다. 클러스터는 다른 클러스터와는 상이한 수의 관측된 오브젝트를 가질 수도 있다. 초기 조건들 (즉, 초기 k 평균)을 획득하기 위하여, 이러한 분할-스킴-발생 태스크는 관측된 위치선들에 대해 계층적 클러스터링 방법 (예를 들어, 분할 또는 응집 클러스터링)을 수행하도록 구현될 수도 있다. 불균일한 분할을 획득하는데 이용될 수도 있는 유사성의 척도의 하나의 예는 로케이션 공간에서의 관측된 오브젝트 위치선들 사이의 유클리디안 거리이다.

[0174] 분할-스킴-발생 태스크에 의해 (예를 들어, 아웃라이어들의 존재 시에) 이용될 수도 있는 k -평균 클러스터링에 대한 하나의 대안은 k -중앙객체 클러스터링이다. 다른 대안은 분포 모드들을 상이한 구역들로 분리하는데 이용될 수도 있는 평균-시프트 클러스터링이다. 태스크 T300 은 상이한 길이들일 수도 있고 및/또는 상이한 사이즈의 구역들에 대응할 수도 있는 다수의 분할 스킴들 (예를 들어, 다수의 코드북들) 중에서 선택하는 단계를 포함할 수도 있다. 예를 들어, 태스크 T300 은 나타낸 구역에 대응하는 복수의 이러한 제 2-레벨 스킴들 중에서, 제 1 레벨의 분할 스킴이 복수의 별개의 구역들 중 하나를 나타내고 또한 제 2 레벨의 분할 스킴을 나타내도록 분할 스킴들의 계층을 이용할 수도 있다. 이러한 계층은 서로 상이하게 복수의 별개의 구역들 중 하나를 서브분할하는데 이용될 수도 있다. 하나의 이러한 예에서, 제 1 코드북은 태스크 T300 이 결정된 위치선에 대한 하나보다 더 많은 코드북을 포함한 메타데이터를 생성할 수도 있도록, 제 1 코드북이 관측된 위치선들의 분포에서 상이한 모드들에 대응하는 상이한 양자화 구역들을 인덱싱하고, 제 2 코드북들은 이들 모달 구역들 중 하나 이상의 각각을 서브-구역들로 분할하는데 이용된다.

[0175] 태스크 T300 은 예를 들어, 중심이 결정된 레퍼런스 위치선에 가장 가까운 구역을 선택하는 것에 의해 로케이션 공간의 대응하는 구역에 결정된 레퍼런스 위치션을 맵핑시키도록 구현될 수도 있다. 인덱스 선택을 위해 이용된 유사성의 척도 (예를 들어, 로케이션 공간에서의 유클리디안 거리)는 파티셔닝을 위해 이용된 유사성의 척도와 동일하거나 또는 상이할 수도 있다. 분할 스킴이 양자화 코드북에 의해 설명되는 경우에 대해, 이러한 맵핑은 코드북 (예를 들어, 선택된 구역을 나타내는 코드워드)로 인덱스를 선택하는 것을 포함할 수도 있다.

- [0176] 도 24 를 참조하면, 직교 좌표들을 이용하여 오브젝트 포지션에 대한 메타데이터를 발생시키기 위한 방법 (M100) 의 특정 애플리케이션이 도시된다. 도 24 의 실시형태는 프레임-바이-프레임 기반으로 다수의 오브젝트들의 결정된 포지션들을 인코딩하는데 이용될 수도 있는 코드북 (2102) 을 예시한다. 결정된 포지션들은 로케이션 공간 (2106) 에서의 각각의 오브젝트의 포지션들을 나타낼 수도 있다. 코드북 (2102) 은 동일하지 않은 사이즈의 구역들을 포함하여, 로케이션 공간의 별개의 구역들로의 분할 (2108) 에서의 대응하는 구역들 (예를 들어, 2142, 2144, 2146) 에 로케이션 공간 (2106) 에서의 포지션들 (예를 들어, 2132, 2134, 2136) 을 맵핑한다.
- [0177] 코드북 (2102) 은 다수의 프레임들 또는 다른 이미지들 (2104) 에 대한 메타데이터를 생성하는데 이용될 수도 있다. 이러한 메타데이터는 비디오 스트림에 통합될 수도 있고 및/또는 별도로 스트리밍되고 및/또는 별개의 파일에 저장될 수도 있다. 결정된 포지션들은 프레임에서의 오브젝트들에 대한 직교 좌표 데이터 (예를 들어, x-좌표 및 y-좌표) 를 포함할 수도 있다. 특정 실시형태에서, 결정된 포지션들은 프레임에서의 오브젝트들에 대한 3 차원 좌표 데이터 (예를 들어, 3 차원 데카르트 좌표 시스템의 z-좌표) 를 또한 포함할 수도 있다. 예시된 실시형태에서, 코드북 (2102) 은 10,000 개의 프레임들 (또는 다른 이미지들) 에 걸쳐 3 개의 오브젝트들에 대한 결정된 포지션들을 클러스터링하는 것에 의해 발생될 수도 있다. 이들 결정된 포지션들은 또한, 메타데이터로서 대응하는 프레임들과 연관되었을 수도 있다. 특정 실시형태에서, 제 1 프레임은 도 20 의 제 1 이미지 (예를 들어, 디스플레이 (1710)) 에 대응할 수도 있고, 제 2 프레임은 도 20 의 제 2 이미지 (예를 들어, 디스플레이 (1730)) 에 대응할 수도 있고, 제 3 프레임은 도 20 의 제 3 이미지 (예를 들어, 디스플레이 (1750)) 에 대응할 수도 있다. 이 예에서 10,000 개의 프레임들에 대한 결정된 포지션들은 코드북 (2102) 을 획득하는데 이용되지만, 다른 실시형태들에서, 코드북 (2102) 은 더 적은 (또는 부가적인) 프레임들에 대한 결정된 포지션들에 기초할 수도 있다. 대안적으로 또는 부가적으로, 코드북은 상술된 바와 같이 (예를 들어, 도 19b 내지 도 19d 에 도시한 바와 같이) 관심-기반 분할 스킴, 또는 관측된 결정된 포지션들의 세트를 클러스터링하는 것에 의해 획득되지 않은 다른 분할 스킴에 기초할 수도 있다. 부가적으로, 결정된 포지션들은 더 적은 (또는 부가적인) 오브젝트들에 대한 포지션 정보를 포함할 수도 있다. 도 24 의 다음의 설명은 제 3 이미지에 대응하는 결정된 포지션들에 기초한다. 그러한, 유사한 기법들이 다른 결정된 포지션들에 적용가능할 수도 있다.
- [0178] 제 3 이미지에서의 결정된 포지션들은 제 1 직교 좌표 데이터 (2132), 제 2 직교 좌표 데이터 (2134), 및 제 3 직교 좌표 데이터 (2136) 를 포함할 수도 있다. 제 1 직교 좌표 데이터 (2132) 는 도 20 의 제 3 이미지에서의 제 1 오브젝트 (1702) 의 x-좌표 포지션 및 제 3 이미지에서의 제 1 오브젝트 (1702) 의 y-좌표 포지션을 포함할 수도 있다. 제 2 직교 좌표 데이터 (2134) 는 제 3 이미지에서의 제 2 오브젝트 (1704) 의 x-좌표 포지션 및 제 3 이미지에서의 제 2 오브젝트 (1704) 의 y-좌표 포지션을 포함할 수도 있다. 제 3 직교 좌표 데이터 (2136) 는 제 3 이미지에서의 제 3 오브젝트 (1706) 의 x-좌표 포지션 및 제 3 이미지에서의 제 3 오브젝트 (1706) 의 y-좌표 포지션을 포함할 수도 있다.
- [0179] 다른 특정 실시형태에서, 직교 좌표 데이터 (2132 내지 2136) 에서의 각각의 오브젝트 (1702 내지 1706) 의 x-좌표들 및 y-좌표들은 제 3 이미지에서의 오브젝트들 (1702 내지 1706) 의 질량 중심 (C_{M3}) 에 상대적일 수도 있다. 예를 들어, 질량 중심 (C_{M3}) 은 원점으로서 지정될 수도 있고, 각각의 오브젝트 (1702 내지 1706) 의 로케이션 (예를 들어, x-좌표들 및 y-좌표들) 은 원점에 상대적일 수도 있다. 이러한 경우에, 이러한 도 19a 에 도시한 바와 같은 농구 코트 또는 다른 스포츠 필드의 예에서, 도 19b 내지 도 19d 에 도시한 것과 같은 불균일한 분할 스킴에 따라 선택된 오브젝트들 (예를 들어, 플레이어들) 의 이러한 질량 중심의 로케이션을 양자화하는 것이 바람직할 수도 있다.
- [0180] 다른 특정 실시형태에서, 특정 오브젝트는 원점으로서 지정될 수도 있고, 다른 오브젝트들의 로케이션들 (예를 들어, x-좌표들 및 y-좌표들) 은 원점에 상대적일 수도 있다. 특정 오브젝트 ("1 차 관심 오브젝트" 라고도 불림) 는 다음 예들 중 임의의 예를 포함할 수도 있다: 공 (예를 들어, 스포츠 이벤트의 비디오에서), 선택된 상대방 (예를 들어, 스포츠 이벤트에서), 범죄 피해자 (예를 들어, 감시 비디오에서), 신부 (예를 들어, 웨딩 비디오에서). 1 차 관심 오브젝트는 시간의 경과에 따라 이동할 수도 있고 (예를 들어, 사람), 또는 공간에서 고정된 로케이션을 갖는 오브젝트 (예를 들어, 스포츠 비디오에서 네트, 바스켓, 또는 다른 골) 일 수도 있다. 비제한적 예에서, 제 1 오브젝트 (1702) 는 원점으로서 지정될 수도 있고, 제 2 오브젝트 (1704) 및 제 3 오브젝트 (1706) 의 로케이션들은 원점에 상대적일 수도 있다.
- [0181] 프로세서는 동일하지 않은 사이즈의 구역들을 포함하여, 로케이션 공간의 별개의 구역들로의 분할 (2108) 에 따

라 메타데이터를 생성하기 위해 직교 좌표 데이터 (2132 내지 2136) 를 인코딩하도록 코드북 (2102) 을 이용할 수도 있다. 예시하기 위해, 프로세서는 제 1 직교 좌표 데이터 (2132) 를 제 1 인코딩된 로케이션 데이터 (2142) 로 인코딩 (예를 들어, 양자화) 할 수도 있다. 예를 들어, 프로세서는 제 1 직교 좌표 데이터 (2132) 를 대응하는 값 (예를 들어, 코드북 (2102) 의 코드워드) 에 맵핑시키고 제 1 직교 좌표 데이터 (2132) 를 값 (예를 들어, 제 1 인코딩된 로케이션 데이터 (2142)) 으로서 인코딩할 수도 있다. 부가적으로, 프로세서는 제 2 직교 좌표 데이터 (2134) 를 제 2 인코딩된 로케이션 데이터 (2144) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 코드북 (2102) 을 이용하여 제 2 직교 좌표 데이터 (2134) 를 대응하는 값에 맵핑시키고 제 2 직교 좌표 데이터 (2134) 를 값 (예를 들어, 제 2 인코딩된 로케이션 데이터 (2144)) 으로서 인코딩할 수도 있다. 유사한 방식으로, 프로세서는 제 3 직교 좌표 데이터 (2136) 를 제 3 인코딩된 로케이션 데이터 (2146) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 코드북 (2102) 을 이용하여 제 3 직교 좌표 데이터 (2136) 를 대응하는 값에 맵핑시키고 제 3 직교 좌표 데이터 (2136) 를 값 (예를 들어, 제 3 인코딩된 로케이션 데이터 (2146)) 으로서 인코딩할 수도 있다.

[0182] 제 1 인코딩된 로케이션 데이터 (2142) 에 기초하여, 도 20 의 제 3 이미지에서의 제 1 오브젝트 (1702) (예를 들어, 오브젝트 1) 의 포지션은 생성된 메타데이터에서 그리드 (2108) 에서의 대응하는 로케이션 (예를 들어, 나타난 구역의 중심) 으로서 표현될 수도 있다. 제 2 인코딩된 로케이션 데이터 (2144) 에 기초하여, 제 3 이미지에서의 제 2 오브젝트 (1704) (예를 들어, 오브젝트 2) 의 로케이션은 생성된 메타데이터에서 그리드 (2108) 에서의 대응하는 로케이션으로서 표현될 수도 있다. 제 3 인코딩된 로케이션 데이터 (2146) 에 기초하여, 제 3 이미지에서의 제 3 오브젝트 (1706) (예를 들어, 오브젝트 3) 의 로케이션은 생성된 메타데이터에서 그리드 (2108) 에서의 대응하는 로케이션으로서 표현될 수도 있다.

[0183] 상기 언급한 바와 같이, 결정된 포지션은 다수의 선택된 오브젝트들의 기하학적 배열 또는 "형성 (formation)" 의 레퍼런스 포지션일 수도 있다. 하나 이상의 로케이션 (또는 "레퍼런스") 코드북들을 이용하는 것에 더하여 또는 그에 대한 대안으로, 태스크 T300 은 레퍼런스 포지션 (예를 들어, 오브젝트들의 질량 중심의 로케이션) 에 상대적인 공간에서의 선택된 오브젝트들 (예를 들어, 그들의 기하학적 배열, 또는 이 배열을 설명하는 형상) 의 형성을 인코딩하는 하나 이상의 형성 (또는 "상대적인") 코드북들을 이용하도록 구현될 수도 있다. 3 개의 선택된 오브젝트들의 비제한적 예에서, 레퍼런스 포지션으로서, 2-D 공간에서 선택된 오브젝트들의 질량 중심의 로케이션을 저장하고 레퍼런스 포지션에 상대적으로, 총 8 개의 자유도들에 대해, 2-D 공간에서 각각의 선택된 오브젝트의 로케이션을 나타내는 메타데이터를 생성하는 것이 바람직할 수도 있다. 이들 값들은 단일 코드워드로서 (예를 들어, 8-D 공간에서) 양자화될 수도 있다. 대안적으로, 이들 값들은 상술된 하나 이상의 로케이션 코드북들을 이용하여 레퍼런스 포지션으로서 (예를 들어, 2-D 공간에서), 그리고 하나 이상의 형성 코드북들 (즉, 후보 기하학적 배열들의 인덱싱된 세트들) 을 이용하여 오브젝트 형성으로서 (예를 들어, 6-D 공간에서) 별도로 양자화될 수도 있다.

[0184] 대안적으로, 형성 또는 상대적 코드북으로의 인덱스는 템플릿들 (즉, 후보 기하학적 배열들) 의 세트의 하나를 식별할 수도 있다. 이러한 경우에, 나타내진 템플릿에 대한 수정 (예를 들어, 스케일링, 회전 또는 다른 배향, 에스펙트비 등) 이 별도로 양자화될 수도 있다. 도 33a 는 3 개의 템플릿들의 예를 도시하고, 도 33b 는 도 33a 의 가장 왼쪽의 템플릿에 대한 수정들의 3 개의 예를 도시한다.

[0185] 형성 코드북은 예를 들어, 가능한 형성들 중에서의 발생 빈도의 차이들; 가능한 형성들 중에서의 관심 레벨의 차이들 (예를 들어, 결정된 형성에 대한 유사성; 가능한 형성의 특정 미리 결정된 이벤트 (예를 들어, 스코어링 이벤트) 또는 패턴 (예를 들어, 풋볼 팀의 특정 시작 형성) 에 대한 상이한 연관성의 정도들; 공간 레퍼런스로부터의 가장 먼 선택된 오브젝트의 거리; 선택된 오브젝트들의 선택된 서브세트의 서로로부터의 거리; 및/또는 선택된 오브젝트들 중 임의의 2 개 사이의 최대 거리에 따라 불균일하게 형성 공간 (즉, 모든 가능한 형성들의 공간의 일부 부분) 을 분할하도록 구성될 수도 있다.

[0186] 코드워드들 (예를 들어, 인코딩된 로케이션 데이터 (2142 내지 2146)) 은 고정-폭 코드워드들 또는 가변-폭 코드워드들일 수도 있다. 특정 실시형태에서, 프로세서는 오브젝트의 로케이션에 기초하여 (예를 들어, 오브젝트가 "고밀도 영역" 에 있는지 또는 "저밀도 영역" 에 있는지 여부에 기초하여) 가변-폭 코드워드들을 이용할 수도 있다. 고밀도 영역들은 오브젝트들 (1702 내지 1706) 중 적어도 하나에 의해 빈번하게 점유되는 도 20 의 이미지들의 영역들에 대응할 수도 있다. 메타데이터에서의 비트들의 수를 감소시키기 위해, 로케이션 공간 (2106) 에서의 고밀도 영역들에 대응하는 코드워드들은 저밀도 영역들에 대응하는 코드워드들보다 더 적은 비트들 (즉, 더 짧은 코드워드) 을 가질 수도 있다 (즉, 덜 빈번하게 점유된 그리드 로케이션들은 더 긴 코드워드들을 갖는다). 대안적으로, 또는 또한, 불균일한 분할 스킴 (2108) 은 고밀도 구역들에서 더 많은 서브-

구역들 (즉, 더 높은 해상도를 제공하기 위해 더 많은 별개의 코드워드들) 및 저밀도 구역들에서 더 적은 서브-구역들 (즉, 더 낮은 해상도를 제공하는, 더 적은 별개의 코드워드들) 을 갖는 계층적 스킴일 수도 있다.

[0187] 오브젝트가 고밀도 영역에 있으면, 오브젝트에 대한 코드워드는 오브젝트의 보다 "정밀한" 로케이션을 발생시키기 위해 상대적으로 작은 구역을 표현할 수도 있다. 예를 들어, 증가된 수의 코드워드들은 프레임-바이-프레임 기반으로 보다 정확한 추적을 획득하기 위해 고밀도 영역들에서 오브젝트들 (1702 내지 1706) 의 로케이션을 추적하는데 이용될 수도 있다. 오브젝트가 저밀도 영역에 있으면, 오브젝트에 대한 코드워드는 상대적으로 큰 구역을 표현할 수도 있다. 특정 실시형태에서, 고밀도 영역들 및 저밀도 영역들은 이력 데이터의 분석 (예를 들어, 스크린의 어느 영역들이 오브젝트들에 의해 빈번하게 점유되는 이력을 갖는지를 분석) 에 기초하여 결정될 수도 있다. 다른 특정 실시형태에서, 고밀도 영역들 및 저밀도 영역들은 미리 결정될 수도 있다.

[0188] 형성 코드북으로의 코드워드는 n 개의 선택된 오브젝트들의 n 개의 포지션들에 대한 고유한 맵핑들 간을 구별할 수도 있다 (예를 들어, 각각의 이러한 맵핑에 대한 고유한 코드워드를 포함할 수도 있다). 예를 들어, 도 34a 는 고유한 코드워드로서 각각 인코딩될 수도 있는 (즉, 상이한 각각의 후보 기하학적 배열들에 의해 표현될 수도 있는) 3 개의 선택된 오브젝트들의 4 개의 상이한 형성들의 예들을 도시한다. 이들 예들에서, 선택된 오브젝트의 각각의 포지션은 그 오브젝트를 고유하게 식별하는 아이콘 (예를 들어, 원, 정사각형, 또는 삼각형) 에 의해 나타내진다. 대안적으로, 형성 코드북으로의 코드워드는 여러 상이한 고유한 맵핑들을 인코딩할 수도 있다. 예를 들어, 도 34b 는 도 34a 의 형성들에 대응하는 불균일한 맵핑들의 예들을 도시하고, 여기서 임의의 선택된 오브젝트의 로케이션은 x 로 나타내진다. 이 경우에, 도 34c 는 단일 불균일한 맵핑 (즉, 도 34b 의 가장 왼쪽에 도시되는 후보 기하학적 배열) 으로 나타내질 수도 있는 선택된 오브젝트들의 6 개의 고유한 형성들을 도시한다.

[0189] 코드북 (2102) 은 고정 코드북 또는 적응적 코드북일 수도 있다. 적응적 코드북은 오브젝트 로케이션들을 대표하는 이력 데이터에 기초하여 코드워드들의 길이를 조정할 수도 있다 (예를 들어, 가변-폭 코드워드들을 발생시킬 수도 있다). 예를 들어, 적응적 코드북은 스크린의 빈번하게 점유된 영역에서의 오브젝트들에 대해 더 짧은 코드워드들을 발생시킬 수도 있고 이력 데이터에 기초하여 스크린의 덜 빈번하게 점유된 영역에서의 오브젝트들에 대해 더 긴 코드워드들을 발생시킬 수도 있다. 그에 반해서, 고정 코드북은 이력 데이터에 기초하여 변화하지 않을 수도 있고, 고정-폭 코드워드들 또는 가변-폭 코드워드들을 이용할 수도 있다.

[0190] 각각의 레퍼런스 포지션들에 대한 가능한 형성들 중에서 관심 레벨들 또는 발생 빈도들의 차이들을 활용하는 다수의 형성 코드북들을 이용하여 형성 데이터를 인코딩하는 것이 바람직할 수도 있다. 예를 들어, 가변 사이즈의 형성 코드북들의 세트 중의 하나는 레퍼런스 포지션의 코드워드에 의존하여 (예를 들어, 레퍼런스 포지션이 센터 코트에 있을 때의 더 작은 코드북 대 레퍼런스 포지션이 골 부근에 있을 때의 더 큰 코드북) 선택될 수도 있다.

[0191] 도 24 의 실시형태는 사용자가 코드북 (2102) 을 이용하여 태스크 T300 에 의해 생성된 메타데이터에 기초하여, 비디오의 특정 부분들과 같은 특정 이미지들을 추출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 예를 들어, 로케이션 공간 (2106) 은 사용자가 상이한 오브젝트들에 대해 상이한 로케이션들 (예를 들어, 파라미터들) 을 선택하는 것을 가능하게 하는 사용자-상호작용 그리드로서 스크린 상에 디스플레이될 수도 있다. 그 선택에 기초하여, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기 (fast-forward), 되감기 등) 해야 하는 일 없이 유사한 오브젝트 로케이션 구성들을 갖는 비디오의 원하는 부분들 (예를 들어, 프레임들) 을 쉽게 로케이팅할 수도 있다. 예를 들어, 방법 (M100) 은 분할 스킴에서 구역들 중 하나를 특정하는 요청을 검색하기 위해 생성된 메타데이터를 비교하는 결과에 기초하여 (예를 들어, 스트림으로부터) 하나 이상의 이미지들을 추출하는 태스크를 포함하도록 구현될 수도 있다. 특정 실시형태에서, 로케이션 공간 (2106) 의 이러한 디스플레이는 도 21 의 클러스터들의 스크린 (1802) 과 실질적으로 유사한 방식으로 이용될 수도 있다. 부가적으로, 생성된 메타데이터 (2142 내지 2146) 는 인코더로부터 디코더에 압축 및 송신될 수도 있다. 디코더의 메타데이터 (2142 내지 2146) 의 압축 및 송신은 디코더가 비교적 낮은 양의 데이터 (예를 들어, 양자화된 값들) 를 이용하여 추적된 오브젝트들의 포지션을 결정하는 것을 가능하게 할 수도 있다. 예를 들어, 디코더는 코드북 (2102) 에서의 양자화된 값들에 기초하여 오브젝트들 (1702 내지 1706) 과 연관된 비디오를 디코딩하기 위해 렌더링 동작들을 수행할 수도 있다.

[0192] 하나의 예에서, 추출된 프레임들은 풋볼 게임 또는 다른 스포츠 이벤트에서, 특정 플레이 또는 플레이어들의 세트, 또는 특정 시작 형성 동안만 발생하는 선택된 플레이어들 간의 선택된 형성을 포함한다. 다른

예에서, 추출된 프레임들은 랩 트랙 상에 근접하여 특정 레이싱 카들 간에 선택된 형성을 포함한다. 이러한 선택적 추출에 의해 구동된 비디오 분석은 또한 특정 형성의 발생과의 이벤트 (이는 이전에 미식별될 수도 있음)의 상관을 검출하는데 이용될 수도 있다. 예를 들어, 잠재적인 애플리케이션들은 어셈블리 (예를 들어, 입법 심의회, 가두데모, 칠폐회모임 회칙) 내에서 사람들을 추적하는 것, 및 (예를 들어, 의사 결정자들, 말쑥꾼들, 촉진자들을 식별하기 위해) 특정된 이벤트들과 선택된 사람들의 근접성을 상관시키는 것을 포함한다.

본 명세서에서 설명한 바와 같은 방법 (M100)의 구현들은 또한 감금된 (예를 들어, 실험실 또는 동물원의) 또는 야생의 동물들 간의 소셜 상호작용의 분석에 적용될 수도 있다.

[0193] 도 25를 참조하면, 극 좌표들을 이용하여 오브젝트 포지션에 대한 메타데이터를 발생시키기 위한 방법 (M100)의 다른 특정 적용이 도시된다. 도 25의 실시형태는 프레임-바이-프레임 기반으로 다수의 오브젝트들의 결정된 포지션들을 인코딩하는데 이용될 수도 있는 코드북 (2202)을 예시한다. 결정된 포지션들은 극성 로케이션 공간 (2206)에 각각의 오브젝트의 포지션들을 나타낼 수도 있다. 코드북 (2202)은 로케이션 공간 (2206)에서의 포지션들 (예를 들어, 2232, 2234, 2236)을 동일하지 않은 사이즈의 구역들을 포함한, 별개의 구역들로의 로케이션 공간 (2206)의 분할 (2208)에서의 대응하는 구역들 (예를 들어, 2242, 2244, 2246)에 맵핑한다.

[0194] 코드북 (2202)은 다수의 프레임들 또는 다른 이미지들 (2204)에 대한 메타데이터를 생성하는데 이용될 수도 있다. 이러한 메타데이터는 비디오 스트림에 통합될 수도 있고 및/또는 별도로 스트리밍되고 및/또는 별개의 파일에 저장될 수도 있다. 결정된 포지션들은 프레임에서의 오브젝트들에 대한 극 좌표 데이터를 포함할 수도 있다. 특정 실시형태에서, 결정된 포지션들은 또한 프레임에서의 오브젝트들에 대한 구형 또는 원통형 좌표 데이터 (예를 들어, 3D 좌표 데이터)를 포함할 수도 있다. 예시된 실시형태에서, 코드북 (2102)은 10,000개의 프레임들 (또는 다른 이미지들)에 걸쳐 3개의 오브젝트들의 결정된 포지션들을 클러스터링하는 것에 의해 발생될 수도 있다. 이들 결정된 포지션들은 또한 메타데이터로서 대응하는 프레임들과 연관되었을 수도 있다. 특정 실시형태에서, 제 1 프레임은 도 20의 제 1 이미지 (예를 들어, 디스플레이 (1710))에 대응할 수도 있고, 제 2 프레임은 도 20의 제 2 이미지 (예를 들어, 디스플레이 (1730))에 대응할 수도 있고, 제 3 프레임은 도 20의 제 3 이미지 (예를 들어, 디스플레이 (1750))에 대응할 수도 있다. 이 예에서, 10,000개의 프레임들에 대한 결정된 포지션들은 코드북 (2202)을 획득하는데 이용되지만, 다른 실시형태들에서, 코드북 (2202)은 더 적은 (또는 부가적인) 프레임들에 대한 결정된 포지션들에 기초할 수도 있다.

대안적으로 또는 부가적으로, 코드북은 상술한 바와 같은 (예를 들어, 도 19b 내지 도 19d에 도시한 바와 같은) 관심-기반 분할 스킴, 또는 관측된 결정된 포지션들의 세트를 클러스터링하는 것에 의해 획득되지 않는 다른 분할 스킴에 기초할 수도 있다. 부가적으로, 결정된 포지션들은 더 적은 (또는 부가적인) 오브젝트들에 대한 포지션 정보를 포함할 수도 있다. 도 25의 다음의 설명은 제 3 이미지에 대응하는 결정된 포지션들에 기초한다. 그러나, 유사한 기법들이 다른 결정된 포지션들에 적용가능할 수도 있다.

[0195] 제 3 이미지에서의 결정된 포지션들은 극성 (polar) 배향 데이터 (2230), 제 1 극 좌표 데이터 (2232), 제 2 극 좌표 데이터 (2234), 및 제 3 극 좌표 데이터 (2236)를 포함할 수도 있다. 극성 배향 데이터 (2230)는 질량 중심 (C_{M3})의 포지션 (예를 들어, 도 20의 제 3 이미지에서의 질량 중심 (C_{M3})의 x-좌표 (C_x) 및 제 3 이미지에서의 질량 중심 (C_{M3})의 y-좌표 (C_y))을 나타낼 수도 있다. 극성 배향 데이터 (2230)는 또한 중심으로서 질량 중심 (C_{M3})을 갖는 원의 반경 (R)을 나타낼 수도 있다. 스케일 팩터로서 구현될 수도 있는 반경 (R)은 각각의 추적된 오브젝트를 포괄할 정도로 충분히 클 수도 있다. 예시적 예는 극성 로케이션 공간 (2206)에 대하여 나타내진다. 하나의 예에서, 태스크 T300은 로케이션 공간 (2206) (예를 들어, 도 24에 도시한 바와 같은 직교 좌표 공간)과는 상이할 수도 있는, 배향 포지션이 나타내는 대응하는 로케이션 공간을 분할하는, 동일하지 않은 사이즈의 구역들을 포함하는, 복수의 구역들 중 하나를 나타내는 메타데이터로서 배향 포지션 (예를 들어, 질량 중심의 포지션)을 인코딩하도록 구현된다.

[0196] 제 1 극 좌표 데이터 (2232)는 제 1 오브젝트 (1702) (오브젝트 1)의 제 1 극 좌표들을 나타낼 수도 있다. 제 1 극 좌표들은 질량 중심 (C_{M3})으로부터의 제 1 오브젝트 (1702)의 제 1 거리 (a) 및 0도 레퍼런스로부터 측정된 제 1 오브젝트 (1702)의 제 1 각도 (θ_a) (예를 들어, 수평으로부터의 각도 변위)를 포함할 수도 있다. 제 2 극 좌표 데이터 (2234)는 제 2 오브젝트 (1704) (오브젝트 2)의 제 2 극 좌표들을 나타낼 수도 있다. 제 2 극 좌표들은 질량 중심 (C_{M3})으로부터의 제 2 오브젝트 (1704)의 제 2 거리 (b) 및 수평으로부터 측정된 제 2 오브젝트 (1704)의 제 2 각도 (θ_b)를 포함할 수도 있다. 제 3 극 좌표 데이터 (2236)는 제 3 오브젝트 (1706) (오브젝트 3)의 제 3 극 좌표들을 나타낼 수도 있다. 제 3 극 좌표들은

질량 중심 (C_{M3}) 으로부터의 제 3 오브젝트 (1706) 의 제 3 거리 (c) 및 수평으로부터 측정된 제 1 오브젝트 (1702) 의 제 3 각도 (θ_c) 를 포함할 수도 있다.

[0197] 특정 실시형태에서, 극성 배향 데이터 (2230) 의 반경 (R) 은 질량 중심 (C_{M3}) 으로부터 더 멀리 떨어져 있는 오브젝트의 거리에 기초하여 정규화될 수도 있다. 예를 들어, 반경 (R) 은 제 1 거리 (a) 가 제 2 거리 (b) 및 제 3 거리 (c) 보다 더 크면 제 1 거리 (a) 와 동일할 수도 있다.

[0198] 프로세서는 동일하지 않은 사이즈의 구역들을 포함한 구역들로의 로케이션 공간의 분할 (2108) 에 따라 메타데이터를 생성하기 위해 극 좌표 데이터 (2232 내지 2236) 를 인코딩하도록 코드북 (2202) 을 이용할 수도 있다. 예시하기 위해, 프로세서는 제 1 극 좌표 데이터 (2232) 를 제 1 인코딩된 로케이션 데이터 (2242) 로 인코딩 (예를 들어, 양자화) 할 수도 있다. 예를 들어, 프로세서는 (극성 배향 데이터 (2230) 에 기초하는) 제 1 극 좌표 데이터 (2232) 를 대응하는 값 (예를 들어, 코드북 (2202) 의 코드워드) 에 맵핑시키고 제 1 극 좌표 데이터 (2232) 를 값 (예를 들어, 제 1 인코딩된 로케이션 데이터 (2242)) 으로서 인코딩할 수도 있다. 부가적으로, 프로세서는 제 2 극 좌표 데이터 (2234) 를 제 2 인코딩된 로케이션 데이터 (2244) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 (예를 들어, 극성 배향 데이터 (2230) 에 기초하는) 제 2 극 좌표 데이터 (2234) 를 코드북 (2202) 을 이용하여 대응하는 값에 맵핑시키고 제 2 극 좌표 데이터 (2234) 를 값 (예를 들어, 제 2 인코딩된 로케이션 데이터 (2244)) 으로서 인코딩할 수도 있다. 유사한 방식으로, 프로세서는 제 3 극 좌표 데이터 (2236) 를 제 3 인코딩된 로케이션 데이터 (2246) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 제 3 극 좌표 데이터 (2236) 를 코드북 (2202) 을 이용하여 대응하는 값에 맵핑시키고 제 3 극 좌표 데이터 (2236) 를 값 (예를 들어, 제 3 인코딩된 로케이션 데이터 (2246)) 으로서 인코딩할 수도 있다.

[0199] 제 1 인코딩된 로케이션 데이터 (2242) 에 기초하여, 도 20 의 제 3 이미지에서의 제 1 오브젝트 (1702) (예를 들어, 오브젝트 1) 의 포지션은 생성된 메타데이터에서 극성 그리드 (2208) 에서의 대응하는 로케이션 (예를 들어, 나타내진 구역의 중심) 으로서 표현될 수도 있다. 제 2 인코딩된 로케이션 데이터 (2244) 에 기초하여, 제 3 이미지에서의 제 2 오브젝트 (1704) (예를 들어, 오브젝트 2) 의 로케이션은 생성된 메타데이터에서 극성 그리드 (2208) 에서의 대응하는 로케이션으로서 표현될 수도 있다. 제 3 인코딩된 로케이션 데이터 (2246) 에 기초하여, 제 3 이미지에서의 제 3 오브젝트 (1706) (예를 들어, 오브젝트 3) 의 로케이션은 생성된 메타데이터에서 극성 그리드 (2208) 에서의 대응하는 로케이션으로서 표현될 수도 있다. 예를 들어, 극성 그리드 (2208) 는 그리드 원점으로부터의 방사상 거리의 범위들에 기초하여 (예를 들어, 질량 중심 (C_{M3}) 으로부터의 방사상 거리의 범위들에 기초하여) 그리고 0도 레퍼런스로부터의 각도 변위의 범위들에 기초하여 동일하지 않은 사이즈의 구역들을 포함한 구역들로 분할될 수도 있다.

[0200] 도 25 의 실시형태는 사용자가 코드북 (2202) 을 이용하여 태스크 T300 에 의해 생성된 메타데이터에 기초하여 특정 이미지들 (예를 들어, 비디오의 프레임들) 을 추출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 예를 들어, 극성 로케이션 공간 (2206) 은 사용자가 상이한 오브젝트들에 대해 상이한 로케이션들 (예를 들어, 파라미터들) 을 선택하는 것을 가능하게 하는 사용자-상호작용 그리드로서 스크린 상에 디스플레이될 수도 있다. 그 선택에 기초하여, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기 (fast-forward), 되감기 등) 해야 하는 일 없이 유사한 오브젝션 로케이션 구성들을 갖는 비디오의 원하는 부분들 (예를 들어, 프레임들) 을 쉽게 로케이팅할 수도 있다. 예를 들어, 방법 (M100) 은 분할 스킴에서 구역들 중 하나를 특정하는 검색 요청과 생성된 메타데이터를 비교하는 결과에 기초하여 (예를 들어, 스토리지로부터) 하나 이상의 이미지들을 추출하는 태스크를 포함하도록 구현될 수도 있다. 특정 실시형태에서, 이러한 극성 로케이션 공간 (2206) 의 디스플레이는 도 21 의 클러스터들의 스크린 (1802) 과 실질적으로 유사한 방식으로 이용될 수도 있다. 부가적으로, 생성된 메타데이터 (2242 내지 2246) 는 인코더로부터 디코더에 압축 및 송신될 수도 있다. 생성된 메타데이터 (2242 내지 2246) 의 디코더에의 압축 및 송신은 디코더가 상대적으로 낮은 양의 데이터 (예를 들어, 양자화된 값들) 를 이용하여 추적된 오브젝트들의 포지션을 결정하는 것을 가능하게 할 수도 있다. 예를 들어, 디코더는 코드북 (2202) 에서의 양자화된 값들에 기초하여 오브젝트들 (1702 내지 1706) 과 연관된 비디오를 디코딩하기 위해 렌더링 동작들을 수행할 수도 있다.

[0201] 코드워드들 (예를 들어, 인코딩된 로케이션 데이터 (2242 내지 2246)) 은 고정-폭 코드워드들 또는 가변-폭 코드워드들일 수도 있다. 특정 실시형태에서, 프로세서는 오브젝트의 로케이션에 기초하여 (예를 들어, 오브젝트가 "고밀도 영역" 에 있는지 또는 "저밀도 영역" 에 있는지 여부에 기초하여) 가변-폭 코드워드들을 이용할 수도 있다. 고밀도 영역들은 오브젝트들 (1702 내지 1706) 중 적어도 하나에 의해 빈번하게 점유되는 도 20 의 이미지들의 영역들에 대응할 수도 있다. 메타데이터에서의 비트들의 수를 감소시키기 위해, 극성 로케이

선 공간 (2206) 에서의 고밀도 영역들에 대응하는 코드워드들은 저밀도 영역들에 대응하는 코드워드들보다 더 적은 비트들 (즉, 더 짧은 코드워드) 을 가질 수도 있다 (즉, 덜 빈번하게 점유된 그리드 로케이션들은 더 긴 코드워드들을 갖는다). 대안적으로, 또는 또한, 불균일한 분할 스킴 (2208) 은 고밀도 영역들에서 더 많은 서브-구역들 (즉, 더 높은 해상도를 제공하기 위해 더 많은 별개의 코드워드들) 을 갖고 저밀도 구역들에서 더 적은 서브-구역들 (즉, 더 낮은 해상도를 제공하는 더 적은 별개의 코드워드들) 을 갖는 계층적 스킴일 수도 있다.

[0202] 오브젝트가 고밀도 영역에 있으면, 오브젝트에 대한 코드워드는 오브젝트의 보다 "정밀한" 로케이션을 발생시키기 위해 상대적으로 작은 구역을 표현할 수도 있다. 예를 들어, 증가된 수의 코드워드들은 프레임-바이-프레임 기반으로 보다 정확한 추적을 획득하기 위해 고밀도 영역들에서 오브젝트들 (1702 내지 1706) 의 로케이션을 추적하는데 이용될 수도 있다. 오브젝트가 저밀도 영역에 있으면, 오브젝트에 대한 코드워드는 상대적으로 큰 구역을 표현할 수도 있다. 특정 실시형태에서, 고밀도 영역들 및 저밀도 영역들은 이력 데이터의 분석 (예를 들어, 스크린의 어느 영역들이 오브젝트들에 의해 빈번하게 점유되는 이력을 갖는지를 분석하는 것) 에 기초하여 결정될 수도 있다. 다른 특정 실시형태에서, 고밀도 영역들 및 저밀도 영역들은 미리 결정될 수도 있다.

[0203] 코드북 (2202) 은 고정 코드북 또는 적응적 코드북일 수도 있다. 적응적 코드북은 오브젝트 로케이션들을 대표하는 이력 데이터에 기초하여 코드워드들의 길이를 조정할 수도 있다 (예를 들어, 가변-폭 코드워드들을 발생시킬 수도 있다). 예를 들어, 적응적 코드북은 스크린의 빈번하게 점유된 영역에서의 오브젝트들에 대해 더 짧은 코드워드들을 발생시킬 수도 있고 이력 데이터에 기초하여 스크린의 덜 빈번하게 점유된 영역에서의 오브젝트들에 대해 더 긴 코드워드들을 발생시킬 수도 있다. 그에 반해서, 고정 코드북은 이력 데이터에 기초하여 변화하지 않을 수도 있고, 고정-폭 코드워드들 또는 가변-폭 코드워드들을 이용할 수도 있다.

[0204] 도 26 을 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법 (2300) 의 특정 실시형태를 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법 (2300) 은 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 또는 이들의 임의의 조합을 이용하여 수행될 수도 있다.

[0205] 방법 (2300) 은 2302 에서, 모바일 디바이스에서, 특정 프레임에서의 제 1 오브젝트의 제 1 로케이션 및 특정 프레임에서의 제 2 오브젝트의 제 2 로케이션을 추적하는 단계를 포함할 수도 있다. 예를 들어, 도 20 을 참조하면, 제 1 오브젝트 (1702) 의 제 1 로케이션은 제 3 프레임에서 추적될 수도 있고 제 2 오브젝트 (1704) 의 제 2 로케이션은 제 3 프레임에서 추적될 수도 있다. 부가적으로, 제 3 오브젝트 (1706) 의 제 3 로케이션은 제 3 프레임에서 추적될 수도 있다.

[0206] 2304 에서, 제 1 로케이션과 연관된 제 1 좌표 데이터가 발생될 수도 있고 제 2 로케이션과 연관된 제 2 좌표 데이터가 발생될 수도 있다. 예를 들어, 도 24 를 참조하면, 프로세서는 제 3 프레임에서의 제 1 오브젝트 (1702) 의 포지션과 연관된 제 1 직교 좌표 데이터 (2132) 및 제 3 프레임에서의 제 2 오브젝트 (1704) 의 포지션과 연관된 제 2 직교 좌표 데이터 (2134) 를 발생시킬 수도 있다. 부가적으로, 프로세서는 제 3 프레임에서의 제 3 오브젝트 (1706) 의 포지션과 연관된 제 3 직교 좌표 데이터 (2136) 를 발생시킬 수도 있다.

[0207] 다른 예로서, 도 25 를 참조하면, 프로세서는 제 3 프레임에서의 제 1 오브젝트 (1702) 의 포지션과 연관된 제 1 극 좌표 데이터 (2232) 및 제 3 프레임에서의 제 2 오브젝트 (1704) 의 포지션과 연관된 제 2 극 좌표 데이터 (2234) 를 발생시킬 수도 있다. 부가적으로, 프로세서는 제 3 프레임에서의 제 3 오브젝트 (1706) 의 포지션과 연관된 제 3 극 좌표 데이터 (2236) 를 발생시킬 수도 있다.

[0208] 2306 에서, 제 1 좌표 데이터 및 제 2 좌표 데이터가 제 1 인코딩된 로케이션 데이터 및 제 2 인코딩된 로케이션 데이터를 생성하기 위해 코드북을 이용하여 인코딩될 수도 있다. 예를 들어, 도 24 를 참조하면, 프로세서는 제 1 인코딩된 로케이션 데이터 (2142) (코드북 (2102) 의 코드워드) 를 발생시키기 위해 제 1 직교 좌표 데이터 (2132) 를 인코딩하도록 코드북 (2102) 을 이용할 수도 있고, 프로세서는 제 2 인코딩된 로케이션 데이터 (2144) 를 발생시키기 위해 제 2 직교 좌표 데이터 (2134) 를 인코딩하도록 코드북 (2102) 을 이용할 수도 있다. 부가적으로, 프로세서는 제 3 인코딩된 로케이션 데이터 (2146) 를 발생시키기 위해 제 3 직교 좌표 데이터 (2132) 를 인코딩하도록 코드북 (2102) 을 이용할 수도 있다.

[0209] 다른 예로서, 도 25 를 참조하면, 프로세서는 제 1 인코딩된 로케이션 데이터 (2242) (예를 들어, 코드북

(2202) 의 코드워드) 로서 메타데이터를 생성하기 위해 제 1 극 좌표 데이터 (2232) 를 인코딩하도록 코드북 (2202) 을 이용할 수도 있고, 프로세서는 제 2 인코딩된 로케이션 데이터 (2244) 로서 메타데이터를 생성하기 위해 제 2 극 좌표 데이터 (2234) 를 인코딩하도록 코드북 (2202) 을 이용할 수도 있다. 부가적으로, 프로세서는 제 3 인코딩된 로케이션 데이터 (2246) 로서 메타데이터를 생성하기 위해 제 3 극 좌표 데이터 (2232) 를 인코딩하도록 코드북 (2202) 을 이용할 수도 있다.

[0210] 2308 에서, 메타데이터 (예를 들어, 제 1 인코딩된 로케이션 데이터 및 제 2 인코딩된 로케이션 데이터) 는 입력 파라미터들에 기초하여 특정 프레임의 추출을 가능하게 하도록 저장될 수도 있다. 예를 들어, 도 24 를 참조하면, 프로세서는 메모리에 제 1 인코딩된 로케이션 데이터 (2142) 및 제 2 인코딩된 로케이션 데이터 (2144) 를 저장할 수도 있다. 모바일 디바이스의 사용자는 파라미터들 (예를 들어, 제 1 오브젝트 (1702) 및 제 2 오브젝트 (1704) 의 원하는 로케이션) 을 입력할 수도 있다. 입력 파라미터들에 기초하여, 모바일 디바이스는 제 1 로케이션 및 제 2 로케이션에 대응하는 입력 파라미터들에 응답하여 재생을 위한 제 3 프레임을 추출할 수도 있다.

[0211] 다른 예로서, 도 25 를 참조하면, 프로세서는 메모리에 제 1 인코딩된 로케이션 데이터 (2242) 및 제 2 인코딩된 로케이션 데이터 (2244) 를 저장할 수도 있다. 모바일 디바이스의 사용자는 파라미터들 (예를 들어, 제 1 오브젝트 (1702) 및 제 2 오브젝트 (1704) 의 원하는 로케이션) 을 입력할 수도 있다. 입력 파라미터들에 기초하여, 모바일 디바이스는 제 1 로케이션 및 제 2 로케이션에 대응하는 입력 파라미터들에 응답하여 재생을 위한 제 3 프레임을 추출할 수도 있다.

[0212] 특정 실시형태에서, 방법 (2300) 은 또한 복수의 프레임들에 걸친 장면의 제 1 영역의 제 1 밀도를 추적하는 단계를 포함할 수도 있다. 제 1 밀도는 제 1 오브젝트 (1702) 또는 제 2 오브젝트 (1704) 가 제 1 영역에 있는 횡수에 기초할 수도 있다. 방법 (2300) 은 또한 복수의 프레임들에 걸친 장면의 제 2 영역의 제 2 밀도를 추적하는 단계를 포함할 수도 있다. 제 2 밀도는 제 1 오브젝트 (1702) 또는 제 2 오브젝트 (1704) 가 제 2 영역에 있는 횡수에 기초할 수도 있다. 방법 (2300) 은 제 1 밀도와 제 2 밀도를 비교하는 단계 및 제 1 밀도가 제 2 밀도보다 더 크면 제 1 영역에 코드북 (2102) 에서의 제 1 양의 메타데이터를 할당하고 제 2 영역에 코드북 (2102) 에서의 제 2 양의 메타데이터를 할당하는 단계를 더 포함할 수도 있다. 제 1 양의 메타데이터는 제 2 양의 메타데이터보다 더 클 수도 있다. 예를 들어, 코드북 (2102) 에서의 더 큰 수의 코드워드들이 제 1 영역 (예를 들어, 더 높은 해상도 영역) 에 대응할 수도 있고 더 작은 수의 코드워드들은 제 2 영역 (예를 들어, 더 낮은 해상도 영역) 에 대응할 수도 있어 덜 밀집한 영역들에 대한 코드워드들의 수를 감소시킬 수도 있다.

[0213] 도 26 의 방법 (2300) 은 사용자가 코드북 (2102) 을 이용하여 인코딩된 비디오의 특정 부분들 (프레임들) 을 추출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 예를 들어, 그리드 (2106) 는 사용자가 상이한 오브젝트들에 대해 상이한 로케이션들 (예를 들어, 파라미터들) 을 선택하는 것을 가능하게 하는 스크린 상에 디스플레이된 사용자-상호작용 그리드일 수도 있다. 그 선택에 기초하여, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기 (fast-forward), 뒤감기 등) 해야 하는 일 없이 유사한 오브젝션 로케이션 구성들을 갖는 비디오의 원하는 부분들을 쉽게 로케이팅할 수도 있다.

[0214] 도 27 을 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법 (2600) 의 특정 실시형태를 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법 (2600) 은 로봇을 이용하여 수행될 수도 있다.

[0215] 방법 (2600) 은, 2602 에서, 로봇에서, 다수의 오브젝트들 사이의 지오메트리를 추적하는 단계를 포함할 수도 있다. 예를 들어, 사용자는 추적될 특정 오브젝트들을 로봇에게 통신할 수도 있다. 특정 실시형태에서, 로봇은 디스플레이 인터페이스를 포함할 수도 있고, 사용자는 디스플레이 인터페이스와의 상호작용들을 통해 로봇에게 통신할 수도 있다. 다른 예로서, 로봇은 환경 팩터들에 기초하여 어떤 오브젝트들이 추적되어야 하는지를 독립적으로 결정할 수도 있다. 예를 들어, 로봇은 모션을 검출하기 위한 센서를 포함할 수도 있다. 검출된 모션에 기초하여, 로봇은 모션과 연관된 오브젝트들을 추적하도록 선택할 수도 있다.

[0216] 2604 에서, 프로세싱 기능이 지오메트리에 기초하여 수행될 수도 있다. 예를 들어, 로봇은 지오메트리에 기초하여 오브젝트들 사이의 관계들을 결정할 수도 있다. 다른 예로서, 로봇은 오브젝트들의 지오메트리에 기초하여 설정 (예를 들어, 장면) 을 결정할 수도 있다.

[0217] 도 28 을 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 특정 실시형태가 도시된다. 도 28 에 나타내진 실시형태에 대하여 설명된 비디오 프로세싱 기법들은 도 1 의 전자 디바이스 (102), 도 2b 의

프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 또는 이들의 임의의 조합에 의해 수행될 수도 있다.

[0218] 도 28 은 제 1 시간 인스턴스에서 캡처되는 장면을 나타내는 이미지 (2700) 를 도시한다. 예를 들어, 이미지 (2700) 는 모바일 디바이스의 스크린 상에 디스플레이될 비디오 스트림의 프레임에 대응할 수도 있다. 모바일 디바이스는 (예를 들어, 모바일 디바이스의 카메라로) 스트림을 캡처하거나 또는 다른 디바이스로부터 스트림을 수신하도록 구성될 수도 있다. 프레임 (2400) 에 나타난 장면은 제 1 오브젝트 (2702), 제 2 오브젝트 (2704), 제 3 오브젝트 (2706), 별, 및 나무를 포함한다. 예시된 실시형태에서, 제 1 오브젝트 (2702) 는 제 1 사람에 대응할 수도 있고, 제 2 오브젝트 (2704) 는 제 2 사람에 대응할 수도 있으며, 제 3 오브젝트 (2706) 는 제 3 사람에 대응할 수도 있다. 다른 실시형태에서, 오브젝트들 (2702 내지 2706) 중 적어도 하나는 모바일 디바이스를 통해 제어되는 로봇에 대응할 수도 있다.

[0219] 도 28 은 또한 모바일 디바이스의 스크린 상의 프레임 (2700) 의 디스플레이 (2710) 를 도시한다. 모바일 디바이스의 스크린 (예를 들어, 터치스크린) 은 뷰파인더에 대응할 수도 있다. 디스플레이 (2710) 는 사용자가 (예를 들어, 태스크 T200 에 의한) 추적을 위해 오브젝트들을 선택하는 것을 가능하게 하는 사용자 선택 메커니즘 (2711) 을 포함할 수도 있다. 예를 들어, 사용자는 스크린 (2710) 상에서 사용자 선택 메커니즘 (2711) 을 터치하고 제 1 오브젝트 (2702) 주위로 제 1 포커스 링 (2712) 을 드래그하여 제 1 오브젝트 (2702) 의 추적을 가능하게 할 수도 있다. 이러한 제 1 포커스 링 (2712) 의 배치에 응답하여, 모바일 디바이스는 추적하기 위한 제 1 오브젝트 (2702) 를 선택할 수도 있다. 유사한 방식으로, 사용자는 또한 스크린 상에서 사용자 선택 메커니즘 (2711) 을 터치하고 제 2 오브젝트 (2704) 및 제 3 오브젝트 (2706) 주위로 제 2 포커스 링 (2714) 및 제 3 포커스 링 (2716) 을 드래그하여, 제 2 오브젝트 (2704) 및 제 3 오브젝트 (2706) 각각의 추적을 가능하게 할 수도 있다. 이러한 제 2 포커스 링 (2714) 및 제 3 포커스 링 (2716) 의 배치에 응답하여, 모바일 디바이스는 추적하기 위한 제 2 오브젝트 (2704) 및 제 3 오브젝트 (2706) 를 각각 선택할 수도 있다.

[0220] 도 28 은 또한 제 2 시간 인스턴스에서 캡처되는 장면을 나타내는 이미지 (예를 들어, 프레임) (2720) 를 도시한다. 이미지 (2720) 에서, 각각의 오브젝트 (2702 내지 2706) 의 포지션은 제 1 시간 인스턴스에서의 장면을 나타내는 이미지 (2700) 에서 나타내진 대응하는 포지션들에 대해 변화되었다. 예를 들어, 제 1 오브젝트 (2702) 는 별에 더 가깝게 이동하였고, 제 2 오브젝트 (2704) 는 나무에 더 가깝게 이동하였고, 제 3 오브젝트 (2706) 는 이미지의 하부에 더 가깝게 이동하였다.

[0221] 도 24 는 또한, 모바일 디바이스의 스크린 상의 이미지 (2720) 의 디스플레이 (2730) 를 도시한다. 디스플레이 (2730) 는 제 1 윈도우 (2732), 제 2 윈도우 (2734), 제 3 윈도우 (2736), 및 제 4 윈도우 (2738) 로 (예를 들어, 디스플레이 태스크 T50 에 의해) 파티셔닝될 수도 있다. 사용자 선택 메커니즘 (2711) 으로 3 개의 오브젝트들 (2702 내지 2706) 을 선택하는 것에 응답하여, 예를 들어, 모바일 디바이스는 스크린 (2730) 을 3 개의 대응하는 윈도우들 (2732 내지 2736) 및 제 4 윈도우 (2738) 로 파티셔닝할 수도 있다.

[0222] 제 1 윈도우 (2732) 는 제 1 오브젝트 (2702) 를 포함하는 이미지 (2720) 의 제 1 부분을 디스플레이할 수도 있다. 이미지 (2720) 의 제 1 부분은 이미지 (2720) 에서의 제 1 오브젝트 (2702) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트 (2702) 의 로케이션을 추적할 수도 있고 제 1 오브젝트 (2702) 의 로케이션 상에 포커싱된 비디오 스트림이 제 1 윈도우 (2732) 에 디스플레이될 수도 있다. 제 2 윈도우 (2734) 는 제 2 오브젝트 (2704) 를 포함하는 이미지 (2720) 의 제 2 부분을 디스플레이할 수도 있다. 이미지 (2720) 의 제 2 부분은 이미지 (2720) 에서의 제 2 오브젝트 (2704) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 2 오브젝트 (2704) 의 로케이션을 추적할 수도 있고 제 2 오브젝트 (2704) 의 로케이션 상에 포커싱된 비디오 스트림이 제 2 윈도우 (2734) 에 디스플레이될 수도 있다. 제 3 윈도우 (2736) 는 제 3 오브젝트 (2706) 를 포함하는 이미지 (2720) 의 제 3 부분을 디스플레이할 수도 있다. 이미지 (2720) 의 제 3 부분은 이미지 (2720) 에서의 제 3 오브젝트 (2706) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 3 오브젝트 (2706) 의 로케이션을 추적할 수도 있고 제 3 오브젝트 (2706) 의 로케이션 상에 포커싱된 비디오 스트림이 제 3 윈도우 (2736) 에 디스플레이될 수도 있다.

[0223] 제 4 윈도우 (2738) 는 각각의 오브젝트 (2702 내지 2706) 사이의 지오메트리 (예를 들어, 질량 중심 (C_{MB})) 의 표시를 디스플레이할 수도 있다. 예를 들어, 제 4 윈도우 (2738) 는 이미지 (2720) 에서의 영역들에 대응하

는 다수의 섹션들 (그레이스케일 라인들로서 표현됨) 을 포함하는 그리드를 포함할 수도 있다. 예를 들어, 그리드의 하부 좌측 섹션은 이미지 (2720) 의 하부 좌측 섹션에 대응할 수도 있다. 도 21 및 도 22 에 대하여 설명한 바와 유사한 방식으로, 모바일 디바이스는 이미지 (2720) 의 오브젝트들의 질량 중심 (C_{M2}) 을 결정할 수도 있다. 질량 중심 (C_{M2}) 은 그리드의 대응하는 섹션에 디스플레이될 수도 있다.

[0224] 도 28 은 또한 (예를 들어, 제 2 시간 인스턴스에 후속하는) 제 3 시간 인스턴스에서 캡처되는 장면을 나타내는 이미지 (2740) 를 도시한다. 이미지 (2740) 에서, 각각의 오브젝트 (2702 내지 2706) 의 포지션은 제 2 시간 인스턴스에서의 장면을 나타내는 이미지 (2720) 에서의 대응하는 포지션들에 대해 변화되었다. 예를 들어, 제 1 오브젝트 (2702) 는 나무에 더 가깝게 이동하였고, 제 2 오브젝트 (2704) 는 장면 (2740) 의 하부에 더 가깝게 이동하였으며, 제 3 오브젝트 (2706) 는 별에 더 가깝게 이동하였다.

[0225] 도 28 은 또한 모바일 디바이스에 의해 발생하는 윈도우들 (2732 내지 2736) 을 이용하는 모바일 디바이스의 스크린 상의 이미지 (2740) 의 디스플레이 (2750) 를 도시한다. 예를 들어, 제 1 윈도우 (2732) 는 제 1 오브젝트 (2702) 를 포함하는 이미지 (2740) 의 제 1 부분을 디스플레이하고, 제 2 윈도우 (2734) 는 제 2 오브젝트 (2704) 를 포함하는 이미지 (2740) 의 제 2 부분을 디스플레이하며, 제 3 윈도우 (2736) 는 제 3 오브젝트 (2706) 를 포함하는 이미지 (2740) 의 제 3 부분을 디스플레이한다.

[0226] 제 4 윈도우 (2738) 는 제 3 시간 인스턴스에서의 이미지 (2740) 를 반영하도록 표시 (예를 들어, 질량 중심 (C_{M2})) 를 업데이트할 수도 있다. 모바일 디바이스는, 프레임-바이-프레임 기반으로, 또는 비-연속 시간 간격들에서, 지오메트리를 추적하고, 업데이트된 지오메트리의 표시를 제 4 윈도우 (2738) 에 디스플레이할 수도 있다. 예를 들어, 제 3 프레임에서, 제 4 윈도우 (2738) 는 제 3 시간 인스턴스에서 이미지 (2740) 에서의 오브젝트들 (2702 내지 2706) 의 질량 중심 (C_{M3}) 을 디스플레이할 수도 있다. 제 4 윈도우 (2738) 는 또한 제 2 프레임과 제 3 프레임 사이의 (예를 들어, 태스크 T200 에 의해 나타난 바와 같은) 오브젝트들의 질량 중심을 추적하는 경로 (예를 들어, 점선들) 를 디스플레이할 수도 있다.

[0227] 도 28 에 나타난 실시형태는 모바일 디바이스가 프레임들 (2700, 2720, 2740) 에서 오브젝트들 (2702 내지 2706) 을 추적하는 것을 가능하게 할 수도 있고, 별개의 윈도우들 (2732 내지 2736) 에서 오브젝트들 (2702 내지 2706) 에 대응하는 개별화된 (예를 들어, 포커싱된) 비디오 스트림들을 디스플레이할 수도 있다. 실시형태는 또한 사용자가 제 4 윈도우 (2738) 에서 오브젝트들 (2702 내지 2706) 의 지오메트리를 뷰잉하는 것을 가능하게 한다. 모바일 디바이스는 단일 카메라를 이용하여 또는 다수의 카메라들을 이용하여 도 28 에 대해 설명되는 기법들을 수행할 수도 있다.

[0228] 도 29 를 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법 (2800) 의 특정 실시형태를 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법 (2800) 은 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 또는 이들의 임의의 조합을 이용하여 수행될 수도 있다.

[0229] 방법 (2800) 은, 2802 에서, 모바일 디바이스에서, 장면에서의 제 1 오브젝트, 장면에서의 제 2 오브젝트, 및 장면에서의 제 3 오브젝트의 선택을 받는 단계를 포함한다. 예를 들어, 도 28 을 참조하면, 사용자는 스크린 (2710) 상에서 사용자 선택 메커니즘 (2711) 을 터치하고 제 1 오브젝트 (2702) 주위로 제 1 포커스 링 (2712) 을, 제 2 오브젝트 (2704) 주위로 제 2 포커스 링 (2714) 을, 그리고 제 3 오브젝트 (2706) 주위로 제 3 포커스 링 (2716) 을 드래그하여, 제 1, 제 2, 및 제 3 오브젝트들 (2702 내지 2706) 각각의 추적을 가능하게 할 수도 있다. 디스플레이된 장면은 모바일 디바이스의 스크린 상에서 캡처되고 디스플레이되는 비디오 스트림에 대응할 수도 있다.

[0230] 2804 에서, 디스플레이는 제 1 윈도우, 제 2 윈도우, 제 3 윈도우, 및 제 4 윈도우로 파티셔닝될 수도 있다. 예를 들어, 도 28 을 참조하면, 디스플레이 (2730) 는 제 1 윈도우 (2732), 제 2 윈도우 (2734), 제 3 윈도우 (2736), 및 제 4 윈도우로 파티셔닝될 수도 있다.

[0231] 2806 에서, 제 1 오브젝트를 포함하는 디스플레이된 장면의 제 1 부분은 제 1 윈도우에 디스플레이될 수도 있다. 예를 들어, 도 28 을 참조하면, 제 1 윈도우 (2732) 는 제 1 오브젝트 (2702) 를 포함하는 이미지 (2720) 의 제 1 부분을 디스플레이할 수도 있다. 이미지 (2720) 의 제 1 부분은 이미지 (2720) 에서의 제 1 오브젝트 (2702) 의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 1 오브젝트

(2702)의 로케이션을 추적할 수도 있고, 제 1 오브젝트 (2702)의 로케이션 상에 포커싱된 비디오 스트림이 제 1 윈도우 (2732)에 디스플레이될 수도 있다.

[0232] 2808에서, 제 2 오브젝트를 포함하는 디스플레이된 장면의 제 2 부분은 제 2 윈도우에 디스플레이될 수도 있다. 예를 들어, 도 28을 참조하면, 제 2 윈도우 (2734)는 제 2 오브젝트 (2704)를 포함하는 이미지 (2720)의 제 2 부분을 디스플레이할 수도 있다. 이미지 (2720)의 제 2 부분은 이미지 (2720)에서의 제 2 오브젝트 (2704)의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 2 오브젝트 (2704)의 로케이션을 추적할 수도 있고, 제 2 오브젝트 (2704)의 로케이션 상에 포커싱된 비디오 스트림이 제 2 윈도우 (2734)에 디스플레이될 수도 있다.

[0233] 2810에서, 제 3 오브젝트를 포함하는 디스플레이된 장면의 제 3 부분은 제 3 윈도우에 디스플레이될 수도 있다. 예를 들어, 도 28을 참조하면, 제 3 윈도우 (2736)는 제 3 오브젝트 (2706)를 포함하는 이미지 (2720)의 제 3 부분을 디스플레이할 수도 있다. 이미지 (2720)의 제 3 부분은 이미지 (2720)에서의 제 3 오브젝트 (2706)의 로케이션에 기초하여 결정될 수도 있다. 예를 들어, 모바일 디바이스는 제 3 오브젝트 (2706)의 로케이션을 추적할 수도 있고, 제 3 오브젝트 (2706)의 로케이션 상에 포커싱된 비디오 스트림이 제 3 윈도우 (2736)에 디스플레이될 수도 있다.

[0234] 2812에서, 각각의 오브젝트 사이의 지오메트리의 표시가 제 4 윈도우에 디스플레이될 수도 있다. 예를 들어, 도 28을 참조하면, 제 4 윈도우 (2738)는 각각의 오브젝트 (2702 내지 2706) 사이의 질량 중심 (C_{M2})을 디스플레이할 수도 있다. 제 4 윈도우 (2738)는 제 3 시간 인스턴스에서의 장면을 반영하도록 질량 중심 (C_{M2})을 업데이트할 수도 있다. 예를 들어, 제 3 프레임에서, 제 4 윈도우 (2738)는 제 3 시간 인스턴스에서 이미지 (2740)에서의 오브젝트들 (2702 내지 2706)의 질량 중심 (C_{M3})을 디스플레이할 수도 있다. 제 4 윈도우 (2738)는 또한 제 2 프레임과 제 3 프레임 사이의 오브젝트들의 질량 중심을 추적하는 경로 (예를 들어, 점선들)를 디스플레이할 수도 있다.

[0235] 도 29의 방법 (2800)은 모바일 디바이스가 장면 (2700, 2720, 2740)에서 오브젝트들 (2702 내지 2706)을 추적하고, 별개의 윈도우들 (2732 내지 2736)에서 오브젝트들 (2702 내지 2706)에 대응하는 개별화된 (예를 들어, 포커싱된) 비디오 스트림들을 디스플레이하는 것을 가능하게 할 수도 있다. 방법 (2800)은 또한 사용자가 제 4 윈도우 (2738)에서 오브젝트들 (2702 내지 2706)의 지오메트리를 뷰잉하는 것을 가능하게 한다. 모바일 디바이스는 단일 카메라를 이용하여 또는 다수의 카메라들을 이용하여 방법 (2800)을 수행할 수도 있다.

[0236] 도 42를 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱을 나타내는 특정 실시형태가 도시된다. 도 42에 나타내진 실시형태에 대하여 설명된 비디오 프로세싱 기법들은 도 1의 전자 디바이스 (102), 도 2b의 프로세서 (264)에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9의 평활화 모듈 (961), 또는 이들의 임의의 조합에 의해 수행될 수도 있다.

[0237] 도 42는 제 1 시간 인스턴스에서의 장면을 나타내는 디지털 이미지 (1100)를 도시한다. 예를 들어, 이미지 (1100)는 모바일 디바이스의 스크린 상에서 캡처되고 디스플레이되는 비디오 스트림에 대응할 수도 있다. 모바일 디바이스는 스트림을 (예를 들어, 모바일 디바이스의 카메라로) 캡처하거나 또는 다른 디바이스로부터 그 스트림을 수신하도록 구성될 수도 있다. 이미지 (1100)에서 나타내진 장면은 제 1 관심 영역 (ROI) (1102), 제 2 ROI (1104), 및 제 3 ROI (1106)를 포함할 수도 있다. 각각의 ROI (1102 내지 1106)는 (예를 들어, 도 1 내지 도 10, 오브젝트 추적 및 검출 모듈들 (204 및 804), 및/또는 태스크 T200을 참조하여) 본 명세서에서 설명된 기법들에 따라 추적되는 하나 이상의 오브젝트들을 포함하는 장면 내의 구역들에 대응할 수도 있다.

[0238] 예시하기 위해, 제 1 ROI (1102)는 제 1 그룹의 추적된 오브젝트들 (예를 들어, 제 1 추적된 오브젝트 (1110) 및 제 3 추적된 오브젝트 (1114))을 포함할 수도 있다. 부가적으로, 제 2 ROI (1104)는 제 2 그룹의 추적된 오브젝트들 (예를 들어, 제 4 추적된 오브젝트 (1120), 제 5 추적된 오브젝트 (1122), 및 제 6 추적된 오브젝트 (1124))을 포함할 수도 있고, 제 3 ROI (1106)는 제 3 그룹의 추적된 오브젝트들 (예를 들어, 제 7 추적된 오브젝트 (1130), 제 8 추적된 오브젝트 (1132), 및 제 9 추적된 오브젝트 (1134))을 포함할 수도 있다. 제 2 추적된 오브젝트 (1112)는 ROI들 (1102 내지 1104)외부에 로케이팅될 수도 있다.

[0239] 메타데이터는 장면 (1100)에서 각각의 추적된 오브젝트 (1110 내지 1114, 1120 내지 1124, 1130 내지 1134)

의 로케이션을 식별하기 위해 (예를 들어, 태스크 T300 에 의해) 발생될 수도 있다. 특정 실시형태에서, 메타데이터는 직교 좌표 데이터에 대응한다. 예를 들어, 제 1 메타데이터는 제 1 추적된 오브젝트 (1110) 의 로케이션 (예를 들어, x-좌표 로케이션 및 y-좌표 로케이션) 을 식별하도록 발생될 수도 있고, 제 2 메타데이터는 제 2 추적된 오브젝트 (1112) 의 로케이션을 식별하도록 발생될 수도 있으며 등등이다. 다른 실시형태들에서, 메타데이터는 극 좌표 데이터, 구형 좌표 데이터, 원통형 좌표 데이터, 또는 3 차원 데카르트 좌표 데이터에 대응할 수도 있다.

[0240] 각각의 ROI (1102 내지 1106) 는 ROI (1102 내지 1106) 에서의 추적된 오브젝트들의 질량 중심에 의해 정의될 수도 있다. 예를 들어, 제 1 ROI (1102) 는 제 1 추적된 오브젝트 (1110) 와 제 3 추적된 오브젝트 (1114) 사이의 제 1 질량 중심 (C_{M1}) 에 의해 정의될 수도 있다. 제 2 ROI (1104) 는 제 4 추적된 오브젝트 (1120), 제 4 추적된 오브젝트 (1122), 및 제 6 추적된 오브젝트 (1124) 사이의 제 2 질량 중심 (C_{M2}) 에 의해 정의될 수도 있다. 제 3 ROI (1106) 는 제 7 추적된 오브젝트 (1130), 제 8 추적된 오브젝트 (1132), 및 제 9 추적된 오브젝트 (1134) 사이의 제 3 질량 중심 (C_{M3}) 에 의해 정의될 수도 있다.

[0241] 도 43 에 대하여 더 상세히 설명한 바와 같이, 코드북은 장면 (1100) 에서의 각각의 추적된 오브젝트 (1110 내지 1114, 1120 내지 1124, 1130 내지 1134) 의 로케이션에 대한 메타데이터를 인코딩하는데 이용될 수도 있다. ROI들 (1102 내지 1106) 중 하나 내측에 로케이팅되는 추적된 오브젝트들에 대한 메타데이터는 ROI들 (1102 내지 1106) 외측에 로케이팅된 추적된 오브젝트들 (예를 들어, 제 2 추적된 오브젝트 (1112)) 에 대한 메타데이터보다 더 높은 비트 레이트로 인코딩될 수도 있다. ROI들 (1102 내지 1106) 중 하나 내측에 로케이팅되는 추적된 오브젝트들에 대한 메타데이터를 더 높은 비트 레이트로 인코딩하는 것은 ROI들 (1102 내지 1106) 에서의 추적된 오브젝트들의 로케이션들이 다른 추적된 오브젝트들의 로케이션보다 더 높은 레벨의 정확도로 인코딩되는 것을 가능하게 할 수도 있다.

[0242] 도 43 을 참조하면, 오브젝트 로케이션에 대한 메타데이터를 인코딩하는 특정 실시형태가 도시된다. 도 43 의 실시형태는 프레임-바이-프레임 기반으로 다수의 오브젝트들의 로케이션에 대한 메타데이터를 인코딩하는데 이용될 수도 있는 코드북 (1202) 을 예시한다. 메타데이터는 각각의 오브젝트의 로케이션들을 그리드에 나타내는데 이용될 수도 있다.

[0243] 코드북 (1202) 은 다수의 프레임들 (1204) 에 대한 메타데이터를 인코딩하는데 이용될 수도 있다. 메타데이터는 프레임에서의 오브젝트들에 대한 직교 좌표 데이터 (예를 들어, x-좌표 및 y-좌표), 프레임에서의 오브젝트에 대한 극 좌표들, 프레임에서의 오브젝트들에 대한 구형 좌표들, 프레임에서의 오브젝트들에 대한 원통형 좌표들, 또는 프레임에서의 오브젝트들에 대한 3 차원 데카르트 좌표들을 포함할 수도 있다. 예시된 실시형태에서, 코드북 (1202) 은 10,000 개의 프레임들에 걸친 3 개의 오브젝트들에 대한 메타데이터를 인코딩할 수도 있다. 특정 실시형태에서, 제 3 프레임은 도 43 의 제 1 프레임 (예를 들어, 장면 (1100)) 에 대응할 수도 있다. 10,000 개의 프레임들에 대한 메타데이터는 코드북 (1202) 에서 나타내지지만, 다른 실시형태들에서, 코드북 (1202) 은 더 적은 (또는 부가적인) 프레임들에 대한 메타데이터를 인코딩할 수도 있다. 부가적으로, 메타데이터는 더 적은 (또는 부가적인) 오브젝트들에 대한 로케이션 정보를 포함할 수도 있다. 도 43 의 다음의 설명은 제 3 프레임에 대응하는 메타데이터에 기초한다. 그러나, 유사한 기법들이 다른 메타데이터에 적용가능할 수도 있다.

[0244] 제 3 프레임에서의 메타데이터는 제 1 좌표 데이터 (1232), 제 2 좌표 데이터 (1234), 제 3 좌표 데이터 (1236), 제 4 좌표 데이터 (1237), ..., 및 제 9 좌표 데이터 (1238) 를 포함할 수도 있다. 제 1 좌표 데이터 (1232) 는 도 42 의 이미지 (1100) 에서의 제 1 추적된 오브젝트 (1110) 의 좌표 포지션을 포함할 수도 있고, 제 2 좌표 데이터 (1234) 는 이미지 (1100) 에서의 제 2 추적된 오브젝트 (1112) 의 좌표 포지션을 포함할 수도 있고, 제 3 좌표 데이터 (1236) 는 이미지 (1100) 에서의 제 3 추적된 오브젝트 (1114) 의 좌표 포지션을 포함할 수도 있고, 제 4 좌표 데이터 (1237) 는 이미지 (1100) 에서의 제 4 추적된 오브젝트 (1120) 의 좌표 포지션을 포함할 수도 있으며, 등등이다.

[0245] 프로세서는 좌표 데이터 (1232 내지 1236) 를 인코딩하기 위해 코드북 (1202) 을 이용할 수도 있다. 예시하기 위해, 프로세서는 제 1 비트 레이트 (예를 들어, 상대적으로 높은 비트 레이트) 를 이용하여 제 1 좌표 데이터 (1232) 를 제 1 인코딩된 로케이션 데이터 (1242) 로 인코딩 (예를 들어, 양자화) 할 수도 있다. 예를 들어, 프로세서는 제 1 좌표 데이터 (1232) 를 대응하는 값 (예를 들어, 코드북 (1202) 의 코드워드) 에 맵핑할 수도 있고 제 1 좌표 데이터 (1232) 를 값 (예를 들어, 제 1 인코딩된 로케이션 데이터 (1242)) 으로서 인코딩

할 수도 있다. 프로세서는 제 1 추적된 오브젝트 (1110) 가 ROI (예를 들어, 도 42 의 제 1 ROI (1102)) 내측에 있기 때문에 제 1 비트 레이트를 이용할 수도 있다.

[0246] 부가적으로, 프로세서는 제 1 비트 레이트보다 더 낮은 제 2 비트 레이트를 이용하여 제 2 좌표 데이터 (1234) 를 제 2 인코딩된 로케이션 데이터 (1244) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 제 2 좌표 데이터 (1234) 를 코드북 (1202) 을 이용하여 대응하는 값에 맵핑할 수도 있고 제 2 좌표 데이터 (1234) 를 값 (예를 들어, 제 2 인코딩된 로케이션 데이터 (1244)) 으로서 인코딩할 수도 있다. 프로세서는 제 2 추적된 오브젝트 (1112) 가 도 42 의 ROI들 (1102 내지 1106) 의 각각의 외측에 있기 때문에 제 2 비트 레이트를 이용할 수도 있다.

[0247] 프로세서는 제 2 비트 레이트보다 더 큰 제 3 비트 레이트를 이용하여 제 3 좌표 데이터 (1236) 를 제 3 인코딩된 로케이션 데이터 (1246) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 코드북 (1202) 을 이용하여 제 3 좌표 데이터 (1236) 를 대응하는 값에 맵핑시키고 제 3 직교 좌표 데이터 (1236) 를 값 (예를 들어, 제 3 인코딩된 로케이션 데이터 (1246)) 으로서 인코딩할 수도 있다. 프로세서는 제 3 추적된 오브젝트 (1114) 가 ROI (예를 들어, 도 42 의 제 1 ROI (1102)) 내측에 있기 때문에 제 3 비트 레이트를 이용할 수도 있다. 특정 실시형태에서, 제 3 비트 레이트는 제 1 비트 레이트와 동일하거나 실질적으로 동일할 수도 있다.

[0248] 프로세서는 제 2 비트 레이트보다 더 큰 제 4 비트 레이트를 이용하여 제 4 좌표 데이터 (1237) 를 제 4 인코딩된 로케이션 데이터 (1247) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 코드북 (1202) 을 이용하여 제 4 좌표 데이터 (1237) 를 대응하는 값에 맵핑시킬 수도 있고 제 4 좌표 데이터 (1237) 를 값 (예를 들어, 제 4 인코딩된 로케이션 데이터 (1247)) 으로서 인코딩할 수도 있다. 프로세서는 제 4 추적된 오브젝트 (1120) 가 ROI (예를 들어, 도 42 의 제 2 ROI (1104)) 내측에 있기 때문에 제 4 비트 레이트를 이용할 수도 있다. 특정 실시형태에서, 제 4 비트 레이트는 제 1 비트 레이트와 실질적으로 동일할 수도 있다.

[0249] 프로세서는 제 2 비트 레이트보다 더 큰 제 9 비트 레이트를 이용하여 제 9 좌표 데이터 (1238) 를 제 9 인코딩된 로케이션 데이터 (1248) 로 인코딩할 수도 있다. 예를 들어, 프로세서는 코드북 (1202) 을 이용하여 제 9 좌표 데이터 (1238) 를 대응하는 값에 맵핑시킬 수도 있고 제 9 좌표 데이터 (1238) 를 값 (예를 들어, 제 9 인코딩된 로케이션 데이터 (1248)) 으로서 인코딩할 수도 있다. 프로세서는 제 9 추적된 오브젝트 (1134) 가 ROI (예를 들어, 도 42 의 제 3 ROI (1106)) 내측에 있기 때문에 제 9 비트 레이트를 이용할 수도 있다. 특정 실시형태에서, 제 9 비트 레이트는 제 1 비트 레이트와 실질적으로 동일할 수도 있다.

[0250] 유사한 인코딩 기법들은 도 42 의 다른 추적된 오브젝트들의 좌표 데이터에 대해 수행될 수도 있다. ROI들 (1102 내지 1106) 중 하나 내측에 로케이팅되는 추적된 오브젝트들에 대한 메타데이터 (예를 들어, 좌표 데이터) 는 ROI들 (1102 내지 1106) 외측에 로케이팅되는 추적된 오브젝트들 (예를 들어, 제 2 추적된 오브젝트 (1112)) 에 대한 메타데이터보다 더 높은 비트 레이트로 인코딩될 수도 있다. 더 높은 비트 레이트로 ROI들 (1102 내지 1106) 중 하나 내측에 로케이팅되는 추적된 오브젝트들에 대한 메타데이터를 인코딩하는 것은 추적된 오브젝트들의 로케이션들이 상대적으로 높은 레벨의 정밀도로 인코딩되는 것을 가능하게 할 수도 있다.

[0251] 도 43 의 실시형태는 사용자가 코드북 (1202) 을 이용하여 인코딩된 데이터에 기초하여 비디오의 특정 부분들을 취출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 예를 들어, 로케이션 공간은 사용자가 상이한 오브젝트들에 대해 상이한 로케이션들 (예를 들어, 파라미터들) 을 선택하는 것을 가능하게 하는 사용자-상호작용 그리드로서 스크린 상에 디스플레이될 수도 있다. 그 선택에 기초하여, 모바일 디바이스의 사용자는 비디오를 통해 내비게이팅 (예를 들어, 플레이, 빨리 감기 (fast-forward), 되감기 등) 해야 하는 일 없이 유사한 오브젝트 로케이션 구성들을 갖는 비디오의 원하는 부분들 (예를 들어, 프레임들) 을 쉽게 로케이팅할 수도 있다. 부가적으로, 인코딩된 로케이션 데이터 (1242 내지 1248) 는 인코더로부터 디코더에 압축 및 송신될 수도 있다. 인코딩된 로케이션 데이터 (1242 내지 1248) 의 디코더에의 압축 및 송신은 디코더가 상대적으로 낮은 양의 데이터 (예를 들어, 양자화된 값들) 를 이용하여 추적된 오브젝트들의 포지션을 결정하는 것을 가능하게 할 수도 있다.

[0252] 도 44 를 참조하면, 오브젝트 추적을 이용하는 비디오 프로세싱 방법 (1300) 의 특정 실시형태를 예시하는 플로우차트가 도시된다. 예시적 실시형태에서, 방법 (1300) 은 도 1 의 전자 디바이스 (102), 도 2b 의 프로세서 (264) 에 의해 구현되는 오브젝트 추적 및 검출 모듈 (204), 프로세서에 의해 구현되는 도 8 의 오브젝트 추적 및 검출 모듈 (804), 프로세서에 의해 구현되는 도 9 의 평활화 모듈 (961), 도 43 의 코드북 (1202), 또는 이들의 임의의 조합을 이용하여 수행될 수도 있다.

- [0253] 방법 (1300) 은 1302 에서, 모바일 디바이스에서, 특정 프레임의 관심 구역에 로케이팅된 제 1 오브젝트의 제 1 로케이션을 (예를 들어, 태스크 T200 에 의해) 추적하는 단계를 포함할 수도 있다. 예를 들어, 도 42 를 참조하면, 제 1 추적된 오브젝트 (1110) 의 제 1 로케이션이 추적될 수도 있다. 1304 에서, 특정 프레임의 관심 구역에 로케이팅되지 않은 제 2 오브젝트의 제 2 로케이션이 추적될 수도 있다. 예를 들어, 도 42 를 참조하면, 제 2 추적된 오브젝트 (1112) 의 제 2 로케이션이 추적될 수도 있다.
- [0254] 1306 에서, 제 1 로케이션과 연관된 제 1 좌표 데이터 및 제 2 로케이션과 연관된 제 2 좌표 데이터가 발생할 수도 있다. 예를 들어, 도 43 을 참조하면, 프로세서는 장면 (1100) 에서의 제 1 추적된 오브젝트 (1110) 의 포지션과 연관된 제 1 좌표 데이터 (1232) 및 장면 (1100) 에서의 제 2 추적된 오브젝트 (1112) 의 포지션과 연관된 제 2 좌표 데이터 (1234) 를 발생시킬 수도 있다.
- [0255] 1308 에서, 제 1 좌표 데이터 및 제 2 좌표 데이터는 제 1 인코딩된 로케이션 데이터 및 제 2 인코딩된 로케이션 데이터를 발생시키기 위해 코드북을 이용하여 (예를 들어 태스크 T300 에 의해) 인코딩될 수도 있다. 예를 들어, 도 43 을 참조하면, 프로세서는 제 1 인코딩된 로케이션 데이터 (1242) (예를 들어, 코드북 (1202) 의 코드워드) 를 발생시키기 위해 제 1 좌표 데이터 (1232) 를 인코딩하도록 코드북 (1202) 을 이용할 수도 있고, 프로세서는 제 2 인코딩된 로케이션 데이터 (1244) 를 발생시키기 위해 제 2 좌표 데이터 (1234) 를 인코딩하도록 코드북 (1202) 을 이용할 수도 있다. 제 1 좌표 데이터 (1232) 는 제 1 비트 레이트 (예를 들어, 높은 비트 레이트) 로 인코딩될 수도 있고 제 2 좌표 데이터 (1234) 는 제 2 비트 레이트 (예를 들어, 낮은 비트 레이트) 로 인코딩될 수도 있다. 프로세서는 제 1 추적된 오브젝트 (1110) 가 ROI (예를 들어, 도 42 의 제 1 ROI (1102)) 내측에 있기 때문에 제 1 좌표 데이터 (1232) 를 인코딩하도록 제 1 비트 레이트를 이용할 수도 있다. 프로세서는 제 2 추적된 오브젝트 (1112) 가 도 42 의 ROI들 (1102 내지 1106) 의 각각의 외측에 있기 때문에 제 2 좌표 데이터 (1234) 를 인코딩하도록 제 2 비트 레이트를 이용할 수도 있다.
- [0256] 1310 에서, 제 1 인코딩된 로케이션 데이터 및 제 2 인코딩된 로케이션 데이터는 입력 파라미터들에 기초하여 특정 프레임의 추출을 가능하게 하도록 저장될 수도 있다. 예를 들어, 도 43 을 참조하면, 프로세서는 메모리에 제 1 인코딩된 로케이션 데이터 (1242) 및 제 2 인코딩된 로케이션 데이터 (1244) 를 저장할 수도 있다. 모바일 디바이스의 사용자는 파라미터들 (예를 들어, 제 1 추적된 오브젝트 (1110) 및 제 2 추적된 오브젝트 (1112) 의 원하는 로케이션) 을 입력할 수도 있다. 입력 파라미터들에 기초하여, 모바일 디바이스는 제 1 로케이션 및 제 2 로케이션에 대응하는 입력 파라미터들에 응답하여 재생을 위한 제 3 프레임을 추출할 수도 있다.
- [0257] 특정 실시형태에서, 방법 (1300) 은 특정 프레임의 관심 구역에 로케이팅된 제 3 오브젝트의 제 3 로케이션을 추적하는 단계를 포함할 수도 있다. 예를 들어, 제 3 추적된 오브젝트 (1114) 의 제 3 로케이션이 추적될 수도 있다. 제 3 로케이션과 연관된 제 3 좌표 데이터가 발생할 수도 있다. 제 3 좌표 데이터는 제 3 인코딩된 로케이션 데이터를 발생시키기 위해 코드북을 이용하여 인코딩될 수도 있다. 제 3 좌표 데이터는 제 2 비트 레이트보다 더 큰 제 3 비트 레이트를 이용하여 인코딩될 수도 있다. 제 3 인코딩된 로케이션 데이터는 입력 파라미터들에 기초하여 특정 프레임의 추출을 가능하게 하도록 저장될 수도 있다.
- [0258] 특정 실시형태에서, 방법 (1300) 은 특정 프레임의 제 2 관심 구역에 로케이팅된 제 4 오브젝트의 제 4 로케이션을 추적하는 단계를 포함할 수도 있다. 예를 들어, 제 4 추적된 오브젝트 (1120) 의 제 4 로케이션이 추적될 수도 있다. 제 4 로케이션과 연관된 제 4 좌표 데이터가 발생할 수도 있다. 제 4 좌표 데이터는 제 4 인코딩된 로케이션 데이터를 발생시키기 위해 코드북을 이용하여 인코딩될 수도 있다. 제 4 좌표 데이터는 제 2 비트 레이트보다 더 큰 제 4 비트 레이트를 이용하여 인코딩될 수도 있다. 제 4 인코딩된 로케이션 데이터는 입력 파라미터들에 기초하여 특정 프레임의 추출을 가능하게 하도록 저장될 수도 있다.
- [0259] 도 44 의 방법 (1300) 은 사용자가 코드북 (1202) 을 이용하여 인코딩된 비디오의 특정 부분들 (프레임들) 을 추출 (예를 들어, 로케이팅) 하는 것을 가능하게 할 수도 있다. 추가적으로, 방법 (1300) 은 ROI 내측의 추적된 오브젝트들에 대한 메타데이터가 ROI 외측의 오브젝트들보다 더 높은 비트 레이트로 인코딩되는 것을 가능하게 할 수도 있다. ROI 내측에 로케이팅된 추적된 오브젝트들에 대한 메타데이터를 더 높은 비트 레이트로 인코딩하는 것은 ROI들에서의 추적된 오브젝트들의 로케이션들이 다른 추적된 오브젝트들의 로케이션보다 더 높은 레벨의 정확도로 인코딩되는 것을 가능하게 할 수도 있다.
- [0260] 도 38a 는 일반적 구성에 따라 메타데이터를 발생시키기 위한 장치 (A100) 의 블록 다이어그램을 도시한다. 장치 (A100) 는 복수의 오브젝트들을 선택하기 위해 적어도 하나의 디지털 이미지를 이용하도록 구성된 관별기 (100) 를 포함하고, 여기서 적어도 하나의 디지털 이미지는 (예를 들어, 태스크 T100 에 관하여 본 명세서에서

설명된 바와 같이) 오브젝트를 물리적 공간에 관하여 나타낸다. 장치 (A100) 는 또한, 로케이션 공간에서의 선택된 오브젝트들의 포지션들을 나타내는 정보에 기초하여, (예를 들어, 태스크 T300 에 관하여 본 명세서에서 설명한 바와 같이) 선택된 오브젝트들의 복수의 후보 기하학적 배열들 중에서 하나를 식별하는 메타데이터를 생성하도록 구성된 선택기 (300) 를 포함한다.

[0261] 도 38b 는 장치 (A100) 의 구현 (A110) 의 블록 다이어그램을 도시한다. 장치 (A110) 는 (예를 들어, 디스플레이 태스크 T50 에 관하여 본 명세서에서 설명한 바와 같이) 적어도 하나의 디지털 이미지를 디스플레이하도록 구성된 디스플레이 모듈 (50) 을 포함한다.

[0262] 도 38c 는 장치 (A100) 의 구현 (A120) 의 블록 다이어그램을 도시한다. 장치 (A120) 는 (예를 들어, 파싱 태스크 T30 에 관하여 본 명세서에서 설명한 바와 같이) 적어도 하나의 디지털 이미지를 생성하기 위해 비디오 스트림을 파싱하도록 구성된 인터페이스 (30) 를 포함한다. 도 38d 는 장치 (A110 및 A120) 의 구현 (A130) 의 블록 다이어그램을 도시한다.

[0263] 도 38e 는 장치 (A100) 의 구현 (A200) 의 블록 다이어그램을 도시한다. 장치 (A200) 는 (예를 들어, 태스크 T200 에 관하여 본 명세서에서 설명한 바와 같이) 로케이션 공간에서의 복수의 오브젝트들에 대한 레퍼런스 포지션을 결정하도록 구성된 오브젝트 추적기 (200) 를 포함한다. 오브젝트 추적기 (200) 는 또한, 비디오 스트림 내에서, (예를 들어, 태스크 T200 에 관하여 본 명세서에서 설명한 바와 같이) 상기 복수의 오브젝트들을 추적하도록 구성될 수도 있다. 도 38f, 도 38g, 및 도 39a 는 오브젝트 추적기 (200) 의 인스턴스들을 포함하는 장치 (A110, A120, 및 A130) 각각의 구현들 (A210, A220, 및 A230) 의 블록 다이어그램들을 도시한다.

[0264] 도 39b 는 장치 (A100) 의 구현 (A300) 의 블록 다이어그램을 도시한다. 장치 (A200) 는 (예를 들어, 태스크 T300 에 관하여 본 명세서에서 설명한 바와 같이) 선택기 (300) 에 의해 생성된 메타데이터를 비디오 스트림의 대응하는 프레임들과 연관시키도록 구성된 비디오 스트림 에디터 (400) 를 포함한다. 도 39c 및 도 39d 는 비디오 스트림 에디터 (400) 의 인스턴스들을 포함하는 장치 (A200 및 A230) 각각의 구현들 (A310 및 A330) 의 블록 다이어그램들을 도시한다.

[0265] 본 명세서에서 설명한 바와 같은 장치 (A100 및/또는 A500) 의 구현들 중 임의의 구현은 다음의 폼 팩터들 중 임의의 것에 있어서의 디바이스와 같은 모바일 디바이스 내에서 구현될 수도 있다: 홀더블들 (holdables) (예를 들어, 스마트폰들), 드라이버블들 (drivables) (예를 들어, 운반체들 또는 로봇들), 웨어러블들 (wearables) (예를 들어, 의류 또는 액세서리들), 및 플라이어블들 (flyables) (예를 들어, 드론들). 모바일 디바이스는 적어도 하나의 디지털 이미지를 디스플레이하도록 구성된 하나 이상의 스크린들 (예를 들어, 터치스크린) 및/또는 적어도 하나의 디지털 이미지 (예를 들어, 비디오 스트림) 를 캡처하도록 구성된 하나 이상의 이미징 유닛들 (예를 들어, 카메라 또는 다른 이미지-캡처링 디바이스) 을 포함할 수도 있다.

[0266] 도 31 을 참조하면, 상술된 비디오 프로세싱 기법들을 수행하도록 동작가능한 컴포넌트들을 포함하는 무선 디바이스 (2900) 의 블록 다이어그램이 도시된다. 디바이스 (2900) 는 메모리 (2932) 에 커플링된, 디지털 신호 프로세서 (DSP) 와 같은 프로세서 (2910) 를 포함한다. 디바이스 (2900) 는 본 명세서에서 설명된 장치 (A100 및/또는 A500) 의 구현들 중 임의의 구현의 엘리먼트들을 포함하도록 구현될 수도 있는 모바일 디바이스의 하나의 예이다.

[0267] 도 31 은 또한, 프로세서 (2910) 및 디스플레이 (2928) 에 커플링되는 디스플레이 제어기 (2926) 를 도시한다. 특정 실시형태에서, 디스플레이는 상술된 뷰파인더 또는 스크린들에 대응할 수도 있다. 메모리 (2932) 는 실행가능한 명령들 (2956) 을 포함하는 유형의 비일시적 프로세서 관독가능 저장 매체일 수도 있다. 명령들 (2956) 은 프로세서 (2910) 와 같은 프로세서에 의해 실행될 수도 있다. 디스플레이 제어기 (2926) 또는 프로세서 (2910) 와 같은 프로세서는 도 3 의 방법 (300), 도 4 의 방법 (400), 도 5 의 방법 (500), 도 6 의 방법 (600), 도 10 의 방법 (1000), 도 12 의 방법 (1200), 도 14 의 방법 (1400), 도 16 의 방법 (1600), 도 19 의 방법 (1900), 도 20 의 방법 (2000, 2010), 도 23 의 방법 (2300), 도 25 의 방법 (2500), 도 26 의 방법 (2600), 도 28 의 방법 (2800), 방법 (M100), 방법 (M110), 방법 (M120), 방법 (M130), 또는 이들의 임의의 조합 중 적어도 일 부분을 수행하도록 명령들 (2956) 을 실행할 수도 있다.

[0268] 코더/디코더 (코덱 (CODEC)) (2934) 는 또한 프로세서 (2910) 에 커플링될 수 있다. 스피커 (2936) 및 마이크 (2938) 는 코덱 (2934) 에 커플링될 수 있다. 도 31 은 또한, 무선 제어기 (2940) 가 프로세서 (2910) 및 안테나 (2942) 에 커플링될 수 있다는 것을 나타낸다. 특정 실시형태에서, 무선 주파수 (RF) 인

터페이스 (2980) 는 무선 제어기 (2940) 와 안테나 (2942) 사이에 배치된다.

- [0269] 특정 실시형태에서, 프로세서 (2910), 디스플레이 제어기 (2926), 메모리 (2932), 코덱 (2934), 및 무선 제어기 (2940) 는 시스템-인-패키지 또는 시스템-온-칩 디바이스 (2922) 에 포함된다. 특정 실시형태에서, 입력 디바이스 (2930) (예를 들어, 상술된 바와 같이 장면들을 캡처하는데 이용되는 카메라) 및 파워 서플라이 (2944) 는 시스템-온-칩 디바이스 (2922) 에 커플링된다. 더욱이, 특정 실시형태에서, 도 31 에 예시된 바와 같이, 디스플레이 (2928), 입력 디바이스 (2930), 스피커 (2936), 마이크로폰 (2938), 안테나 (2942), 파워 서플라이 (2944), 및 RF 인터페이스 (2980) 는 시스템-온-칩 디바이스 (2922) 의 외부에 있다. 예를 들어, RF 인터페이스 (2980) 은 별개의 칩일 수도 있다. 그러나, 디스플레이 (2928), 입력 디바이스 (2930), 스피커 (2936), 마이크로폰 (2938), 안테나 (2942), 파워 서플라이 (2944), 및 RF 인터페이스 (2980) 각각은, 인터페이스 또는 제어기와 같은, 시스템-온-칩 디바이스 (2922) 의 컴포넌트에 커플링될 수 있다.
- [0270] 포지션-관련 메타데이터로 비디오 파일들에 주석을 다는 다수의 예들이 (예를 들어, 태스크 T300 과 관련하여) 상술된다. 이러한 주석은 예를 들어, 추출 및 아카이빙을 위한 프레임 분류 및 클러스터링과 같은 애플리케이션들을 위해, 오브젝트 추적에 의해 획득되는 인코딩 포지션들을 포함할 수도 있다.
- [0271] 이러한 기능성은 사용자에게 관심이 있는 콘텐츠의 일부를 식별하는 메타데이터를 (예를 들어, 비디오 스트림이 캡처되고 있을 때) 생성하는데 이용될 수도 있다. 메타데이터에 기초한, 특정된 콘텐츠의 신속한 취출을 위한 메커니즘을 제공하는 것에 의해 하나 이상의 비디오 스트림들의 이러한 관심-기반 프리-필터링의 결과들을 레버리징하는 것이 바람직할 수도 있다. 예를 들어, 수동적 검토의 필요성을 감소 또는 제거하면서 (예를 들어, 큰 목록의 비디오 파일들 중에서) 원하는 콘텐츠의 자동 추출을 지원하는 것이 바람직하다.
- [0272] 도 45a 는 태스크들 (X100, X200, 및 X300) 을 포함하는 일반적 구성에 따른 이미지 취출의 방법 (M500) 의 플로우차트를 도시한다. 태스크 X100 은 복수의 선택된 오브젝트들을 식별하는 정보를 획득하고, 태스크 X200 은 복수의 후보 기하학적 배열들 중에서 하나의 후보 기하학적 배열을 선택한다. 그 선택에 응답하여, 태스크 T300 은 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별한다. 태스크들 (X100, X200, 및 X300) 은 예를 들어, 도 1 의 전자 디바이스 (102), 프로세서 (2910), 무선 디바이스 (2900), 또는 이들의 임의의 조합에 의해 수행될 수도 있다.
- [0273] 태스크 X100 은 복수의 선택된 오브젝트들을 식별하는 정보를 획득한다. 태스크 X100 은 (예를 들어, 터치스크린, 이를 테면 터치스크린 (116) 을 통해) 사용자로부터 정보를 획득하도록 구현될 수도 있다. 이 경우에, 사용자는 터치스크린 상에 디스플레이된 메뉴로부터 오브젝트들 중 하나 이상을 선택할 수도 있다. 예를 들어, 이러한 메뉴는 각각이 특정 사람에 대응하는 (예를 들어, 사용자의 연락처 리스트의 사람들의 썸네일 이미지들) 아이콘들의 디스플레이일 수도 있다.
- [0274] 다른 예에서, 사용자는 하나 이상의 비디오 파일들과 연관된 메타데이터에 기초하여 발생된 터치스크린 상의 디스플레이로부터 오브젝트들 중 하나 이상을 선택할 수도 있다. 예를 들어, 스크린은 (예를 들어, 캡처 동안) 비디오 파일 내에서 추적된 오브젝트들의 아이콘들의 메뉴를 디스플레이하도록 구성될 수도 있다. 이러한 경우에, 비디오 파일과 연관된 메타데이터는 추적된 오브젝트를 식별하는 라벨을 비디오 파일의 하나 이상의 프레임들의 각각에서의 오브젝트의 대응하는 포지션과 연관시키는 정보를 포함할 수도 있다.
- [0275] 유사하게, 사용자는 연관된 메타데이터를 갖는 이미지 (예를 들어, 비디오 파일의 프레임) 로부터 오브젝트들 중 하나 이상을 선택할 수도 있다. 이러한 경우에, 터치스크린 인터페이스는 오브젝트에 대한 포지션 메타데이터가 대응하는 비디오 파일에서 이용가능하다는 것 (즉, 오브젝트가 추적되었다는 것) 을 나타내기 위해 이미지 내의 오브젝트를 하이라이팅하는 것에 의해 이미지를 디스플레이하도록 구성될 수도 있다. 하나의 이러한 예에서, 터치스크린 인터페이스는 (예를 들어, 오브젝트들 주위에 라소를 드로잉하는 것에 의해, 또는 오브젝트들 주위에 박스를 생성하기 위해 핀치 및 줌 액션들을 이용하는 것에 의해) 사용자가 다수의 오브젝트들을 한꺼번에 선택하는 것을 허용하도록 구성될 수도 있다.
- [0276] 이러한 터치스크린 입력에 부가적으로 또는 대안적으로, 태스크 X100 은 스피치 인식에 의해 사용자로부터 정보를 획득하도록 구현될 수도 있다. 사용자는 예를 들어, 스크린 상에 디스플레이된 바와 같은 이용가능한 오브젝트들의 리스트 (예를 들어, 상술한 바와 같은 아이콘들의 메뉴) 로부터 오브젝트의 이름을 관독함으로써 오브젝트들 중 하나 이상을 선택할 수도 있다. 대안적으로, 사용자는 특정 오브젝트에 대한 포지션 메타데이터가 검색될 적어도 하나의 비디오 파일에 존재한다는 것을 단순히 가정할 수도 있고, 이 선택을, 오브젝트의

이름을 말하거나, (예를 들어, 연락처 리스트로부터) 오브젝트와 연관된 아이콘을 선택하거나, 또는 테스트로부터 오브젝트의 이름을 입력하는 것에 의해 입력할 수도 있다.

[0277] 태스크 X200 은 (예를 들어, 도 35 에 관하여 본 명세서에서 설명한 바와 같이) 복수의 후보 기하학적 배열들 중에서 하나를 선택한다. 태스크 X200 은 (예를 들어, 터치스크린, 이를 테면 터치스크린 (116) 을 통해) 사용자로부터 획득된 정보에 기초하여 이러한 선택을 수행하도록 구현될 수도 있다. 하나의 예에서, 사용자는 원하는 배열을 나타내는 이미지 (예를 들어, 비디오 파일의 프레임) 를 특정한다. 다른 예에서, 사용자는 원하는 배열을, 이러한 배열에 아이콘들 (예를 들어, 일반적 아이콘들, 또는 원하는 오브젝트들의 아이콘들) 을 배치하는 것에 의해 설명한다. 이들 2 개의 예들 중 어느 하나를 채용하는 터치스크린 구성이 단일 사용자 동작에서 태스크들 (X100 및 X200) 양자에 사용자 선택 정보를 제공하는데 이용될 수도 있다는 것을 유의한다. 예를 들어, 사용자는 추적된 오브젝트들의 아이콘들의 디스플레이를 이용하여 원하는 배열을 설명하고 그 설명을 생성하기 위해 그들을 이용하기 위해 선택된 오브젝트들을 나타낼 수도 있다. 추가 예에서, 사용자는 이러한 배열들의 디스플레이 중에서 원하는 후보 기하학적 배열을 선택할 수도 있다.

[0278] 부가적으로 또는 대안적으로, 태스크 X200 은 스피치 인식에 의해 사용자로부터 정보를 획득하도록 구현될 수도 있다. 이러한 경우에, 사용자는 예를 들어, '가까운 (close)', '근방의 (near)', '먼 (far)', '다음 (next)', '떨어져 (away)' 등과 같은 설명어들을 이용하여 원하는 배열을 나타낼 수도 있다. 예를 들어, 스피치 인식 인터페이스는 다음의 형태들 중 임의의 하나 이상에서 스피치 요청을 인식하도록 구성될 수도 있다: "Joe 및 Kate 바로 옆의 Tom 을 발견해라", "Jill 및 Sam 으로부터 떨어져 있는 Tom 을 보여 주라", "Sara 에 가깝고 Jill 로부터 먼 Tom 을 디스플레이해라", "Sara 바로 옆이고 나무로부터 먼 Tom 을 검색해라". 다시, 이러한 동작은 단일 사용자 동작에서 태스크들 (X100 및 X200) 양자에 사용자 선택 정보를 제공하는데 이용될 수도 있다.

[0279] 상기 예들 중 임의의 예에서, 태스크 X200 은 예를 들어, 사용자에게 의해 나타내지는 배열에 (예를 들어, 도 35 에 관하여 상술한 바와 같이) 적절한 환경 메트릭들을 적용하는 것에 의해, 후보 기하학적 배열들 중 특정 하나를 선택하도록 구현될 수도 있다. 추가 예에서, 사용자는 이러한 배열들의 디스플레이 중에서 원하는 후보 기하학적 배열을 선택할 수도 있다.

[0280] 그 선택에 응답하여, 태스크 X300 은 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별한다. 하나의 예에서, 획득된 정보는, 복수의 선택된 오브젝트들의 각각에 대해, 그 오브젝트와 연관되는 라벨을 포함하고, 태스크 X300 은 복수의 디지털 이미지들 (예를 들어, 비디오 파일) 과 연관되는 메타데이터 내의 라벨들을 검색하도록 구현된다. 이러한 라벨은 오브젝트의 이름 및 그 오브젝트를 나타내기 위해 메타데이터 내에서 이용되는 임의의 다른 고유한 또는 고유하지 않은 데이터 스트링일 수도 있다. 예를 들어, 사람에 대한 라벨은 사용자의 연락처 리스트의 사람들의 썸네일 이미지와 연관되는 데이터 스트링일 수도 있다.

[0281] 복수의 디지털 이미지들과 연관된 메타데이터는 또한 특정 후보를 나타내는 라벨 (예를 들어, 데이터 스트링) 을 포함하도록 구성될 수도 있다. 이러한 경우에, 태스크 X300 은 메타데이터 내에서, 선택된 후보 기하학적 배열과 연관되는 라벨을 검색하도록 구현될 수도 있다.

[0282] 대안적으로, 메타데이터는 임의의 후보 기하학적 배열을 나타내는 이러한 데이터 스트링이 없을 수도 있지만, 각각의 선택된 오브젝트의 포지션을 나타내는 정보를 여전히 포함할 수도 있다. 이러한 경우에, 태스크 X200 은 이 포지션 메타데이터에 기초하여, 오브젝트들이 선택된 후보 기하학적 배열에 따라 배열되는지 여부를 판정하도록 구현될 수도 있다. 이러한 태스크 X200 의 구현은 예를 들어, 이러한 결정을 수행하기 위해 (예를 들어, 도 35 에 관하여) 본 명세서에서 설명한 바와 같이 하나 이상의 환경 메트릭들을 적용할 수도 있다. 많은 경우들에서, 이미지에 대한 메타데이터에 의해 나타내진 포지션들 사이의 거리들의 비교는 추가 고려사항으로부터 이미지를 제거하는데 충분할 수도 있다.

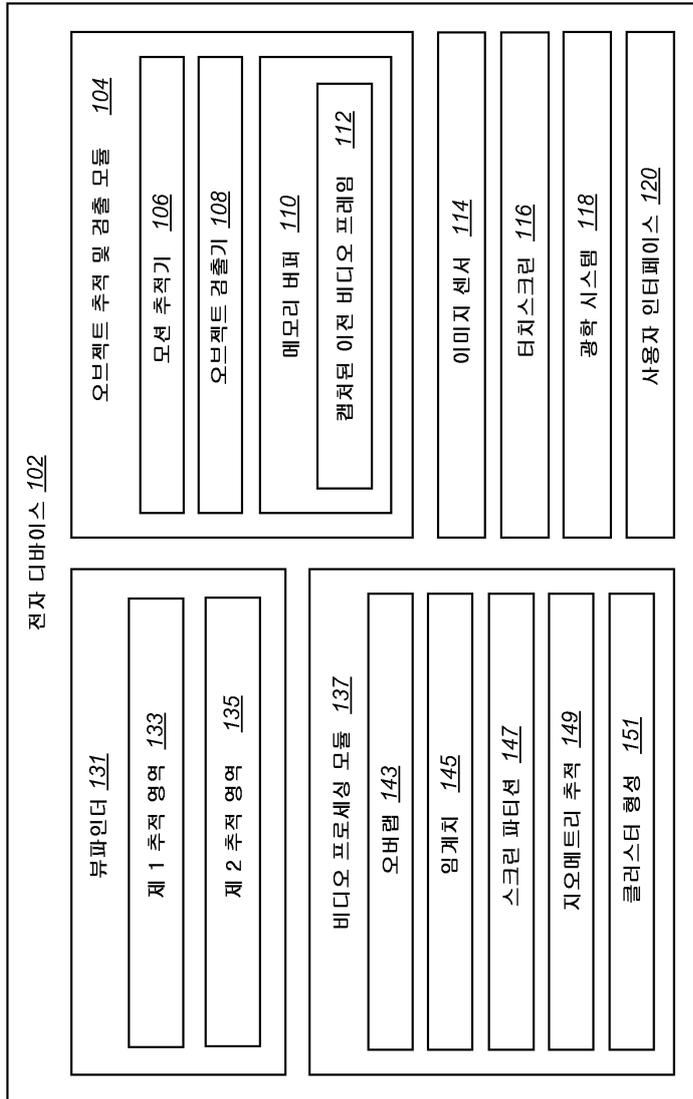
[0283] 선택된 오브젝트들의 기하학적 배열에 기초하여 정보 취출을 위한 비디오 파일 또는 스트림 (또는 디지털 이미지들의 다른 컬렉션) 을 필터링하기 위해 본 명세서에서 설명한 바와 같이 방법 (M100) 의 구현을 이용하는 것이 바람직할 수도 있다. 예를 들어, 특정 배열에서 사람들이 발견되는 이미지들을 식별하기 위해 이러한 구현을 이용하는 것이 바람직할 수도 있다. 이러한 취출의 잠재적인 적용들은 다른 이벤트들과의 이러한 발생들의 상관을 지원하기 위해 선택된 오브젝트들 (예를 들어, 사람들) 의 근접성들을 나타내는 것을 포함할 수도 있다.

- [0284] 다수의 아카이브된 비디오 파일들을 통하여 (예를 들어, 단일 사용자 선택 동작에 응답하여) 검색하기 위해 방법 (M100) 의 구현을 적용하는 것이 바람직할 수도 있다. 이러한 다수의 비디오 파일들의 그룹의 각각에서의 메타데이터의 구성은 하나의 파일에서 다른 파일 간에 상이할 수도 있다는 것이 가능하다. 예를 들어, 특정 비디오 파일과 연관된 메타데이터는 오브젝트-식별 메타데이터 및 포지션-표시 메타데이터를 포함할 수도 있지만, 그 둘 사이의 연관성은 없을 수도 있다. 이러한 경우에, 비디오 파일은 그 파일에서 특정 사람들이 보여진다는 것을 나타내기 위해 태깅될 수도 있고, 파일은 또한, 비디오 스트림 내에서 선택된 사람들을 추적하는 것으로부터 획득된 메타데이터를 포함할 수도 있지만, 추적된 사람들이 태그에서 식별된 동일한 사람이 아닐 수도 있다. 그럼에도 불구하고 이러한 그룹 (예를 들어, 모든 파일들이 선택된 오브젝트들을 나타내는 것으로서 식별되는 그룹) 에 방법 (M100) 의 구현의 단일 인스턴스를 적용하는 것이 바람직할 수도 있다. 예를 들어, 이러한 방법이 추적된 오브젝트들의 원하는 배열을 갖는 프레임들을 식별하는 것으로 결론지어졌다면, 식별된 프레임들의 세트는 추적된 오브젝트들이 또한 선택된 오브젝트들인 그 프레임들만을 구분하도록 검토될 수도 있다.
- [0285] 태스크들 (T200 및 T300) 에 관하여 상술한 바와 같이, 비디오 프레임과 연관된 메타데이터는 로케이션 공간에서의 기하학적 배열에 대한 레퍼런스 포지션을 나타낼 수도 있다. 방법 (M500) 은 (예를 들어, 선택된 후보 기하학적 배열의 포지션의 사용자의 표시에 기초하여) 선택된 후보 기하학적 배열에 대한 레퍼런스 포지션을 선택하도록 구현될 수도 있다. 이러한 경우에, 태스크 X300 은 선택된 오브젝트들이 로케이션 공간에서의 레퍼런스 포지션에 상대적인 것이 아니라 후보 기하학적 배열에 따라 배열되는 이미지들로부터의 후보 기하학적 배열에 따라, 로케이션 공간에서의 레퍼런스 포지션에 상대적인, 선택된 오브젝트들이 배열되는 이미지들을 구별하도록 구현될 수도 있다.
- [0286] 도 45b 는 일반적 구성에 따른 이미지 취출을 위한 장치 (A500) 의 블록 다이어그램을 도시한다. 장치 (A500) 는 (A) 복수의 선택된 오브젝트들을 식별하는 정보 및 (B) 복수의 후보 기하학적 배열들 중에서 하나의 선택을 획득하도록 구성된 판별기 (Z100) 를 포함한다. 판별기 (Z100) 는 예를 들어, 스피치 인식 인터페이스 및/또는 터치스크린 인터페이스를 통해, 정보 및/또는 선택을 획득하도록 구성될 수도 있다.
- [0287] 장치 (A500) 는 또한, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하도록 구성된 검색 엔진 (Z300) 을 포함한다. 하나의 예에서, 획득된 정보는, 복수의 선택된 오브젝트들의 각각에 대해, 그 오브젝트와 연관되는 라벨을 포함하고, 검색 엔진 (Z300) 은 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 라벨들을 검색하도록 구현된다. 부가적으로 또는 대안적으로, 검색 엔진 (Z300) 은 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 후보 기하학적 배열과 연관되는 라벨을 검색하도록 구현될 수도 있다.
- [0288] 도 45c 는 다른 일반적 구성에 따른 이미지 취출을 위한 장치 (MF500) 의 블록 다이어그램을 도시한다. 장치 (MF500) 는 복수의 선택된 오브젝트들을 식별하는 정보를 획득하는 수단 (Y100) 및 복수의 후보 기하학적 배열들 중에서 하나를 선택하는 수단 (Y200) 을 포함한다. 장치 (MF500) 는 또한, 복수의 디지털 이미지들 중에서, 선택된 후보 기하학적 배열에 따라 배열된 복수의 선택된 오브젝트들을 나타내는 적어도 하나의 디지털 이미지를 식별하는 수단 (Y300) 을 포함한다. 하나의 예에서, 획득된 정보는 복수의 선택된 오브젝트들의 각각에 대해, 그 오브젝트와 연관되는 라벨을 포함하고, 수단 (Y300) 은 복수의 디지털 이미지들과 연관되는 메타데이터 내에서 라벨들을 검색하도록 구현된다. 부가적으로 또는 대안적으로, 수단 (Y300) 은 복수의 디지털 이미지들과 연관되는 메타데이터 내에서, 선택된 후보 기하학적 배열과 연관되는 라벨을 검색하도록 구현될 수도 있다.
- [0289] 도 40a 는 일반적 구성에 따라 메타데이터를 발생시키기 위한 장치 (MF100) 의 블록 다이어그램을 도시한다. 장치 (MF100) 는 복수의 오브젝트들을 선택하기 위해 적어도 하나의 디지털 이미지를 이용하는 수단 (F100) 을 포함하고, 여기서 적어도 하나의 디지털 이미지는 (예를 들어, 태스크 T100 에 관하여 본 명세서에서 설명한 바와 같이) 물리적 공간에 관하여 오브젝트들을 나타낸다. 장치 (MF100) 는 또한, 로케이션 공간에서의 선택된 오브젝트들의 포지션들을 나타내는 정보에 기초하여, (예를 들어, 태스크 T300 에 관하여 본 명세서에서 설명한 바와 같이) 선택된 오브젝트들의 복수의 후보 기하학적 배열들 중에서 하나를 식별하는 메타데이터를 생성하는 수단 (F300) 을 포함한다.
- [0290] 도 40b 는 장치 (MF100) 의 구현 (MF110) 의 블록 다이어그램을 도시한다. 장치 (MF110) 는 (예를 들어, 디스플레이 태스크 T50 에 관하여 본 명세서에서 설명한 바와 같이) 적어도 하나의 디지털 이미지를 디스플레이 하는 수단 (F50) 을 포함한다.

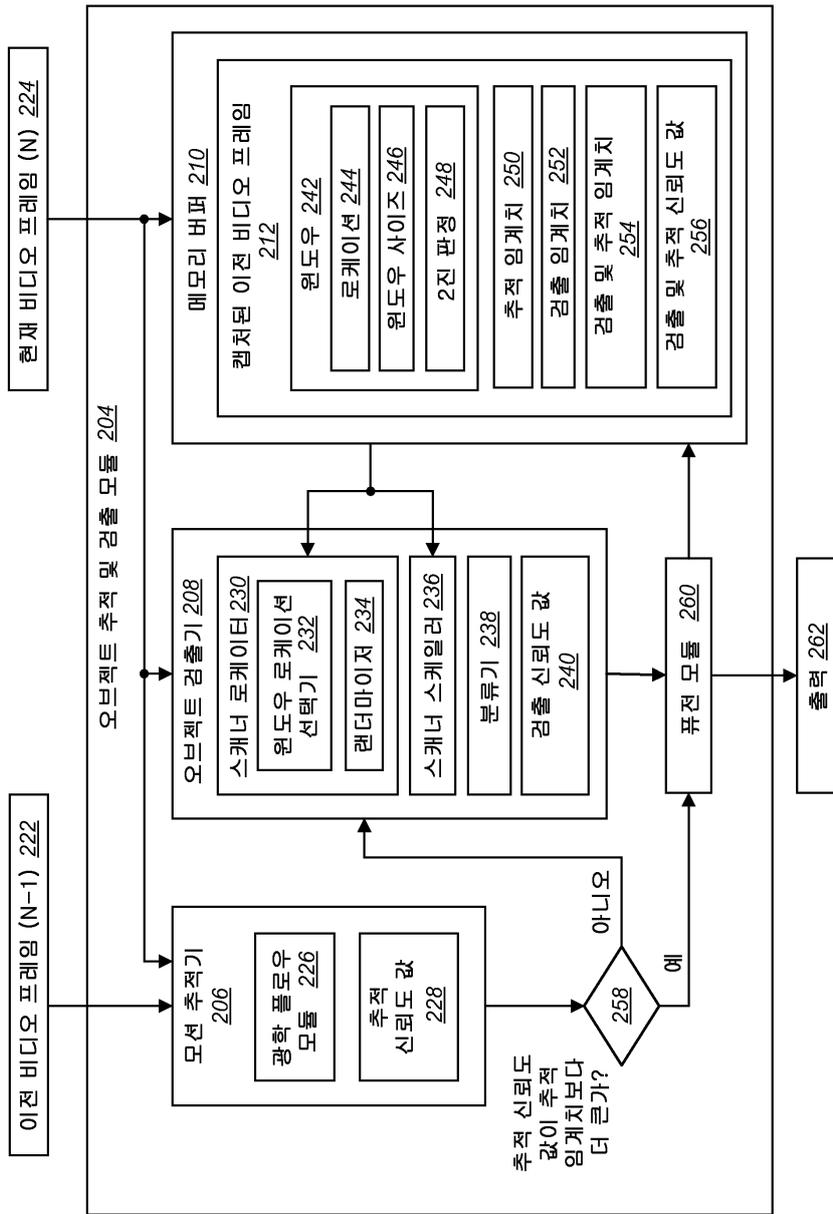
- [0291] 도 40c 는 장치 (MF100) 의 구현 (MF120) 의 블록 다이어그램을 도시한다. 장치 (MF120) 는 (예를 들어, 파싱 태스크 T30 에 관하여 본 명세서에서 설명한 바와 같이) 적어도 하나의 디지털 이미지를 생성하기 위해 비디오 스트림을 파싱하는 수단 (F30) 을 포함한다. 도 40d 는 장치 (MF110 및 MF120) 의 구현 (MF130) 의 블록 다이어그램을 도시한다.
- [0292] 도 40e 는 장치 (MF100) 의 구현 (MF200) 의 블록 다이어그램을 도시한다. 장치 (MF200) 는 (예를 들어, 태스크 T200 에 관하여 본 명세서에서 설명한 바와 같이) 로케이션 공간에서 복수의 오브젝트들에 대한 레퍼런스 포지션을 결정하는 수단 (F200) 을 포함한다. 도 41b 는 (예를 들어, 태스크 T200 에 관하여 본 명세서에서 설명한 바와 같이) 비디오 스트림 내에서, 상기 복수의 오브젝트들을 추적하는 수단 (F210) 을 포함하는 장치 (MF100) 의 구현 (MF400) 의 블록 다이어그램을 도시한다. 도 40f, 도 40g, 및 도 41a 는 수단 (F200) 의 인스턴스들을 포함하는, 장치 (MF110, MF120, 및 MF130) 각각의 구현들 (MF210, MF220, 및 MF230) 의 블록 다이어그램들을 도시한다. 이러한 장치는 또한, 수단 (F200) 에 추가하여 또는 대안적으로 수단 (F210) 을 포함하도록 구현될 수도 있다.
- [0293] 도 41c 는 장치 (MF100) 의 구현 (MF300) 의 블록 다이어그램을 도시한다. 장치 (MF200) 는 (예를 들어, 태스크 T300 에 관하여 본 명세서에서 설명한 바와 같이) 비디오 스트림의 대응하는 프레임들과 수단 (F300) 에 의해 생성된 메타데이터를 연관시키는 수단 (F400) 을 포함한다. 도 41d 및 도 41e 는 수단 (F400) 의 인스턴스들을 포함하는, 장치 (MF200 및 MF230) 각각의 구현들 (MF310 및 MF330) 의 블록 다이어그램들을 도시한다.
- [0294] 당업자들은, 본 명세서에 개시된 실시형태들과 관련되어 설명된 다양한 예시적인 논리 블록들, 구성들, 모듈들, 회로들, 및 알고리즘 단계들은 전자 하드웨어, 프로세서에 의해 실행되는 컴퓨터 소프트웨어, 또는 이들 양쪽의 조합들로서 구현될 수도 있다는 것을 추가로 인식할 것이다. 예를 들어, 본 명세서에 개시된 장치 (예를 들어, 장치 (MF100, MF110, MF120, MF130, MF200, MF210, MF220, MF230, MF300, MF310, MF330, MF400, MF500, A100, A110, A120, A130, A200, A210, A220, A230, A300, A310, A330, A500, 및 A540) 중 임의의 장치) 의 다양한 구현들의 하나 이상의 엘리먼트들은 마이크로프로세서들, 임베딩된 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA들 (field-programmable gate arrays), ASSP들 (application-specific standard products), 및 ASIC들 (application-specific integrated circuits) 과 같은 로직 엘리먼트들 (예를 들어, 프로세서 (2610)) 의 하나 이상의 고정된 또는 프로그램가능 어레이들 상에서 실행하도록 배열된 명령들의 하나 이상의 세트들로서 전체적으로 또는 부분적으로 구현될 수도 있다. 다양한 예시적인 컴포넌트들, 블록들, 구성들, 모듈들, 회로들, 및 단계들은 이들의 기능성의 관점에서 일반적으로 상술되었다. 이러한 기능성이 하드웨어로서 또는 프로세서 실행가능 명령들로서 구현되는지 여부는 전체 시스템에 부과되는 설계 제약들 및 특정 애플리케이션에 의존한다. 당업자들은 각각의 특정 애플리케이션에 대한 다양한 방법들로 설명된 기능성을 구현할 수도 있지만, 이러한 구현 관점들은 본 개시물의 범위로부터 벗어남을 야기시키는 것으로 해석되어서는 안 된다.
- [0295] 본 명세서에 개시된 실시형태들과 관련되어 설명된 방법 또는 알고리즘의 단계들은 직접 하드웨어로, 프로세서에 의해 실행되는 소프트웨어 모듈로, 또는 이들 둘의 조합으로 구현될 수도 있다. 소프트웨어 모듈은 랜덤 액세스 메모리 (RAM), 플래시 메모리, 판독-전용 메모리 (ROM), 프로그램가능 판독-전용 메모리 (PROM), 소거가능 프로그램가능 판독-전용 메모리 (EPROM), 전기적 소거가능 프로그램가능 판독-전용 메모리 (EEPROM), 레지스터들, 하드 디스크, 착탈식 디스크, 콤팩트 디스크 판독-전용 메모리 (CD-ROM), 또는 당업계에 공지된 임의의 다른 형태의 비순시적 저장 매체에 상주할 수도 있다. 예시적인 저장 매체는, 프로세서가 저장 매체로부터 정보를 판독할 수 있고 정보를 저장 매체에 기입할 수 있도록 프로세서에 커플링된다. 대안적으로, 저장 매체는 프로세서와 일체적일 수도 있다. 프로세서 및 저장 매체는 주문형 집적 회로 (ASIC) 에 상주할 수도 있다. ASIC 는 컴퓨팅 디바이스 또는 사용자 단말기에 상주할 수도 있다. 대안으로, 프로세서 및 저장 매체는 컴퓨팅 디바이스 또는 사용자 단말기에 개별 컴포넌트들로서 상주할 수도 있다.
- [0296] 개시된 실시형태들의 이전 설명은 당업자가 개시된 실시형태들을 제조하거나 이용하는 것을 가능하게 하도록 제공된다. 이들 실시형태들에 대한 다양한 수정들은 당업자들에게 쉽게 명백할 것이며, 본 명세서에 정의된 원리들은 본 개시물의 범위로부터 벗어남이 없이 다른 실시형태들에 적용될 수도 있다. 따라서, 본 개시물은 본 명세서에 도시된 실시형태들로 제한되도록 의도된 것이 아니라, 다음의 청구항들에 의해 정의된 원리들 및 신규한 피처들과 일치하는 가능한 가장 넓은 범위를 따르도록 하기 위한 것이다.

도면

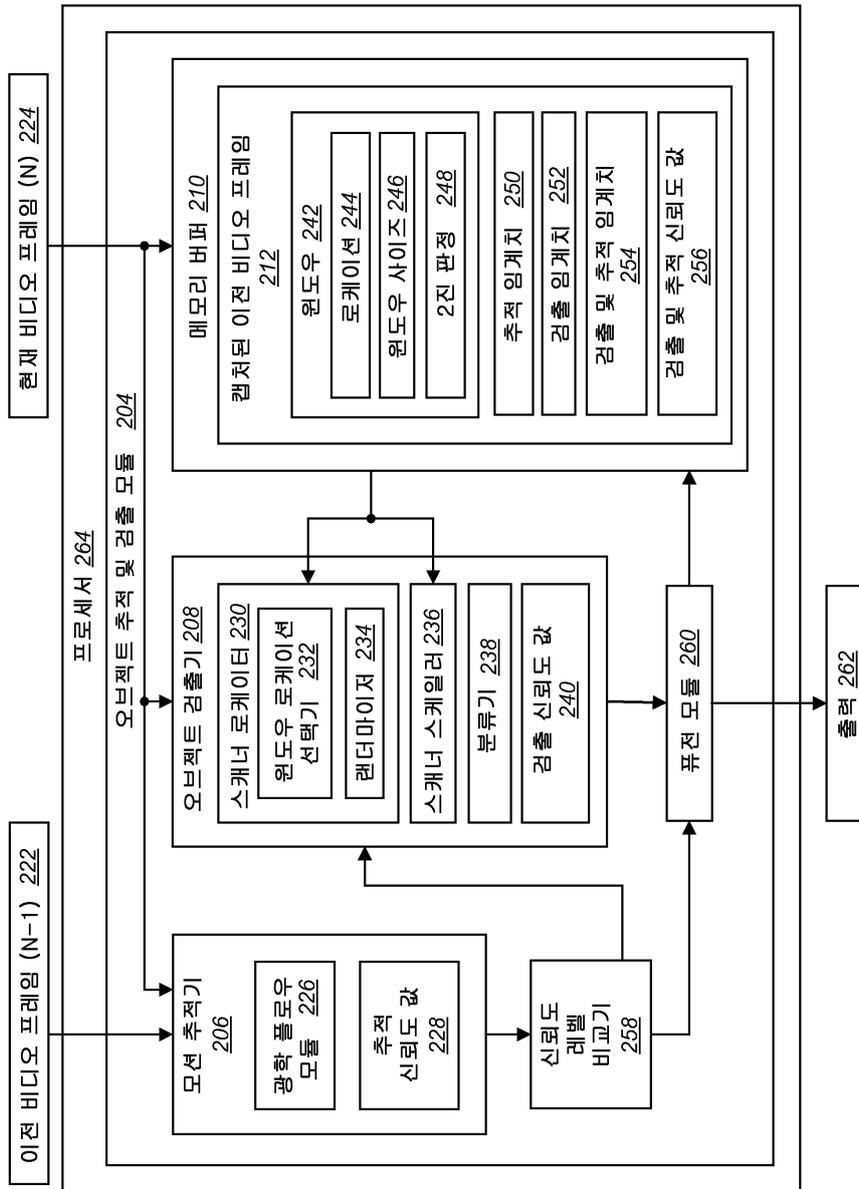
도면1



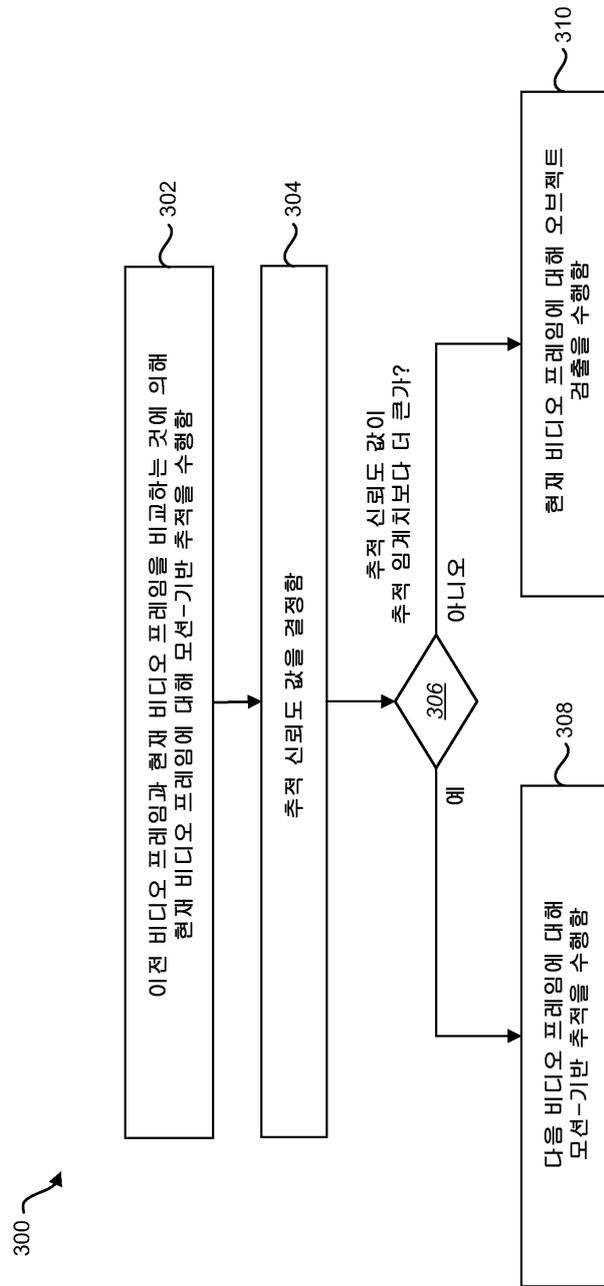
도면2a



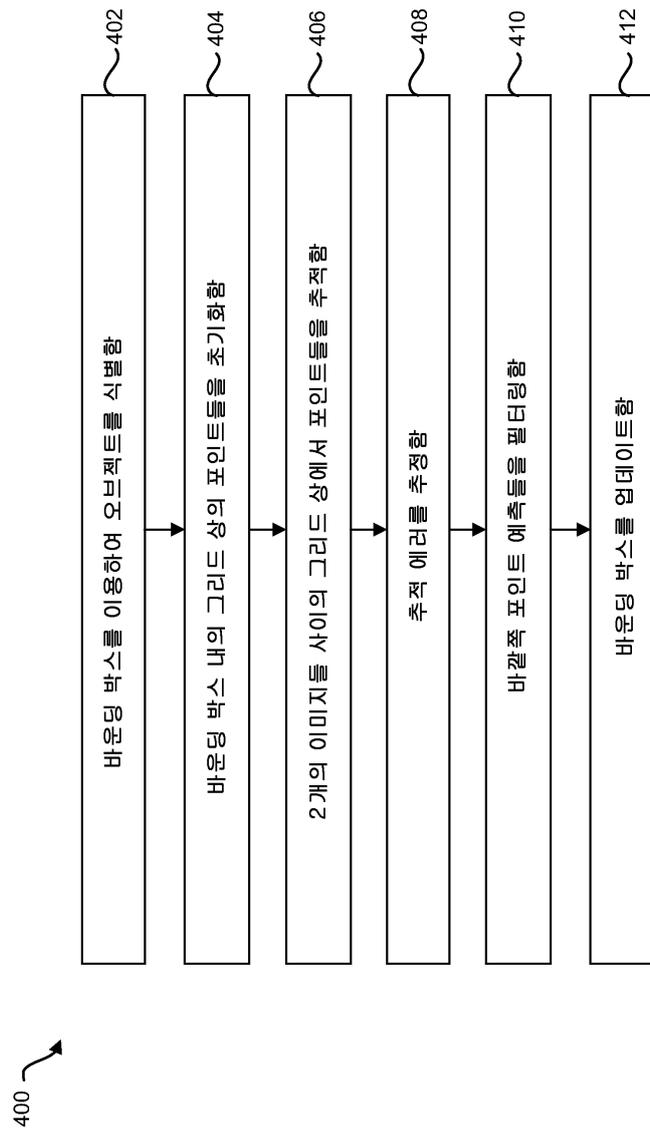
도면2b



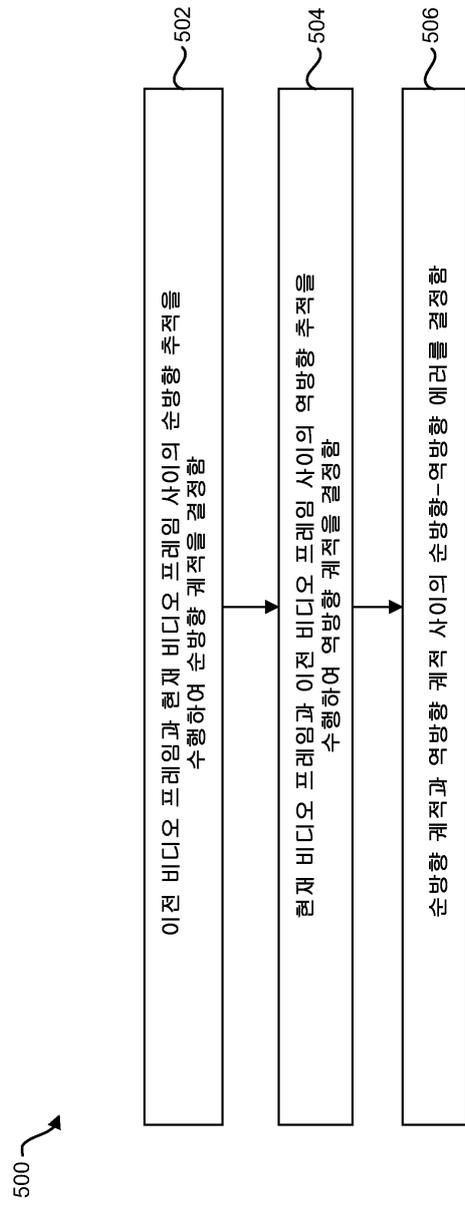
도면3



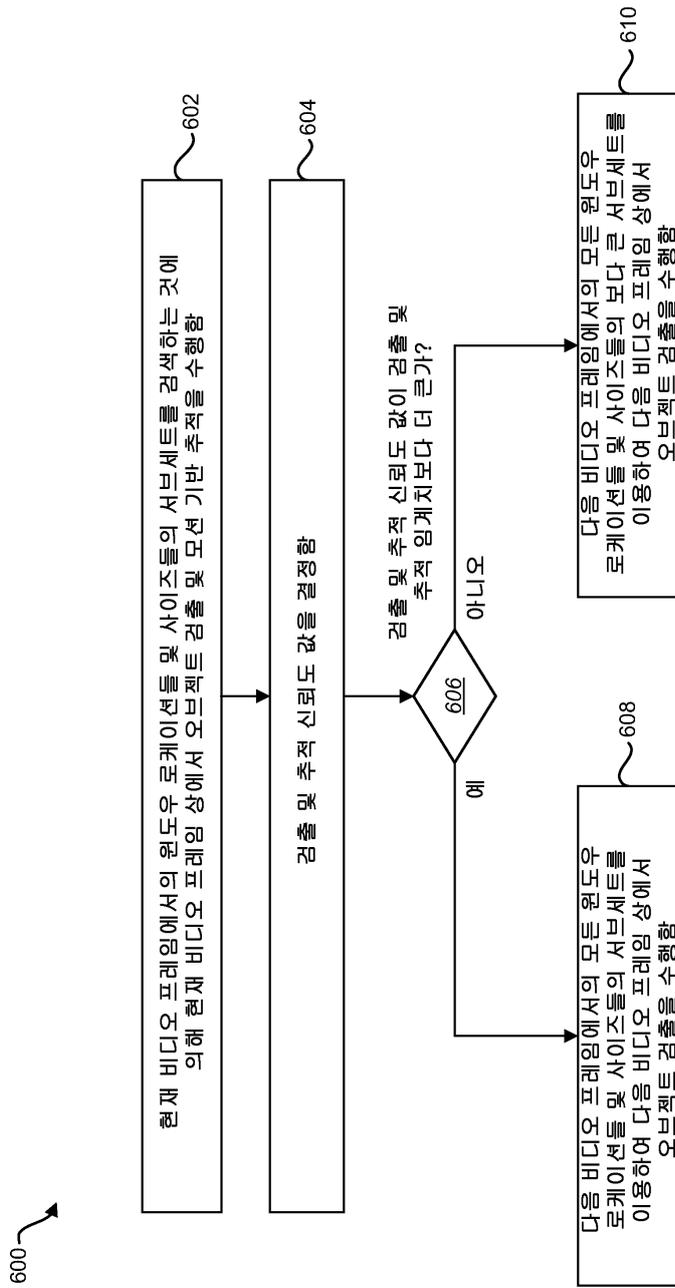
도면4



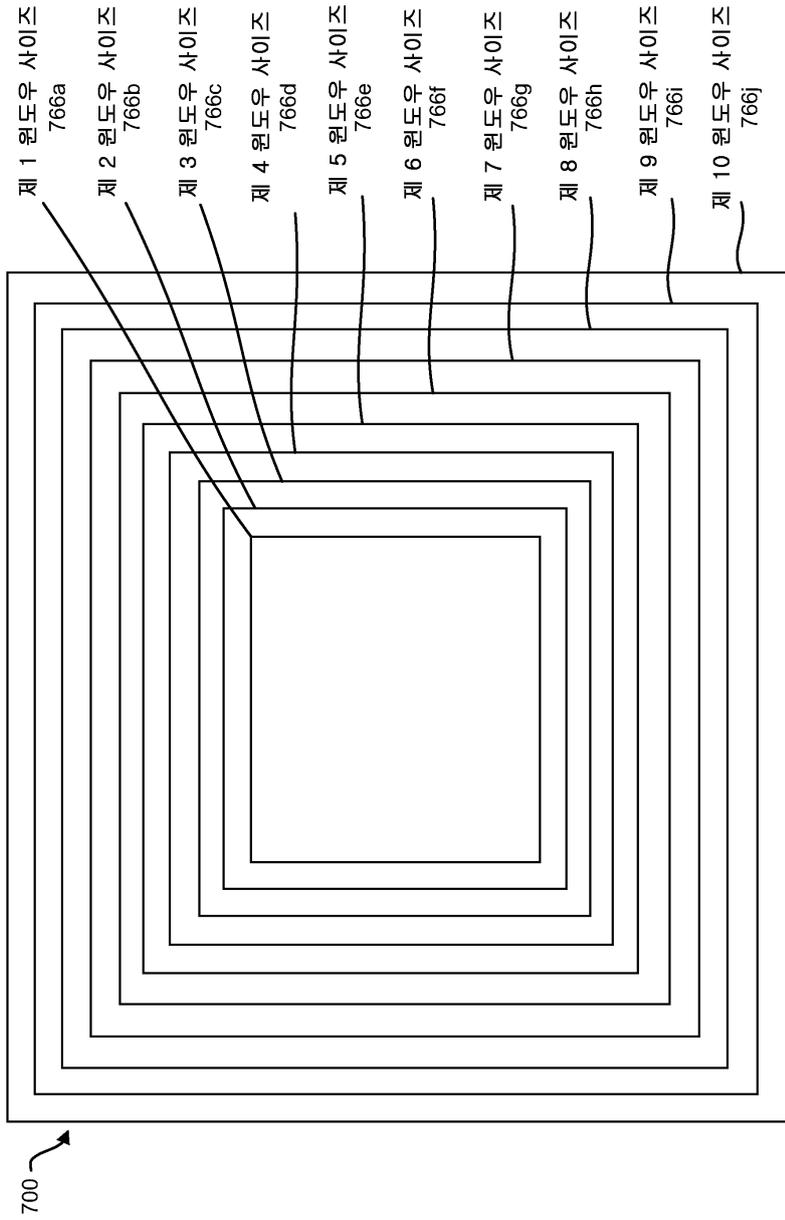
도면5



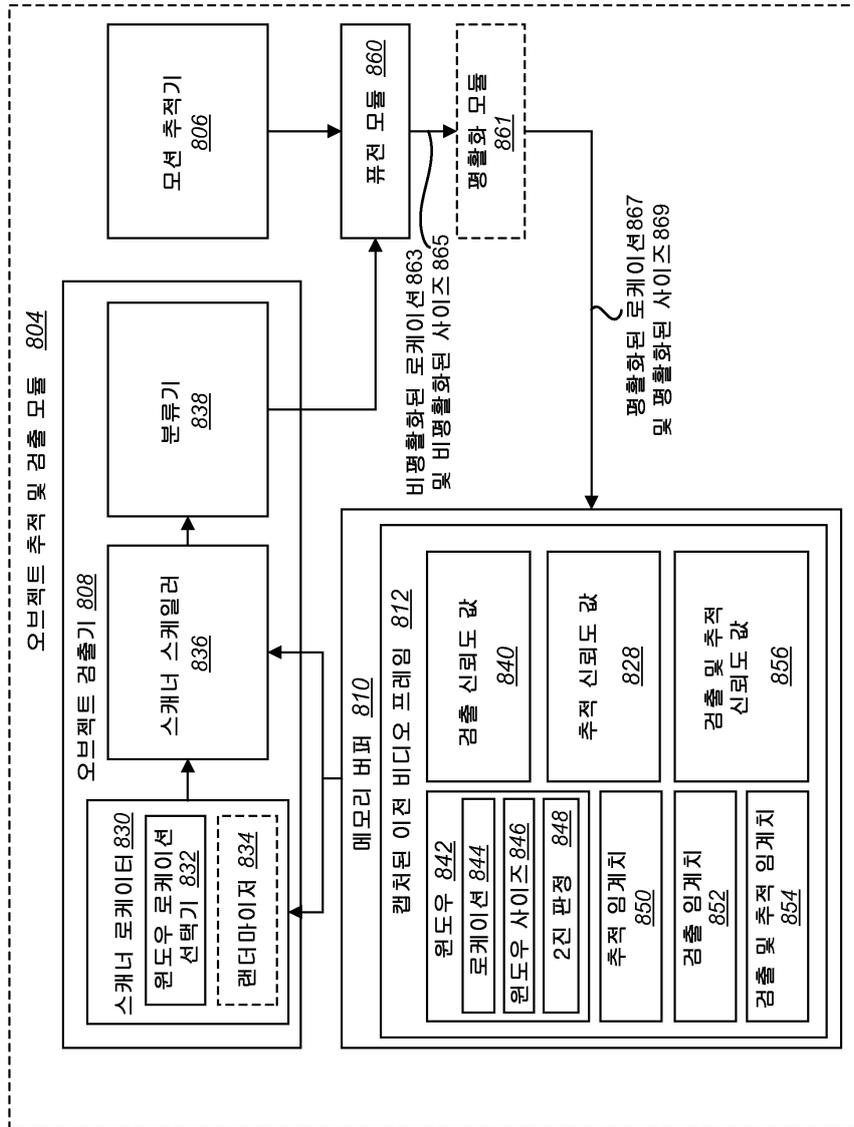
도면6



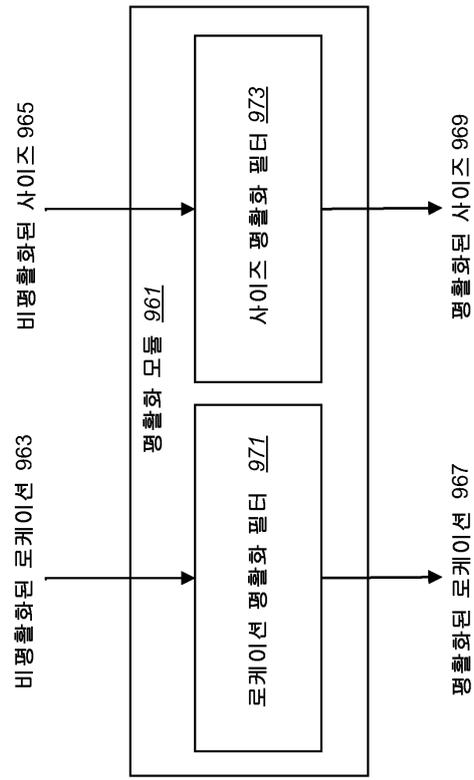
도면7



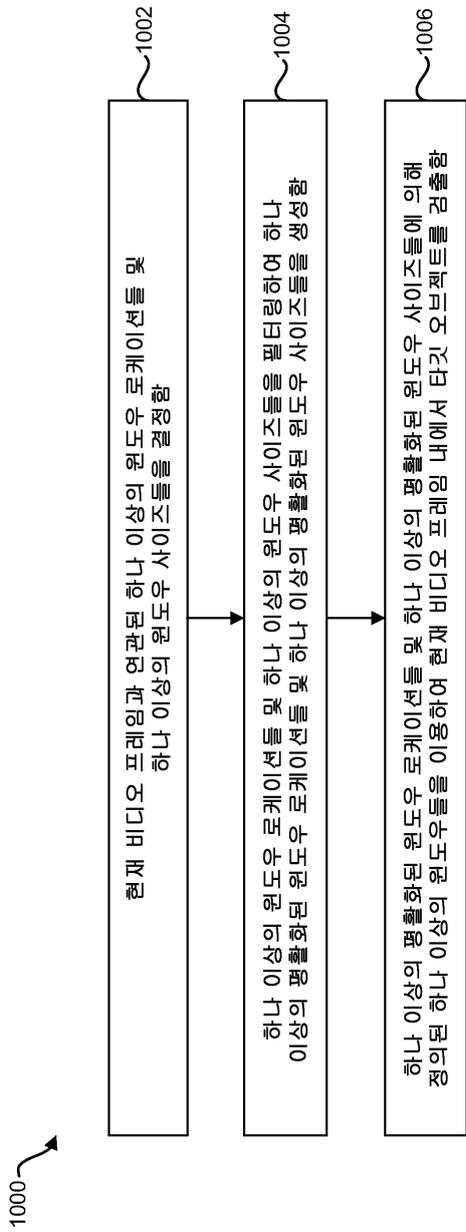
도면8



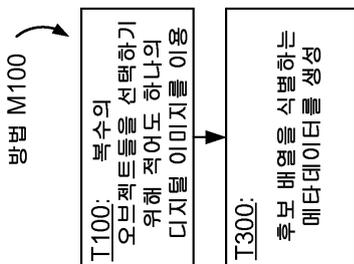
도면9



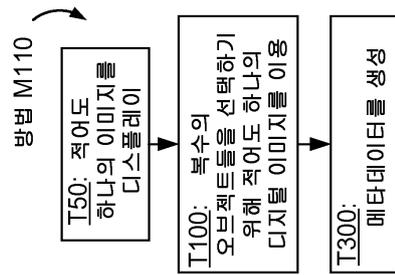
도면10



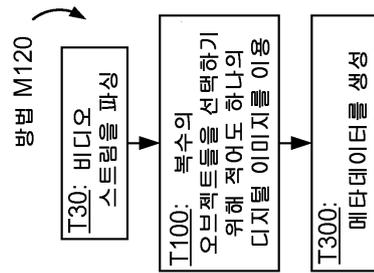
도면11a



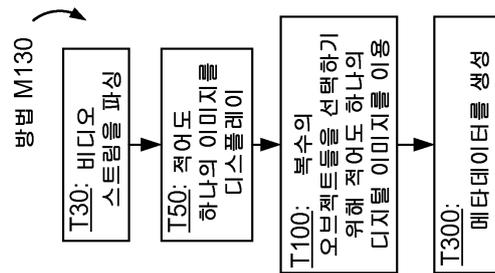
도면11b



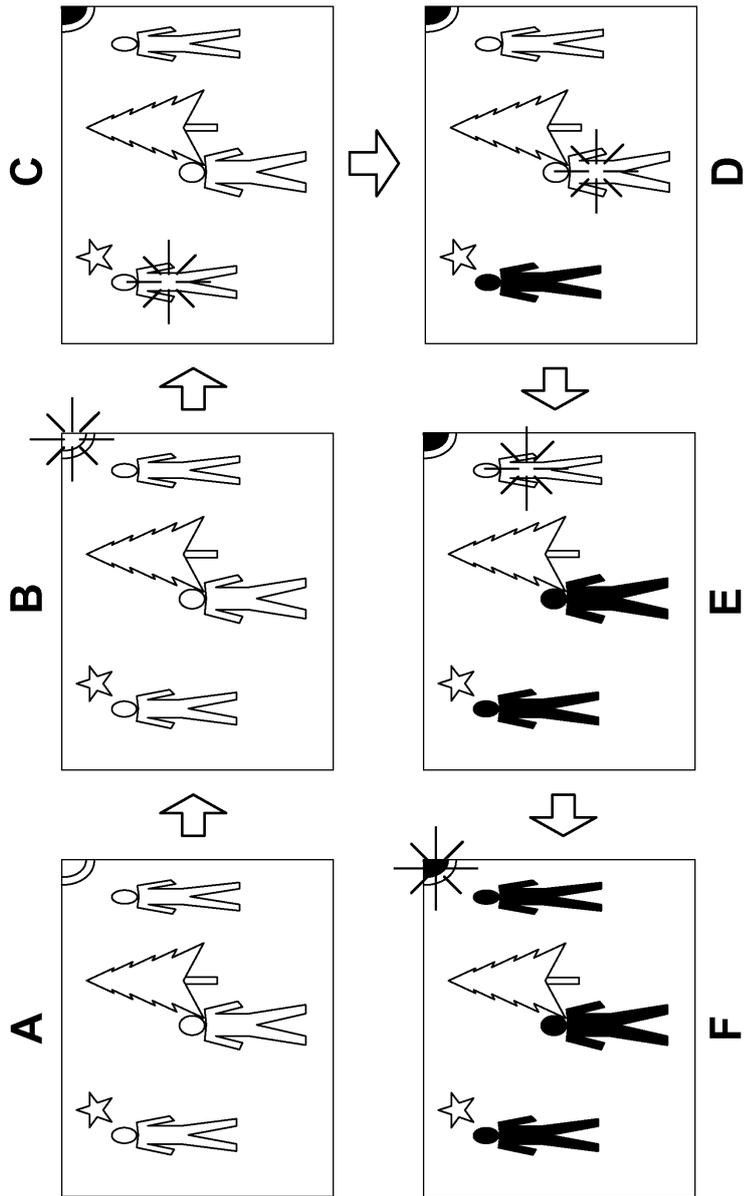
도면11c



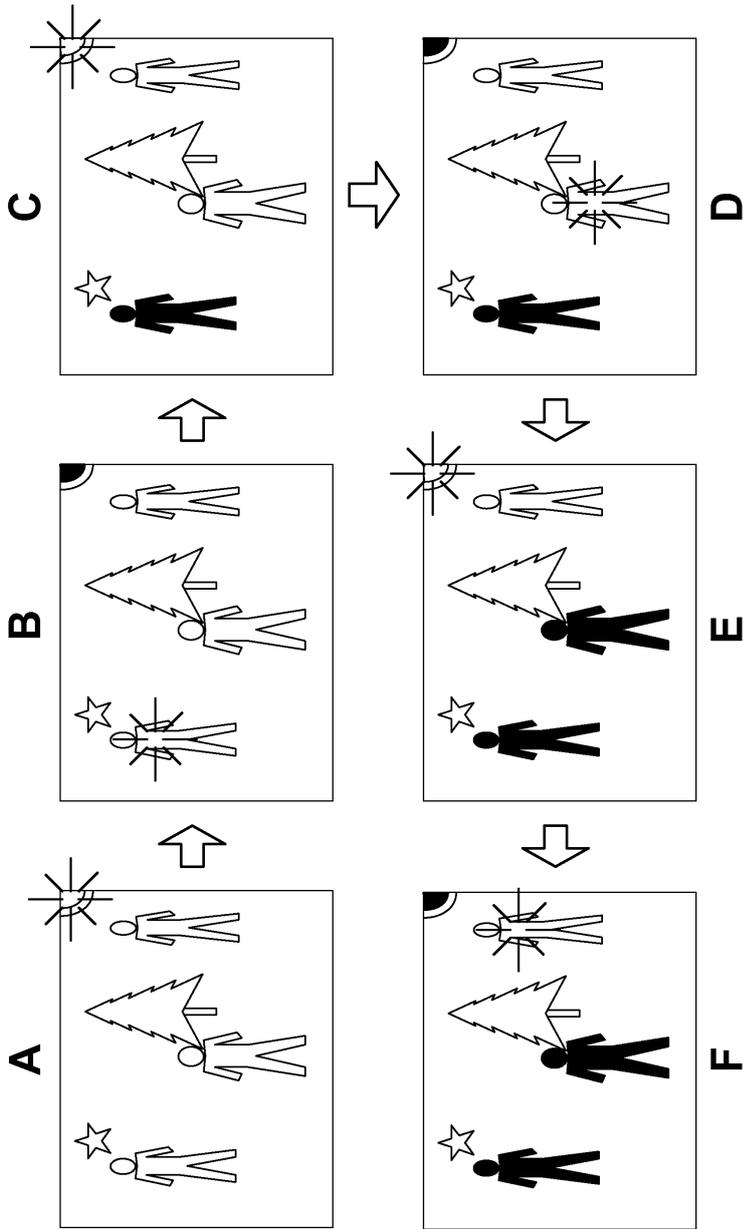
도면11d



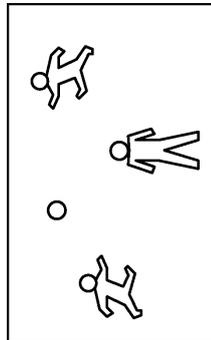
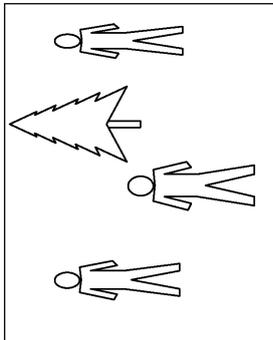
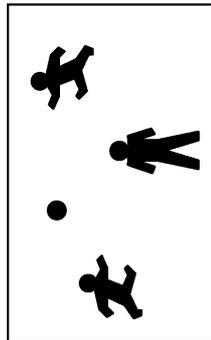
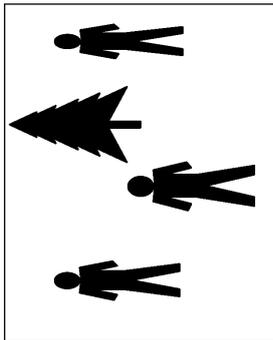
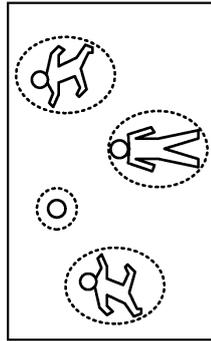
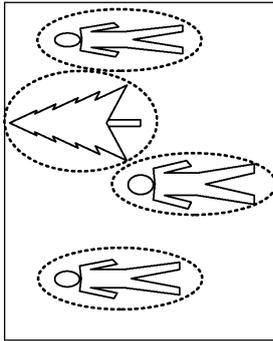
도면12



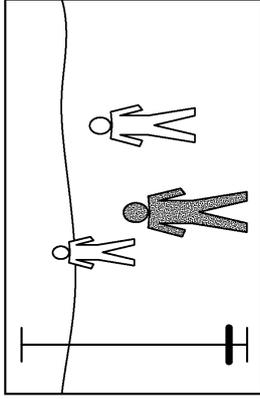
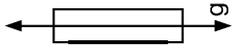
도면13



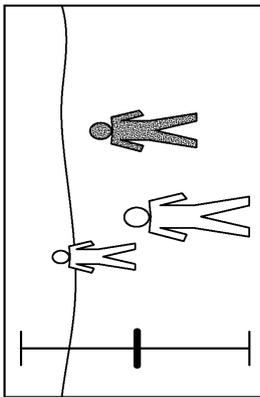
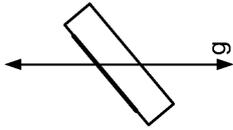
도면14



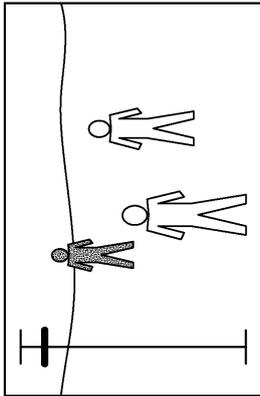
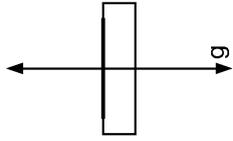
도면15a



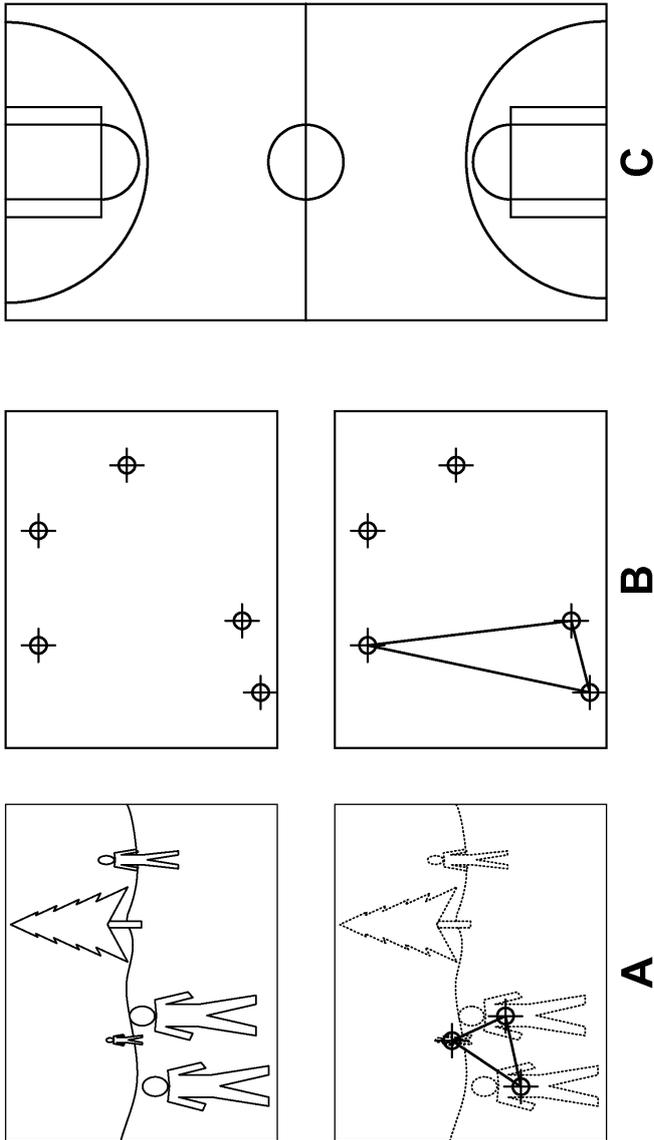
도면15b



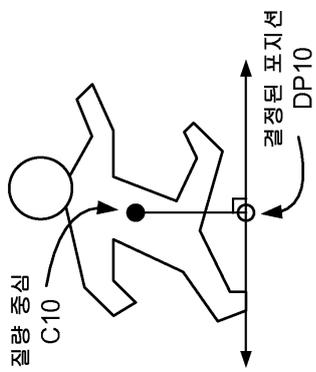
도면15c



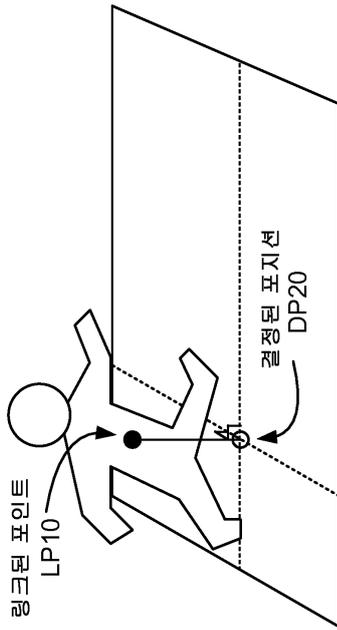
도면16



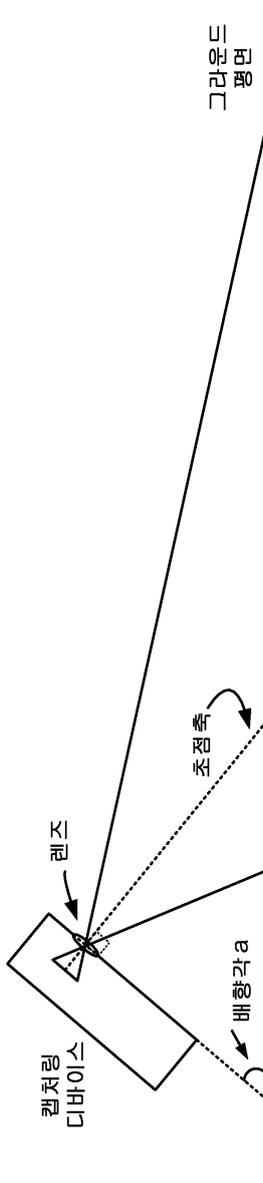
도면17a



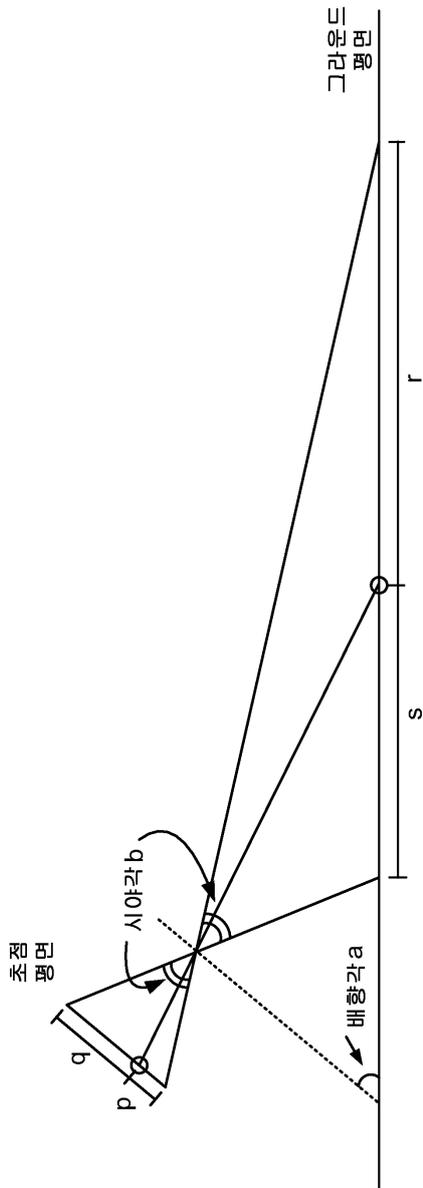
도면17b



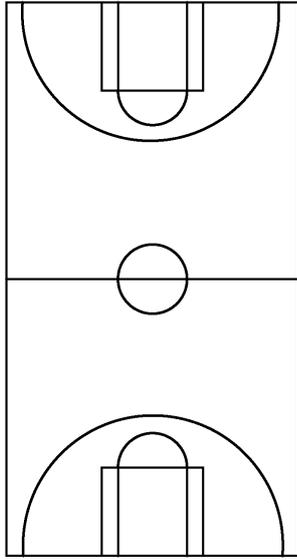
도면18a



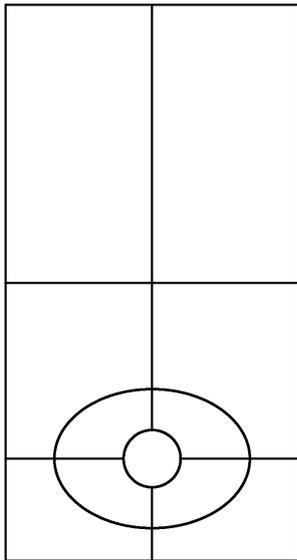
도면18b



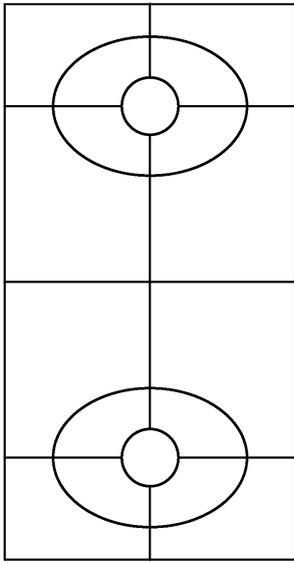
도면19a



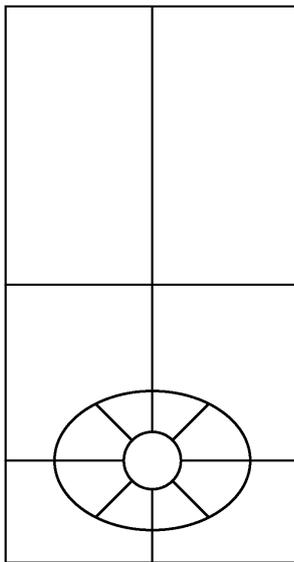
도면19b



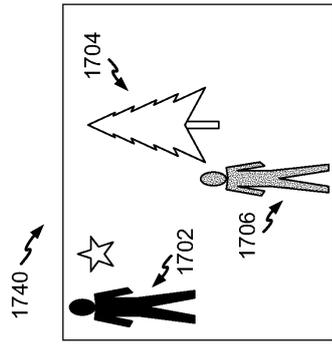
도면19c



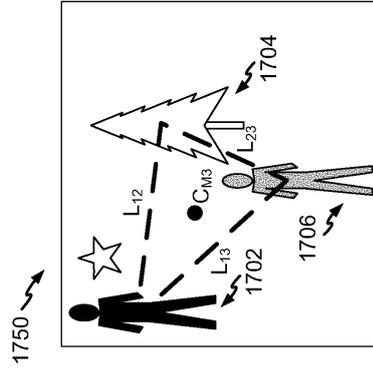
도면19d



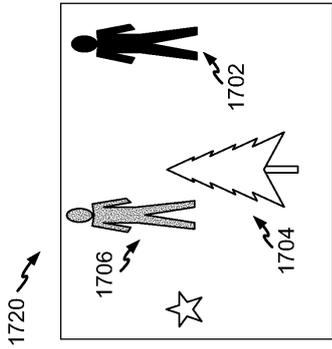
도면20



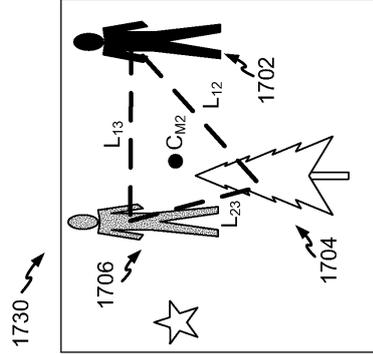
제 3 시간 인스턴스에서의 장면



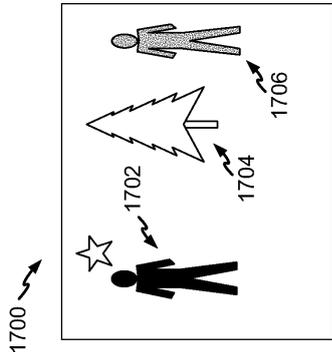
스크린 (프레임 3)



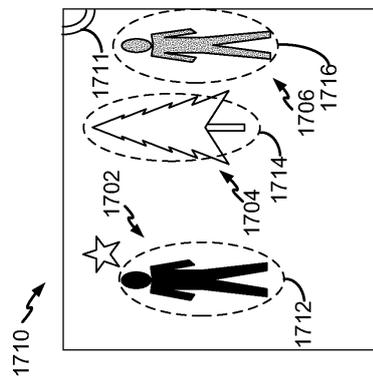
제 2 시간 인스턴스에서의 장면



스크린 (프레임 2)

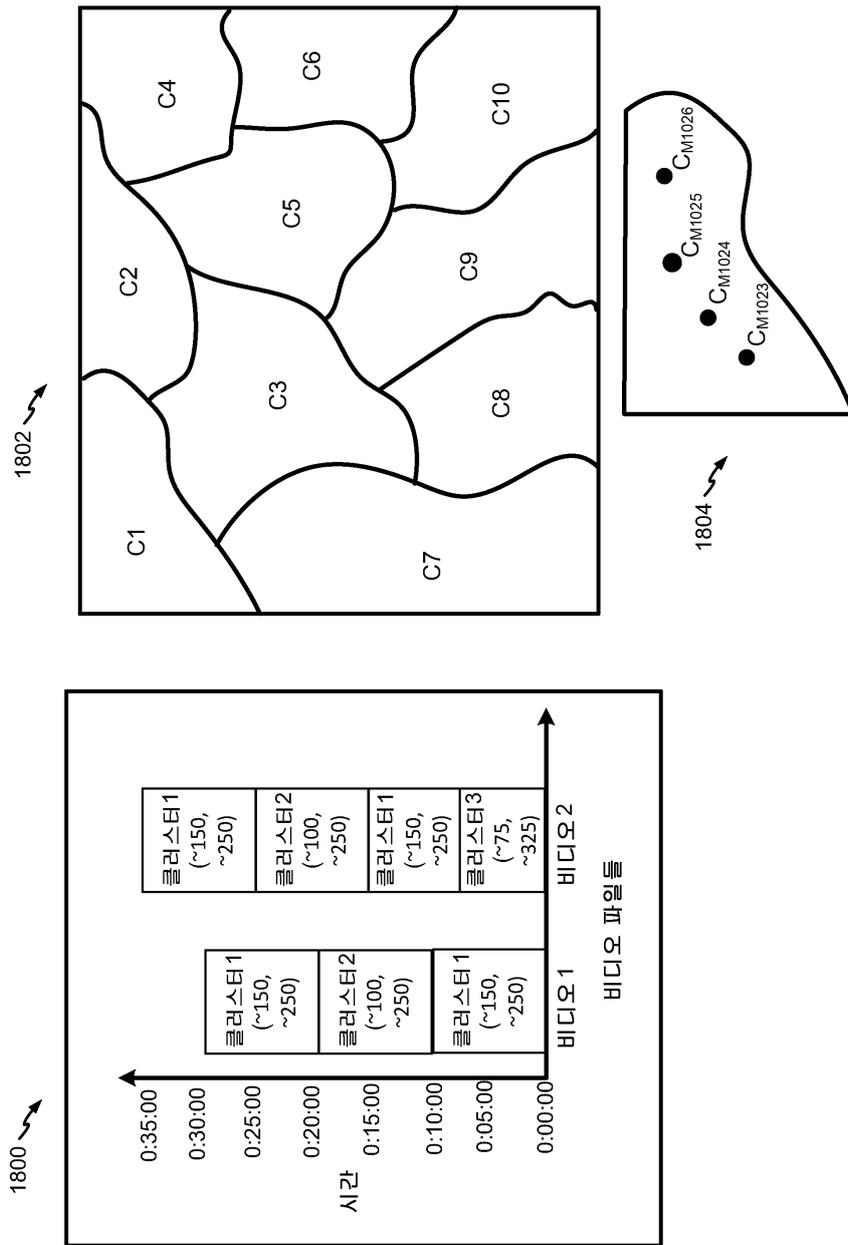


제 1 시간 인스턴스에서의 장면

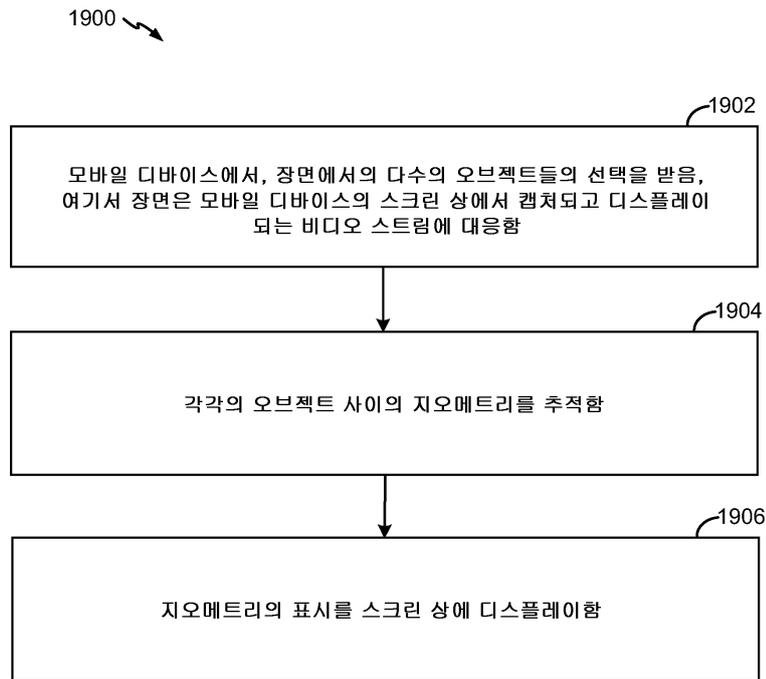


스크린 (프레임 1)

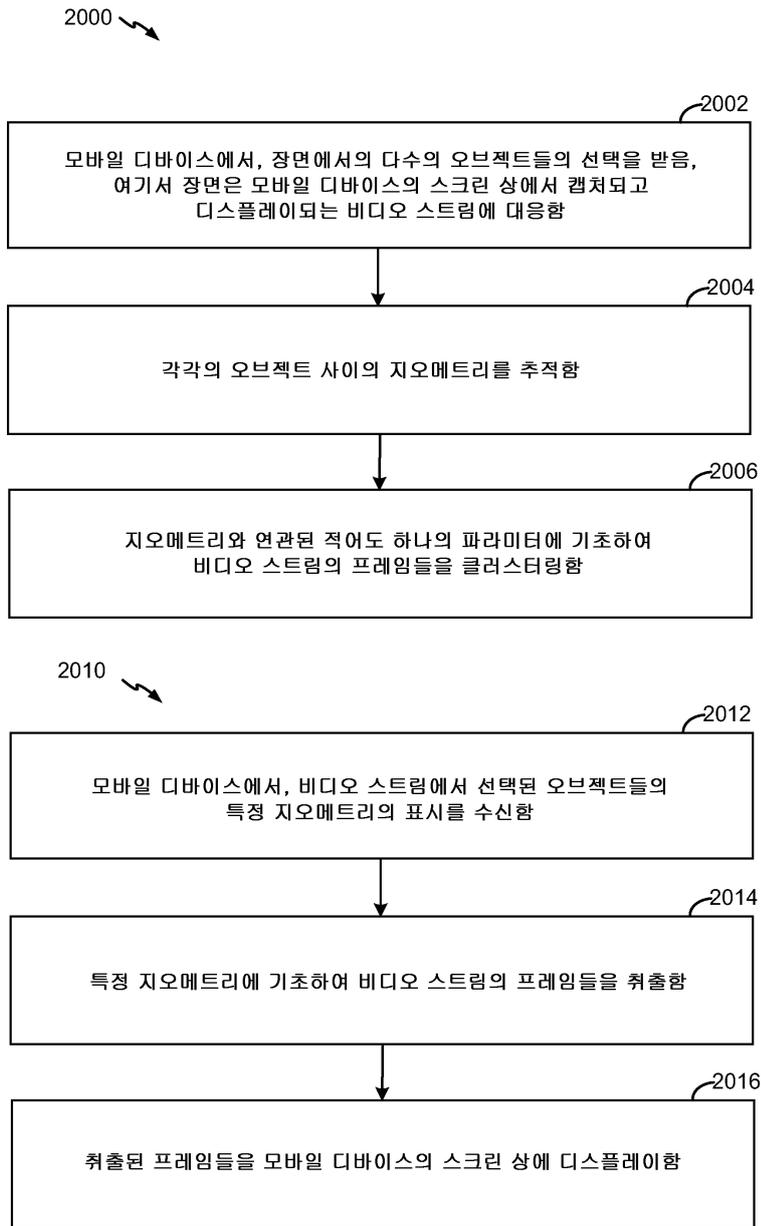
도면21



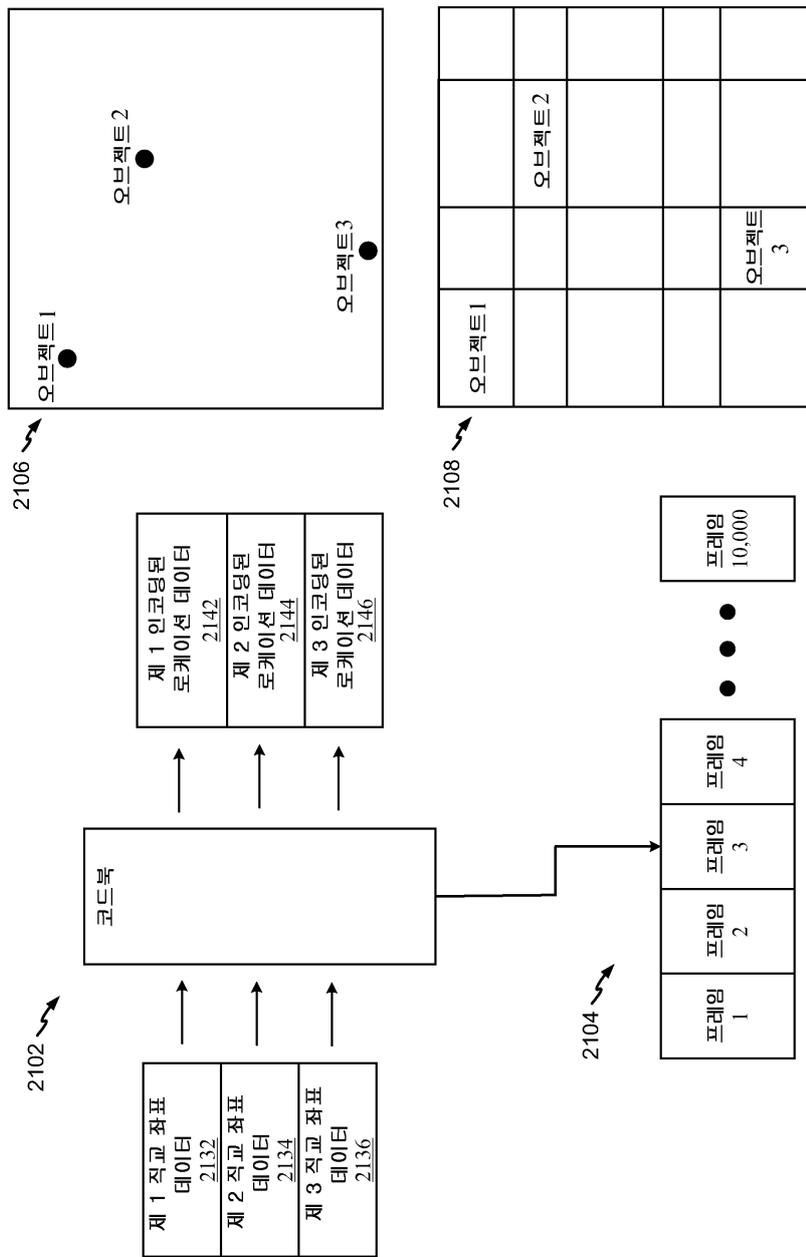
도면22



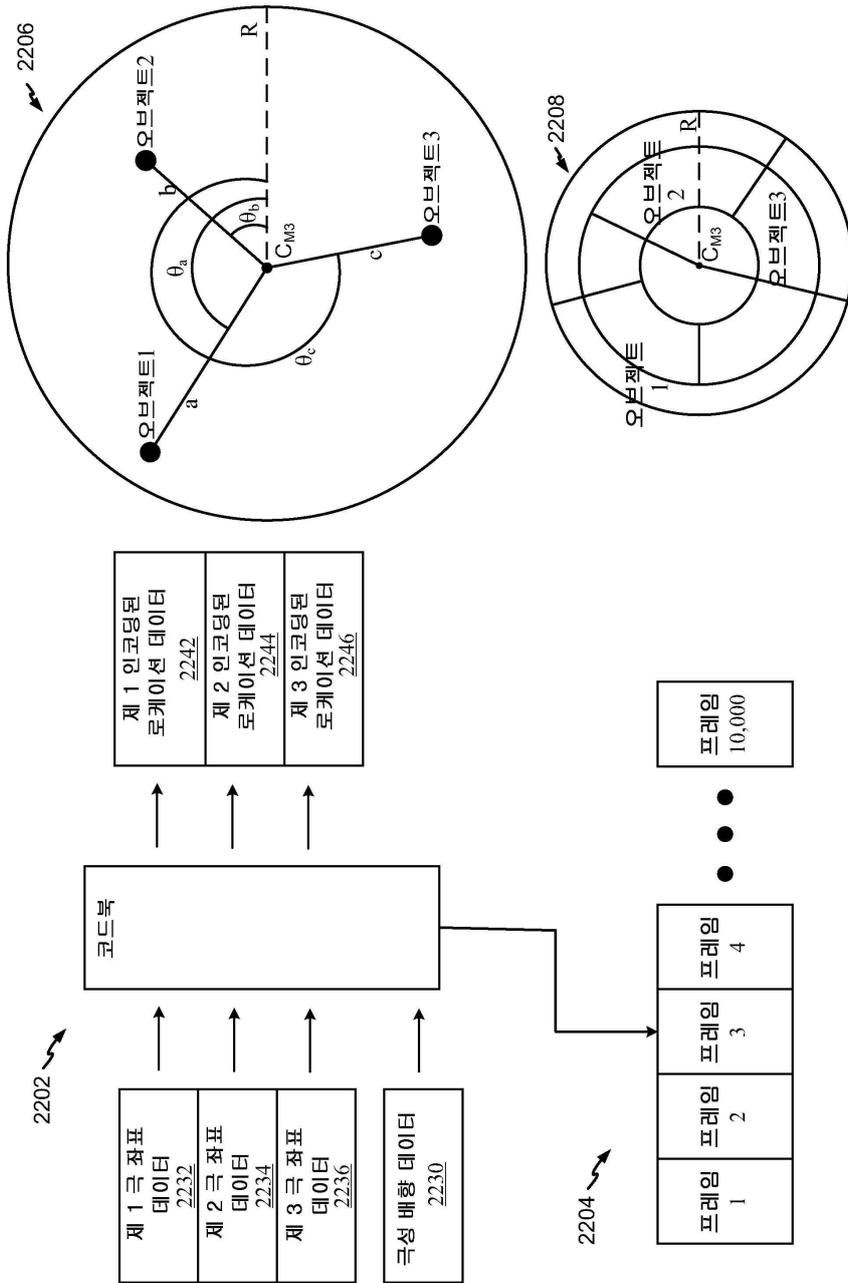
도면23



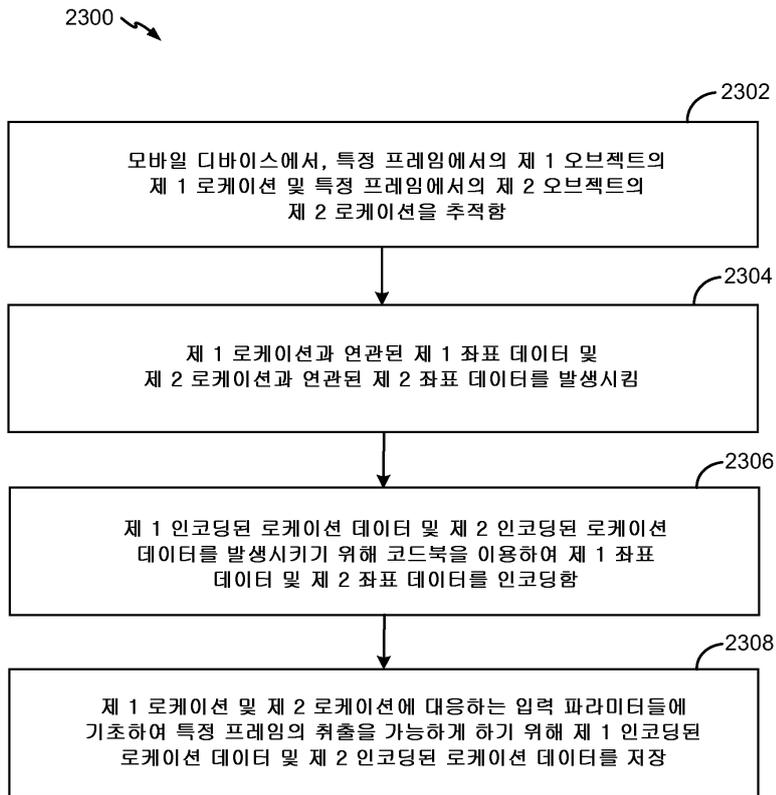
도면24



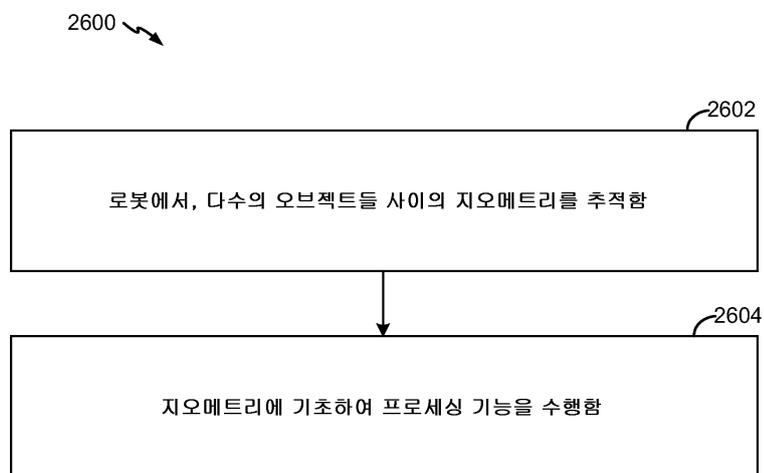
도면25



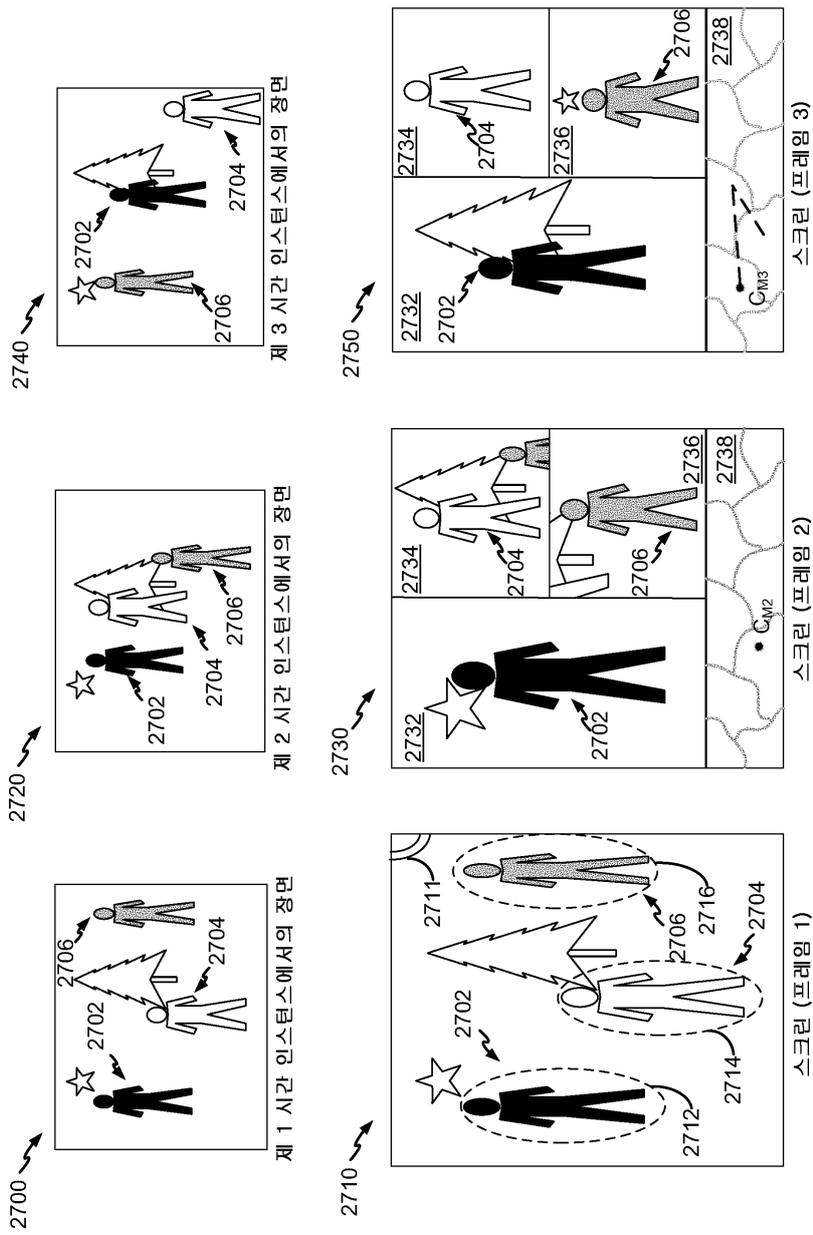
도면26



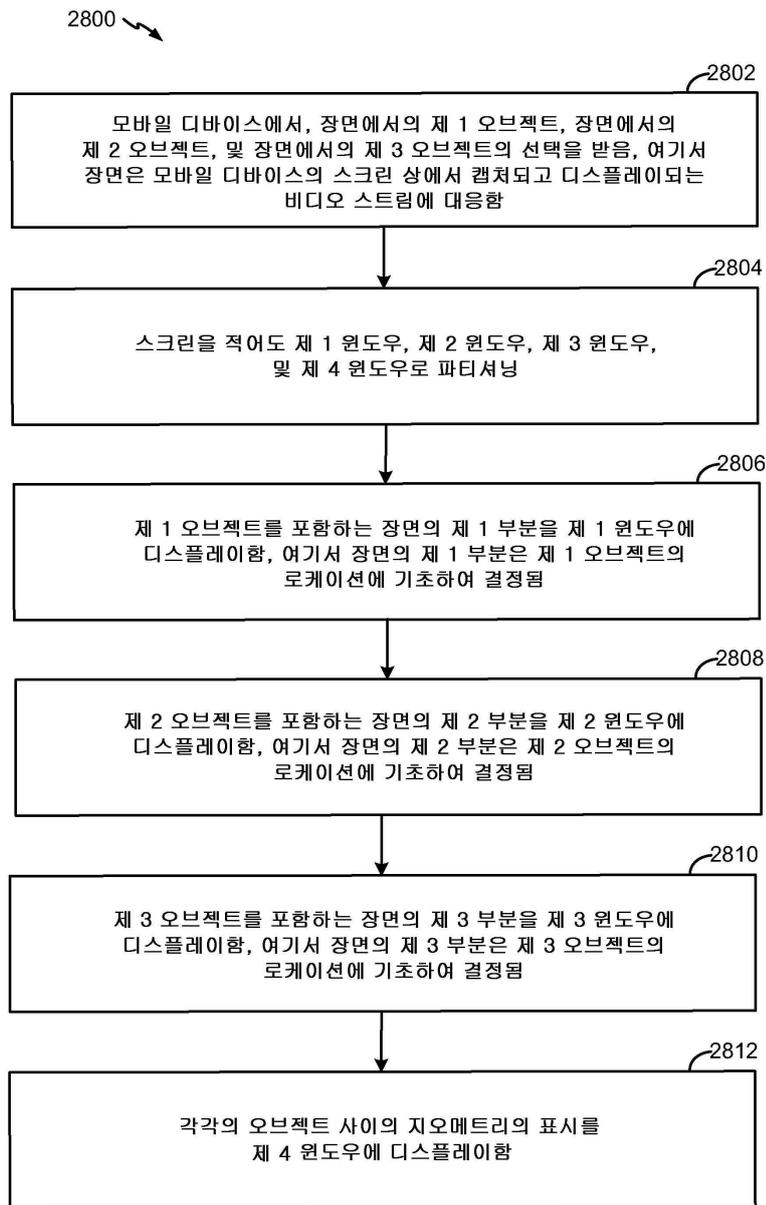
도면27



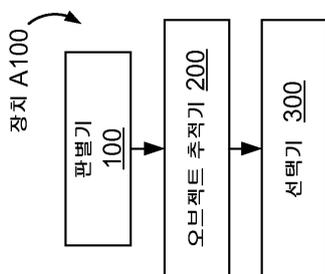
도면28



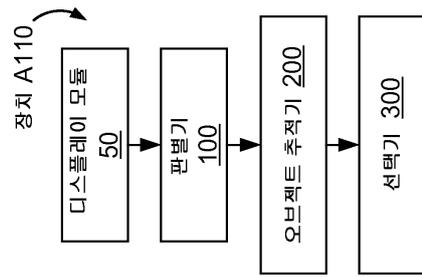
도면29



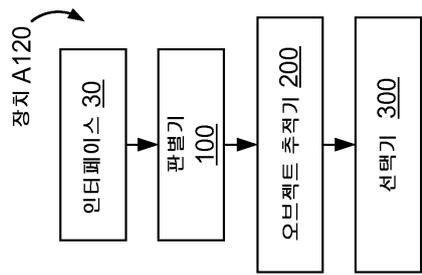
도면30a



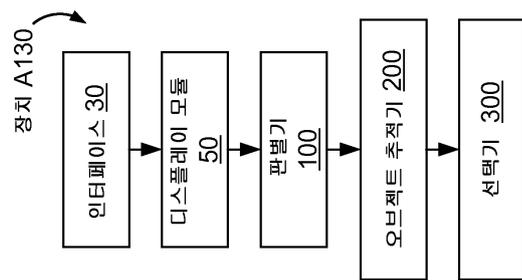
도면30b



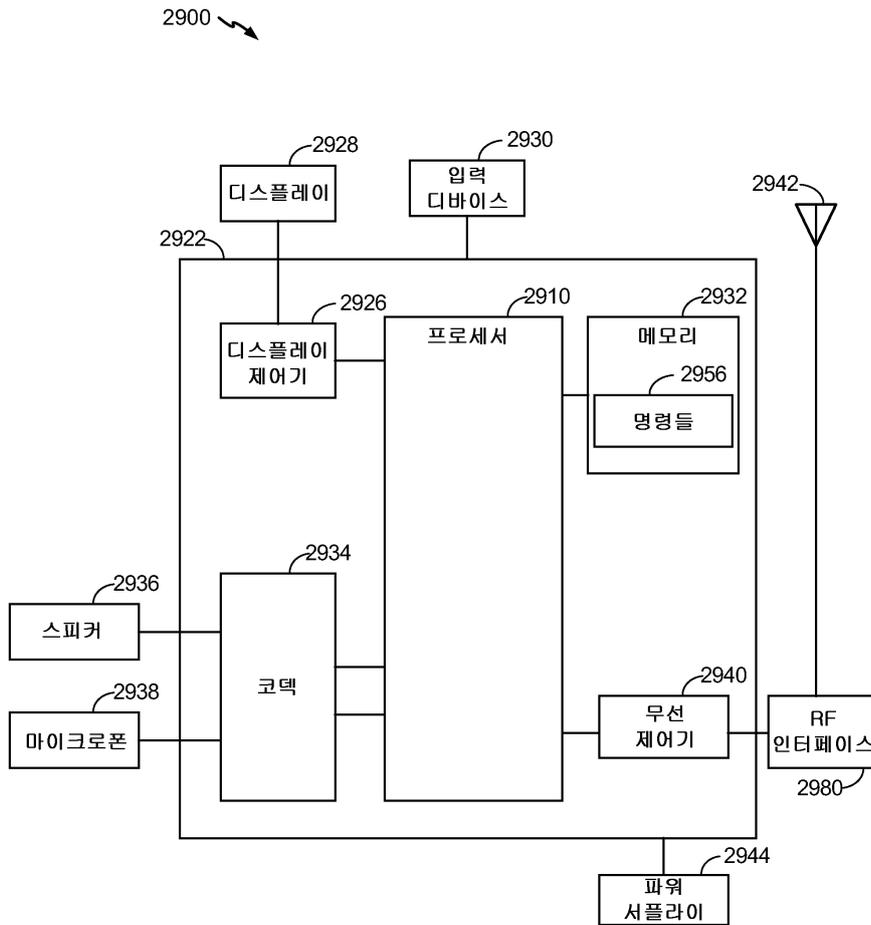
도면30c



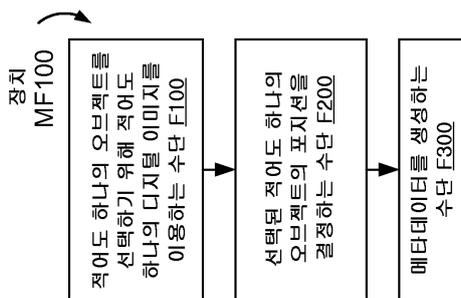
도면30d



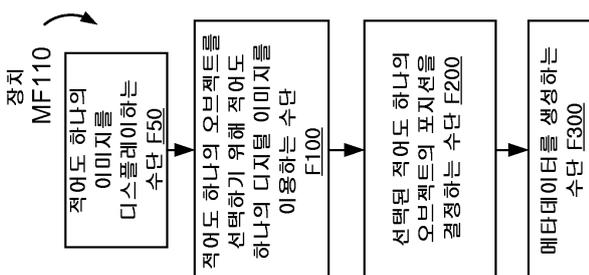
도면31



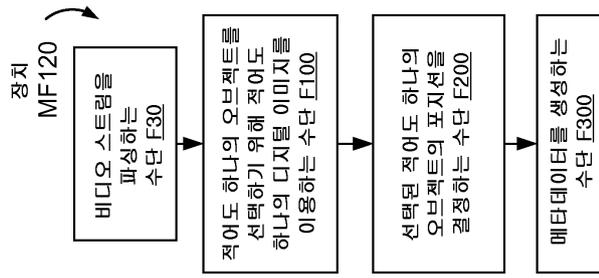
도면32a



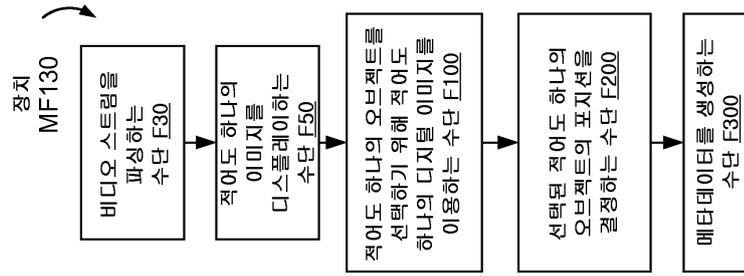
도면32b



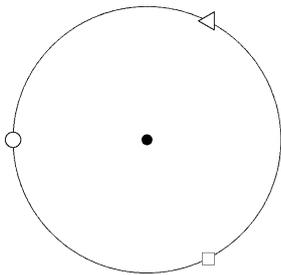
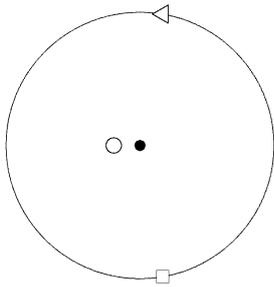
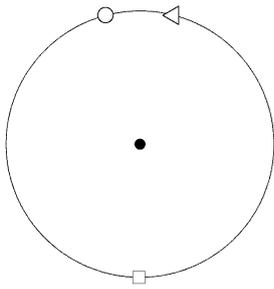
도면32c



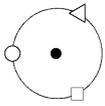
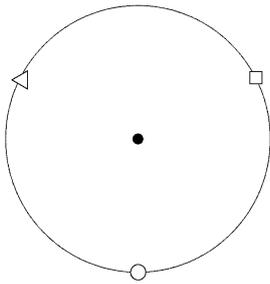
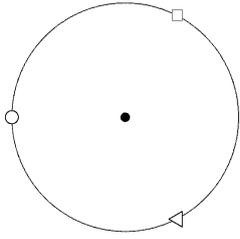
도면32d



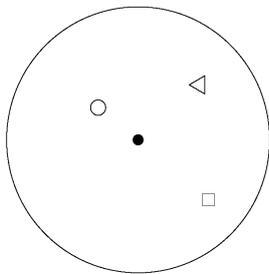
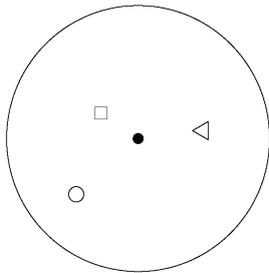
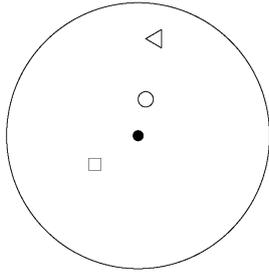
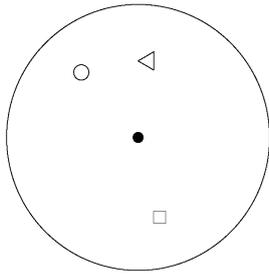
도면33a



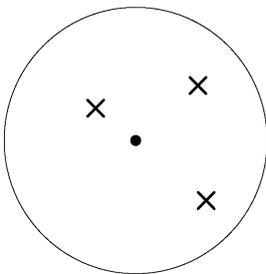
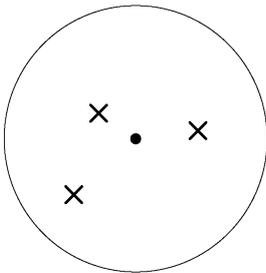
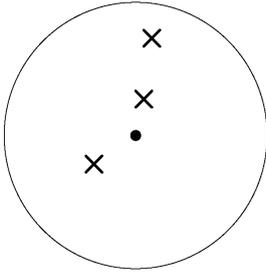
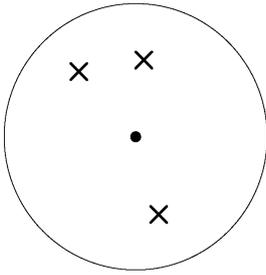
도면33b



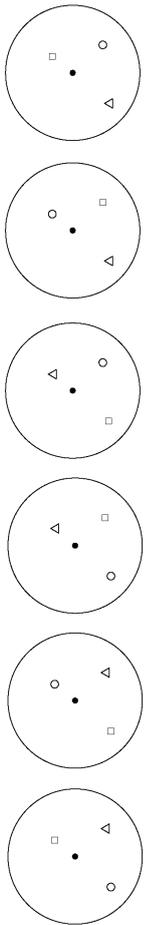
도면34a



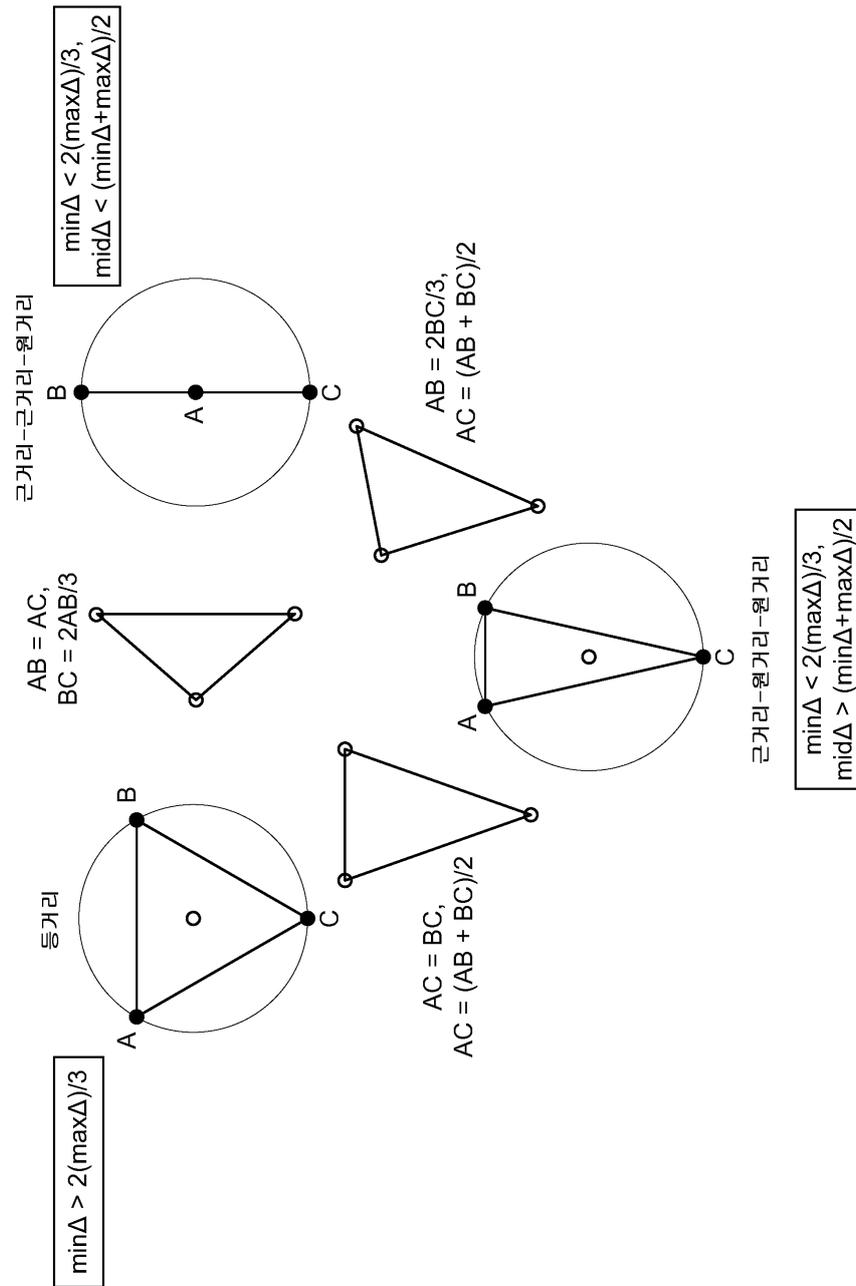
도면34b



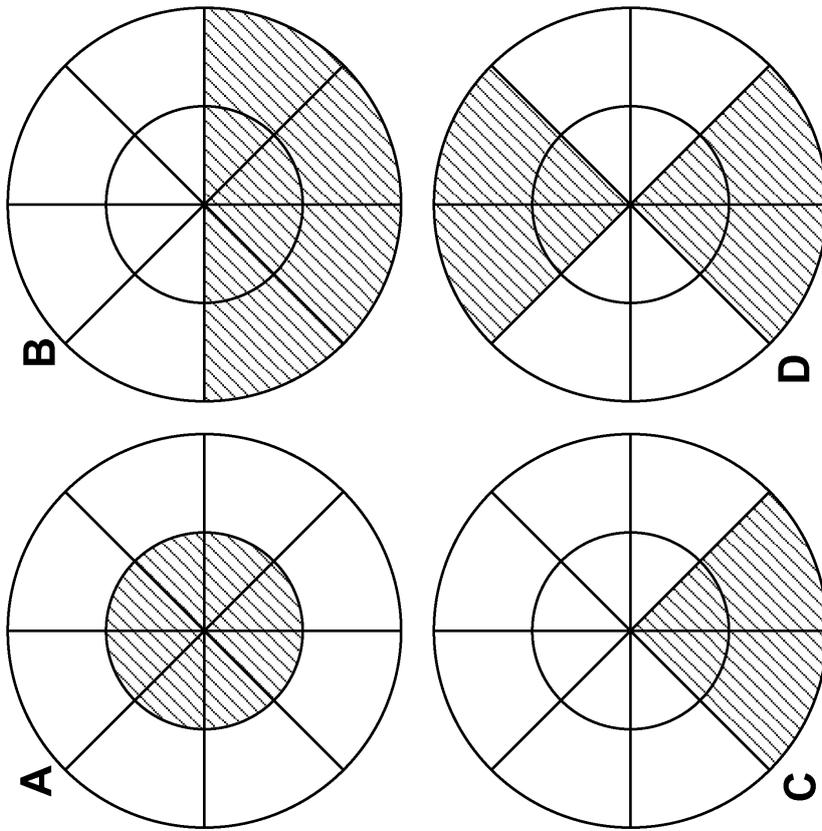
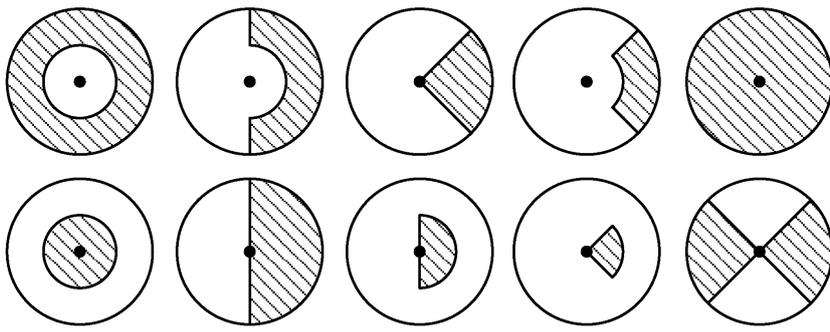
도면34c



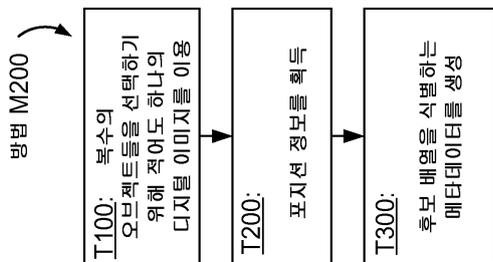
도면35



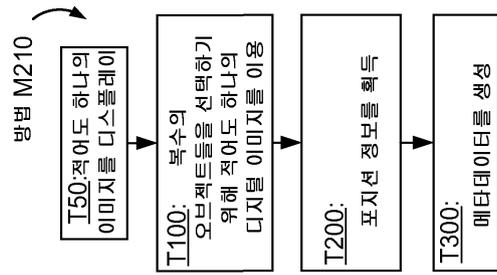
도면36



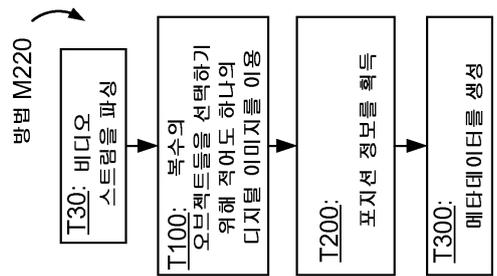
도면37a



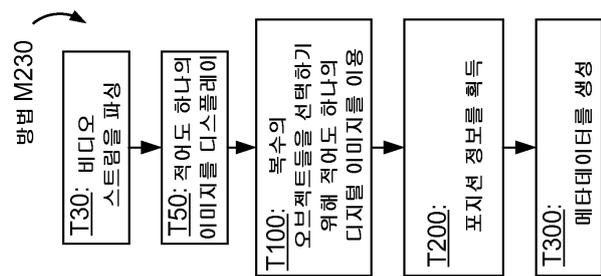
도면37b



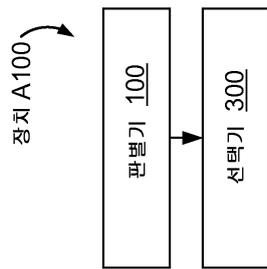
도면37c



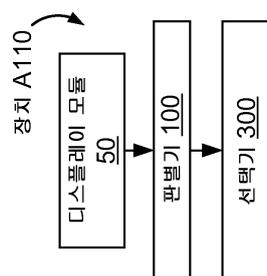
도면37d



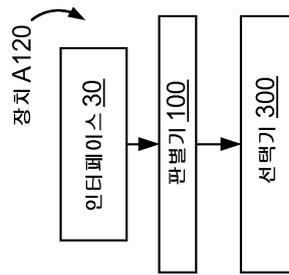
도면38a



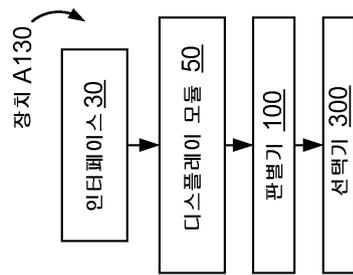
도면38b



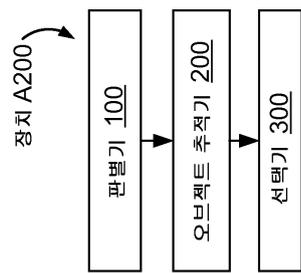
도면38c



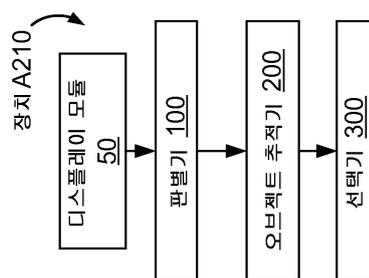
도면38d



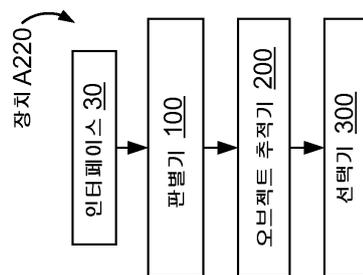
도면38e



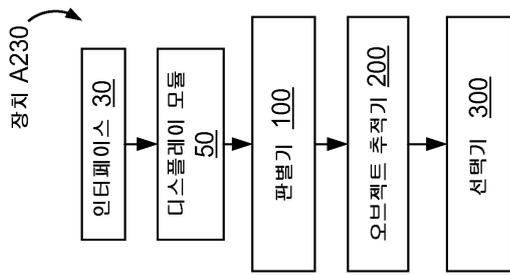
도면38f



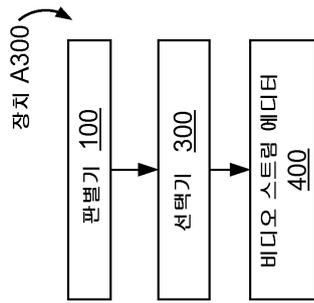
도면38g



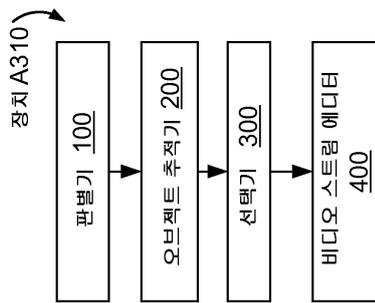
도면39a



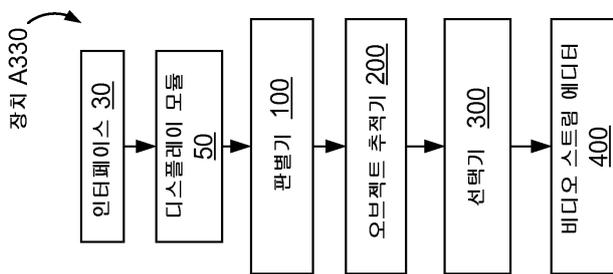
도면39b



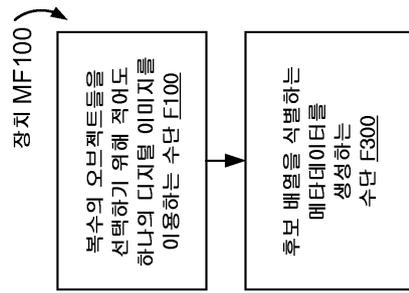
도면39c



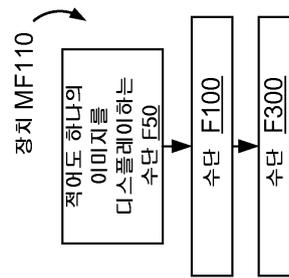
도면39d



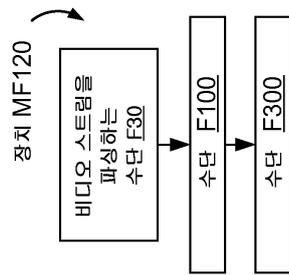
도면40a



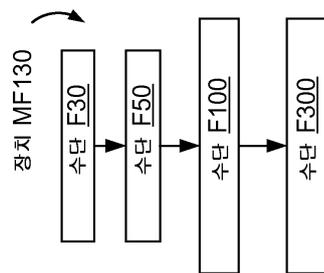
도면40b



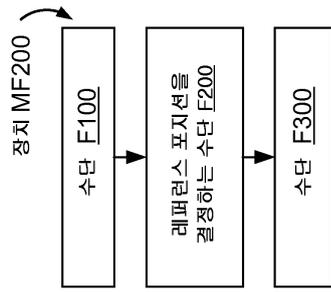
도면40c



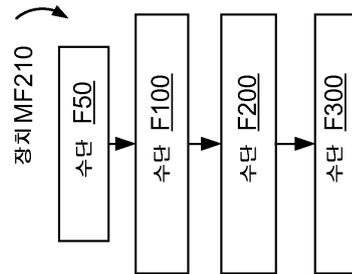
도면40d



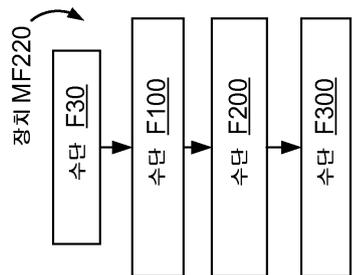
도면40e



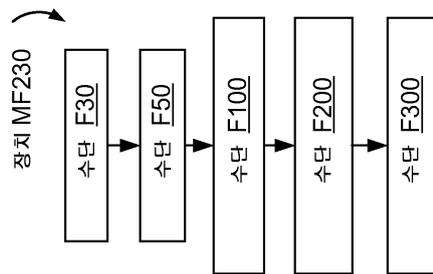
도면40f



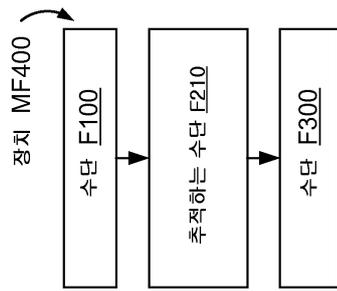
도면40g



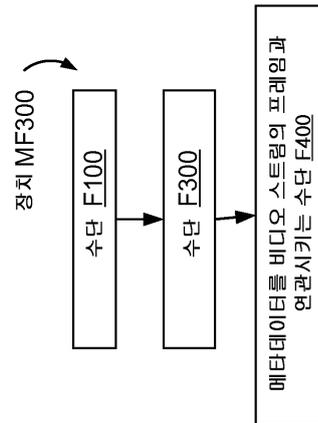
도면41a



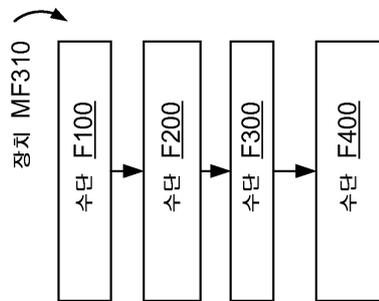
도면41b



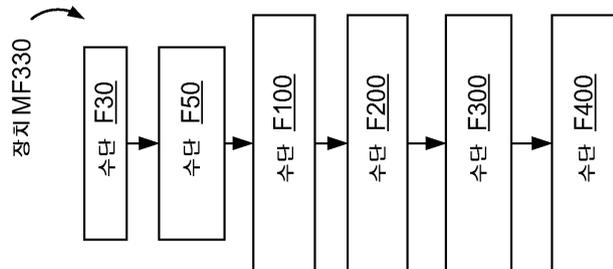
도면41c



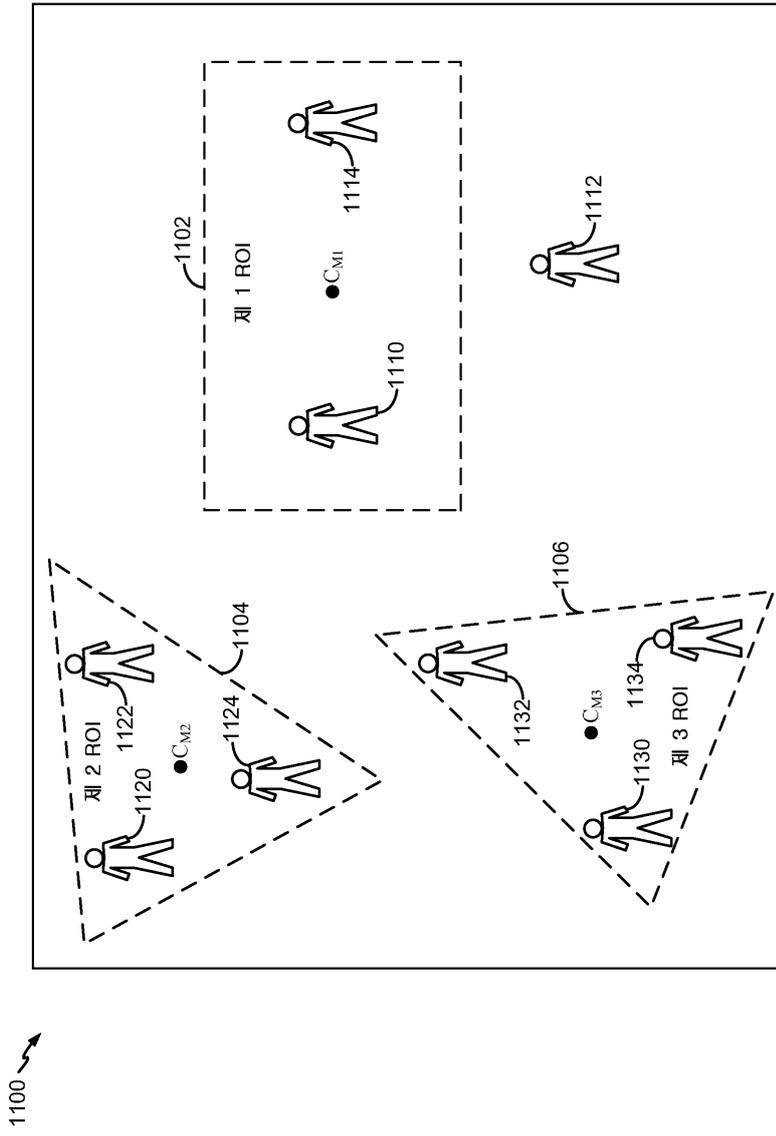
도면41d



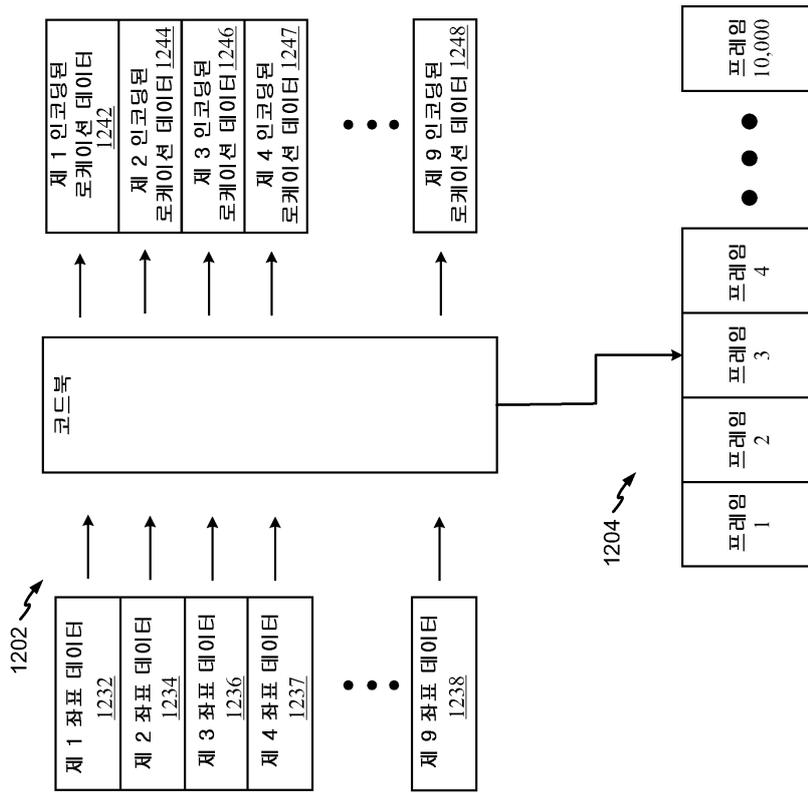
도면41e



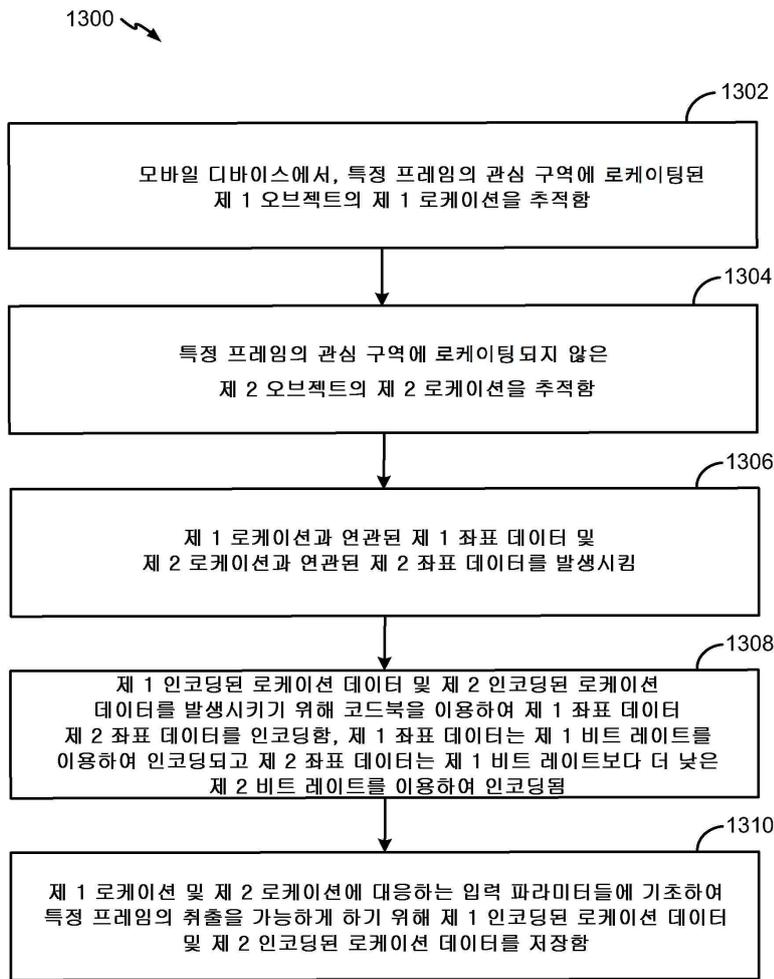
도면42



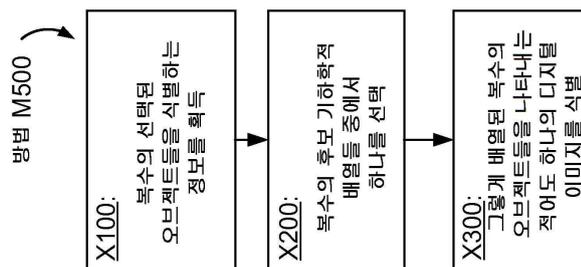
도면43



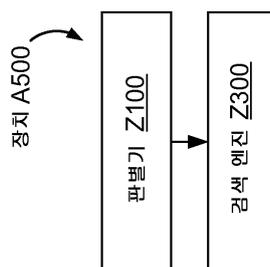
도면44



도면45a



도면45b



도면45c

