



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0003776
(43) 공개일자 2015년01월09일

- (51) 국제특허분류(Int. Cl.)
HO4N 19/51 (2014.01) *HO4N 19/50* (2014.01)
- (21) 출원번호 10-2014-7030447
- (22) 출원일자(국제) 2013년03월25일
심사청구일자 2014년10월29일
- (85) 번역문제출일자 2014년10월29일
- (86) 국제출원번호 PCT/EP2013/056189
- (87) 국제공개번호 WO 2013/144049
국제공개일자 2013년10월03일
- (30) 우선권주장
12305389.4 2012년03월30일
유럽특허청(EPO)(EP)
- (71) 출원인
알까멜 루슨트
프랑스 92100 불론뉴-비영꾸르 루뜨 둘 라 렌느
148/152
- (72) 발명자
론다오 알페이스 패트리스
벨기에 1400 비-니벨스 뒤 세인트 바르브 110 비
티이 1
마크 장-프랑소아
벨기에 비-1083 갠쇼렌 드레브 뒤 샤토 32
벨지프 니코
벨기에 비-2018 앤트워펜 렘브란드스트라아트 20
비4
- (74) 대리인
장훈

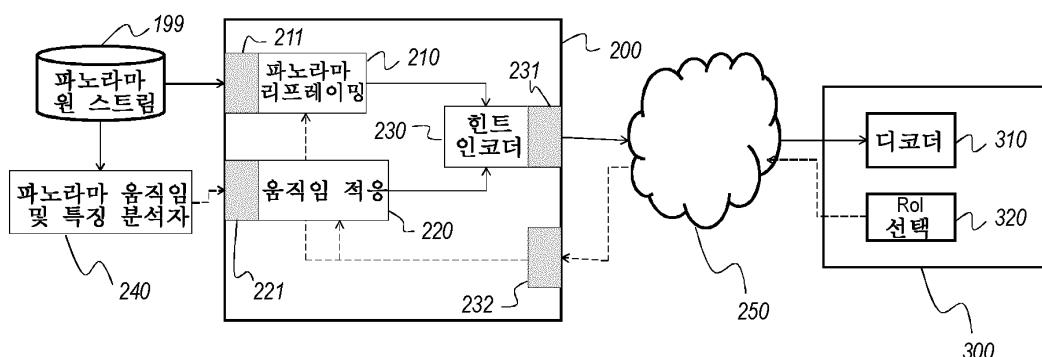
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **비디오 스트림의 선택된 공간 부분을 인코딩하는 방법 및 장치**

(57) 요약

원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하는 방법으로서, 상기 방법은 상기 선택된 공간 부분에 속하는 화상 요소 정보를 획득하는 단계; 상기 선택된 공간 부분의 주변에 있는 상기 원래의 비디오 스트림의 상호 보완적인 공간 부분으로부터 도출된 인코딩 힌트들을 획득하는 단계; 및 상기 인코딩 힌트들을 사용하여 상기 선택된 공간 부분을 인코딩하는 단계를 포함한다.

대표도 - 도2



특허청구의 범위

청구항 1

원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법에 있어서, 상기 선택된 공간 부분에 속하는 화상 요소 정보를 획득하는 단계; 상기 선택된 공간 부분 주변의 상기 원래의 비디오 스트림의 상호 보완적인 공간 부분으로부터 도출된 인코딩 힌트들을 획득하는 단계; 및 상기 인코딩 힌트들을 사용하여 상기 선택된 공간 부분을 인코딩하는 단계를 포함하는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 2

제 1 항에 있어서,
상기 원래의 비디오 스트림을 획득하는 단계;
상기 독립적인 비디오 스트림의 상기 선택된 공간 부분을 나타내는 선택 정보를 획득하는 단계; 및
상기 선택 정보에 따라 상기 원래의 비디오 스트림으로부터 상기 선택된 공간 부분에 속하는 상기 화상 요소 정보를 추출하는 단계를 추가로 포함하는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 3

제 2 항에 있어서,
상기 원래의 비디오 스트림에 속하는 움직임 및 특징 정보를 획득하는 단계;
상기 움직임 및 특징 정보에서 상기 선택된 공간 부분 주변의 영역에 속하는 관련 특징들을 식별하는 단계; 및
상기 식별된 관련 특징들로부터 상기 인코딩 힌트들을 도출하는 단계를 추가로 포함하는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,
상기 인코딩은 크기 조정 가능한 비디오 코덱에 의해 수행되는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 5

제 4 항에 있어서,
상기 원래의 비디오 스트림은 상기 인코딩을 위해 상기 기저층으로서 제공되는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 6

제 4 항 또는 제 5 항에 있어서,
상기 크기 조정 가능한 비디오 코덱은 H.264 SVC 코덱인, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 7

제 1 항 내지 제 6 항 중 어느 한 항에 있어서,
상기 선택된 공간 부분에 속하는 상기 화상 요소 정보의 상기 획득 단계는 인코딩된 비디오 스트림을 획득하는

단계를 포함하고, 상기 인코딩 단계는 상기 인코딩된 비디오 스트림을 트랜스코딩하는 단계를 포함하는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법.

청구항 8

실행될 때, 제 1 항 내지 제 7 항 중 어느 한 항에 따른 방법을 수행하도록 적응되는, 컴퓨터 프로그램.

청구항 9

데이터 처리 장치가 제 1 항 내지 제 7 항 중 어느 한 항에 따른 방법 단계들을 실행하게 하는 명령들을 포함하는, 컴퓨터 판독가능 데이터 저장 매체.

청구항 10

원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치에 있어서,

상기 원래의 비디오 스트림을 수신하기 위한 비디오 입력 인터페이스;

상기 선택된 공간 부분을 나타내는 선택 정보를 수신하기 위한 선택 정보 입력 인터페이스;

상기 원래의 비디오 스트림으로부터 도출된 특정 정보를 수신하기 위한 움직임 및 특정 정보 입력 인터페이스;

상기 비디오 입력 인터페이스 및 상기 선택 정보 입력 인터페이스에 동작가능하게 결합된 파노라마 리프레이머로서, 상기 선택 정보에 따라 상기 원래의 비디오 스트림으로부터 상기 선택된 공간 부분에 속하는 화상 요소 정보를 추출하도록 구성되는, 상기 파노라마 리프레이머;

상기 움직임 및 특정 정보 입력 인터페이스 및 상기 선택 정보 입력 인터페이스에 동작가능하게 결합된 움직임 어댑터로서, 상기 움직임 및 특정 정보에서 상기 선택된 공간 부분 주변의 영역에 속하는 관련 특징들을 식별하고, 상기 식별된 관련 특징들로부터 인코딩 힌트들을 도출하도록 구성되는, 상기 움직임 어댑터; 및

상기 파노라마 리프레이머 및 상기 움직임 어댑터에 동작가능하게 결합된 힌트 인코더로서, 상기 인코딩 힌트들을 사용하여 상기 선택된 공간 부분을 인코딩하도록 구성된, 상기 힌트 인코더를 포함하는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치.

청구항 11

제 10 항에 있어서,

상기 힌트 인코더는 크기 조정가능한 비디오 코덱에 의해 상기 선택된 공간 부분을 인코딩하도록 구성되는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치.

청구항 12

제 11 항에 있어서,

상기 힌트 인코더는 상기 인코딩을 위해 상기 원래의 비디오 스트림을 상기 기저층으로서 제공하도록 구성되는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치.

청구항 13

제 11 항 또는 제 12 항에 있어서,

상기 크기 조정가능한 비디오 코덱은 H.264 SVC 코덱인, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치.

청구항 14

제 10 항 내지 제 13 항 중 어느 한 항에 있어서,

상기 비디오 입력 인터페이스는 상기 비디오 스트림을 인코딩된 비디오 스트림으로서 획득하도록 구성되고, 상기 힌트 인코더는 상기 인코딩된 비디오 스트림을 트랜스코딩하도록 구성되는, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치.

청구항 15

원래의 비디오 스트림의 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 시스템에 있어서, 주변부의 특징들에 속하는 움직임 정보를 추출하도록 구성된 특정 분석자 및 제 10 항 내지 제 14 항 중 어느 한 항에 따른 다수의 장치를 포함하고, 상기 특정 분석자는 상기 장치의 각각의 움직임 및 특징 정보 인터페이스들에 결합되는, 원래의 비디오 스트림의 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 시스템.

명세서

기술 분야

[0001] 본 발명은 비디오 스트림 인코딩의 분야에 관한 것이다.

배경 기술

[0002] 텔레비전으로 방송되는 스포츠 경기들, 콘서트들, 극 제작물들 등과 같은 특정 형태들의 비디오 콘텐트에 대하여, 시청자는 특정한 시간들에 고품질 개요 화면을 시청하는 데 관심이 있을 수 있고, 다른 시간들에 동작의 특정 부분들에 줌 인하기를 원할 수 있다. 알려진 클라이언트측 팬/틸트/줌(PTZ) 기능들은 확대된 이미지 부분의 품질 또는 해상도의 손실을 야기한다. 알려진 서버측 팬/틸트/줌(PTZ) 기능들은 서버측에서 대량의 계산을 야기한다.

발명의 내용

해결하려는 과제

[0003] 본 발명의 실시예들의 목적은 결과 이미지가 매우 높은 품질을 보유하고, 동시에 계산 자원들이 더 효율적으로 사용되는 방식으로 클라이언트측 PTZ 기능을 제공하는 것이다.

과제의 해결 수단

[0004] 본 발명의 일 양태에 따라, 원래의 비디오 스트림 중 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 방법이 제공되고, 상기 방법은: 선택된 공간 부분에 속하는 화상 요소 정보를 획득하는 단계; 선택된 공간 부분 주변의 원래의 비디오 스트림의 상호 보완적인 공간 부분으로부터 도출된 인코딩 힌트들을 획득하는 단계; 및 인코딩 힌트들을 사용하여 선택된 공간 부분을 인코딩하는 단계를 포함한다.

[0005] 본 발명에 따른 방법의 이점은, 선택된 부분에 대한 인코딩 프로세스가 더 큰 화상에 관한 정보, 특히 선택된 부분 밖이지만, 다음의 프레임들의 코딩에 영향을 미치는 성질이 되기 충분히 가까운 영역에 나타나는 원래의 비디오 스트림의 특징들을 고려함으로써 더 효율적으로 수행될 수 있다는 것이다. 관련된 주변의 특징들은 이들 특징들의 움직임, 및 선택된 부분의 "움직임", 즉, 원래의 비디오 스트림에 관한 임의의 패닝, 틸팅, 또는 주밍에 의해 유도되는 이미지의 "움직임"에 기초하여 바람직하게 선택된다.

[0006] 본 발명에 따른 방법의 다른 이점은, 인코딩 힌트들의 도출이 발생하는 장소와 관계없이, 인코딩 스테이지에서 개선이 획득되는 것이다. 이러한 셋업은 힌트 도출 프로세스의 부분, 특히 원래의 비디오 퍼드에 관한 특징 및 움직임 추출이 나머지 방법 단계들로부터 분리되는 것을 허용한다. 이러한 방식으로, 이들 분리된 단계들의 집중화가 가능하게 행해지고, 다수의 인코딩들이 수행되는 경우, 효율적으로 증가되는 것을 초래한다.

[0007] 일 실시예에서, 본 발명에 따른 방법은 원래의 비디오 스트림을 획득하는 단계; 독립적인 비디오 스트림의 선택된 공간 부분을 나타내는 선택 정보를 획득하는 단계; 및 선택 정보에 따른 원래의 비디오 스트림으로부터 선택된 공간 부분에 속하는 화상 요소 정보를 추출하는 단계를 추가로 포함한다.

[0008] 원래의 비디오 스트림의 정보가 감소된 비디오를 생성하기 위해 사용되고, 원래의 비디오 스트림의 품질(예를 들면, 해상도)이 감소된 비디오 스트림에서 가장 크게 가능한 정도로 유지될 수 있는 것이 본 실시예의 이점이다.

[0009] 일 실시예에서, 본 발명에 따른 방법은 원래의 비디오 스트림에 속하는 움직임 및 특징 정보를 획득하는 단계; 움직임 및 특징 정보에서 선택된 공간 부분의 주변 영역에 속하는 관련 특징들을 식별하는 단계; 및 식별된 관

련 특징들로부터 인코딩 힌트들을 도출하는 단계를 추가로 포함한다.

[0010] 원래의 비디오 스트림의 정보가 움직임 및 특징 정보를 생성하기 위해 사용되고, 이러한 이용가능한 정보의 최대치가 인코딩 힌트들을 생성하기 위해 고려될 수 있다는 것이 본 실시예의 이점이다. 바람직하게는, 움직임 적응자는 후보 특징의 움직임 벡터를 선택된 공간 부분의 팬/틸트/줌 이동을 나타내는 움직임 벡터와 비교함으로써 관련 특징들을 식별하도록 구성된다. 따라서, 선택된 부분의 주변의 영역의 특징들은 그들의 출현이 예측될 수 있는 경우 관심 영역으로 이동할 때 검출될 수 있고, 새로운 값들은 상기 특징의 예측된 출현을 최적으로 처리하기 위해 가변적인 인코딩 파라미터들에 대해 선택될 수 있다.

[0011] 본 발명에 따른 방법의 일 실시예에서, 인코딩은 크기 조정 가능한 비디오 코딩에 의해 수행된다.

[0012] 일 특정 실시예에서, 원래의 비디오 스트림은 인코딩을 위한 기저층으로서 제공된다.

[0013] 일 특정 실시예에서, 크기 조정 가능한 비디오 코덱은 H.264 SVC 코덱이다.

[0014] 본 발명에 따른 방법의 일 실시예에서, 선택된 공간 부분에 속하는 화상 요소 정보의 획득 단계는 인코딩된 비디오 스트림을 획득하는 단계를 포함하고, 상기 인코딩 단계는 인코딩된 비디오 스트림을 트랜스코딩하는 단계를 포함한다.

[0015] 원래의 비디오 스트림 및/또는 선택된 공간 부분은 원래 인코딩된 스트림들로서만 이용가능할 수 있다. 상기 경우에서, 신중한 트랜스코딩은 원래의 비디오 스트림의 디코딩 및 재코딩보다 더 효율적이다.

[0016] 본 발명의 일 양태에 따라, 실행될 때, 본 발명의 실시예들에 따른 방법을 수행하도록 적용된 컴퓨터 프로그램이 제공된다.

[0017] 본 발명의 일 양태에 따라, 데이터 처리 장치가 본 발명의 실시예들에 따라 본 방법의 단계들을 실행하도록 하기 위한 명령들을 포함하는 컴퓨터 관독가능 저장 매체가 제공된다.

[0018] 본 발명의 일 양태에 따라, 원래의 비디오 스트림의 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치가 제공되고, 상기 장치는: 원래의 비디오 스트림을 수신하기 위한 비디오 입력 인터페이스; 선택된 공간 부분을 나타내는 선택 정보를 수신하기 위한 선택 정보 입력 인터페이스; 원래의 비디오 스트림으로부터 도출된 특징 정보를 수신하기 위한 움직임 및 특징 정보 입력 인터페이스; 비디오 입력 인터페이스 및 선택 정보 입력 인터페이스에 동작가능하게 결합된 파노라마 리프레이머로서, 선택 정보에 따라 원래의 비디오 스트림으로부터 선택된 공간 부분에 속하는 화상 요소 정보를 추출하도록 구성되는, 상기 파노라마 리프레이머; 움직임 및 특징 정보 입력 인터페이스 및 선택 정보 입력 인터페이스에 동작가능하게 결합된 움직임 어댑터로서, 움직임 및 특징 정보에서 선택된 공간 부분 주변의 영역에 속하는 관련 특징들을 식별하고, 식별된 관련 특징들로부터 인코딩 힌트들을 도출하도록 구성되는, 상기 움직임 어댑터; 및 파노라마 리프레이머 및 움직임 어댑터에 동작가능하게 결합된 힌트 인코더로서, 인코딩 힌트들을 사용하여 선택된 공간 부분을 인코딩하도록 구성되는, 상기 힌트 인코더를 포함한다.

[0019] 본 발명에 따른 장치의 일 실시예에서, 힌트 인코더는 크기 조정 가능한 비디오 코덱에 의해 선택된 공간 부분을 인코딩하도록 구성된다.

[0020] 일 특정 실시예에서, 힌트 인코더는 원래의 비디오 스트림을 인코딩을 위한 기저층으로서 제공하도록 구성된다.

[0021] 일 특정 실시예에서, 크기 조정 가능한 코덱은 H.264 SVC 코덱이다.

[0022] 본 발명에 따른 장치의 일 실시예에서, 비디오 입력 인터페이스는 비디오 스트림을 인코딩된 비디오 스트림으로서 획득하도록 구성되고, 힌트 인코더는 인코딩된 비디오 스트림을 트랜스코딩하도록 구성된다.

[0023] 본 발명은 원래의 비디오 스트림의 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 시스템을 또한 제공하고, 상기 시스템은, 주변 부분의 특징들에 속하는 움직임 정보를 추출하도록 구성된 특징 분석자, 및 청구항 제 9 항 내지 제 13 항 중 어느 한 항에 따른 다수의 장치를 포함하고, 특징 분석자는 장치의 각각의 움직임 및 특징 정보 인터페이스들에 결합된다.

발명의 효과

[0024] 본 발명의 실시예들에 따른 장치 및 시스템들의 효과들 및 이점들은 실질적으로 본 발명들의 실시예들에 따른 대응하는 방법들의 효과들 및 이점들을 준용하여 동일하다.

도면의 간단한 설명

[0025] 도 1은 본 발명의 일 실시예에 따른 방법의 플로차트.

도 2는 본 발명의 실시예들에 따른 장치 및 시스템의 개략도.

도 3은 본 발명의 실시예들에서 수행되는 움직임 벡터 비교를 도시하는 도면.

발명을 실시하기 위한 구체적인 내용

[0026] 본 발명의 실시예들에 따른 장치 및/또는 방법들의 몇몇 실시예들은 단지 예로서 및 첨부하는 도면들을 참조하여 여기에 기술된다.

[0027] 본 발명은 고해상도 비디오 파노라마로부터 팬-틸트-줌(PTZ) 상호 작용 방식으로 개인화된 관심 영역(RoI)을 요청할 수 있는 다수의 사용자들을 서빙하는 프록시로 구성된 시스템에 관한 것이다. 본 발명은 또한 동일한 비디오 파노라마 소스로부터 이를 개인화된 비디오 스트리밍들의 비디오 인코딩의 프록시 측에서 계산 비용의 최적화에 관한 것이다. 용어 "파노라마"는 일반적으로, 큰 고해상도 비디오 해상도, 통상 4000 개 이상의 픽셀들 폭을 지정하도록 여기에 사용되고, 이는 광각 뷰의 원통형 또는 구형 매핑 또는 물리적 공간의 비디오로서의 표현을 포함할 수 있다. 선택적으로, 파노라마는 함께 통합되고 응화되는 상이한 비디오 소스들로부터 구성될 수 있다.

[0028] 현재 해결책들은 원래의 데이터에서 시간 t 에서 원하는 RoI 프레임을 생성하고, H.264, 또는 움직임 보상 및/또는 인트라 코딩을 갖는 WebM 유사 코덱을 사용하여 이를 인코딩하기 위해 사용자 요청된 크롭핑 및 PTZ 동작들을 수행한다.

[0029] 불행히도, 이러한 기술들을 사용하여, 움직임 보상/추정 및 인트라 예측 모드들은 일부가 공통이거나 또는 중첩하는 RoI를 요청할지라도 모든 사용자들에 대해 계산되어야 한다. 이는 프록시가 입력 비디오 파노라마가 모든 사용자들에 대하여 동일하다는 사실로부터 이익을 얻을 수 없기 때문에 나쁜 시스템 크기 조정 가능성을 초래한다. 이동 디바이스들에 대한 비디오 적응을 위한 하드웨어 가속을 트랜스코딩에 제공하는 시스템들에 대해서도 동일하게 언급될 수 있다.

[0030] 본 기술에서, 비디오 코딩 용어는 H.264 권고 사항의 환경에 공통적인 것이 사용될 것이다. 숙련된 독자는, 이 것이 단지 텍스트를 명확하게 하기 위해 수행되는 것이고, 용어의 이러한 선택은 본 발명의 범위를 특정 인코딩 표준이 적용되는 시스템들에 대해 제한하여 의도하는 것은 아님을 이해할 것이다.

[0031] 본 발명의 실시예들에서, RoI 움직임 추정 및 인트라 예측의 총 계산 복잡성은 파노라마로부터 직접 관련 움직임 및 인트라 예측 방향들을 계산함으로써 감소된다. 결과적으로 계산 속도가 증진된다. 새로운 클라이언트들을 추가하는 것이 작은 계산 복잡성 증가를 초래하기 때문에, 결과의 시스템의 크기 조정 가능성을 주요 이점이다.

[0032] 본 발명의 실시예들은 이러한 이점이 다음의 두 개의 단계들을 실행함으로써 달성될 수 있다는 본 발명자들의 통찰력에 기초한다:

[0033] 1) 움직임 분석(또는 내부 예측 직접 추정)은 각각의 파노라마 화소에 대한 전해상도에서 전체 파노라마상에 수행된다.

[0034] 2) RoI 관련 움직임은 이후, 가능한 RoI 사용자 요청된 범위 및 해상도 변화들을 보상함으로써 이를 파노라마 움직임 벡터들로부터 계산될 수 있다. 실제로 이를 보상된 움직임 벡터들은 상기 RoI의 코딩에 전념된 인코더에 힌트들로서 전송된다. 인코더는 필요한 경우 이러한 움직임 벡터 "힌트"를 개선할 수 있다.

[0035] 도 1은 본 발명의 일 실시예에 따른 방법의 플로차트를 제공한다. 숙련자는 다수의 예시적인 단계들이 반드시 단일 엔티티에 의해 수행되는 것이 아님을 이해할 것이다. 또한, 동시에 발생하는 것으로 도시된 단계들은, 특정 순서가 확실히 요구되는 것이 상세한 설명으로부터 명백하지 않으면, 연속하여 행해질 수 있고, 그 반대도 마찬가지이다.

[0036] 도시된 실시예에 따라, 원래의 비디오 스트리밍은 한편으로 선택된 영역(130)에 대한 관련 화상 요소 정보를 추출하기 위해서, 및 다른 한편으로 움직임 및 특징 정보(150)를 추출하기 위해서 획득되어(100) 사용된다.

[0037] 선택된 영역에 대한 화상 요소 정보의 추출을 수행하기 위하여, 관심 영역(RoI)은, 일반적으로 비디오 클라이언트 디바이스를 통해 시청자에 의해, 그렇지만 대안적으로 또는 추가로 인간 디렉터 또는 자동화된 스크립트에 의해 선택되어야 하고(110), 이러한 셙션 정보는 추출측에서 획득되어야 한다(120). 선택 동작은 원래의 비디오

스트림에 관하여 패닝, 주밍, 및 틸팅으로 한정되지 않고, 선택된 영역의 화상비는 바람직하게는 시청 디바이스와 연관된 고정 화상비(예를 들면, 3×4 또는 16×9)로 한정된다. 화상 요소 정보는 선택된 영역내 임의의 형태의 비디오 이미지들의 표현을 포함하는 것을 의미한다. 이는 압축되지 않은 비디오 이미지들의 스트림, 또는 인코딩된 비디오 스트림일 수 있다.

[0038] 본 발명에 따른 방법의 실시예들은 움직임이 가능하게는 다중 해상도 접근법에서 가장 높은 해상도로 계산되는 이점을 나타내고, 여기서 모든 데이터는 이용가능하고(즉, 파노라마) 따라서, 가장 높은 정확도를 달성한다. 본 발명에 따른 방법의 실시예들은 상이한 사용자들에 대하여 중첩하는 ROI들 또는 정렬된 ROI들이, 움직임 데이터가 파노라마상에서 미리 계산되기 때문에, 그들의 각각의 인코더들에 대한 움직임 추정 노력의 중복을 요구하지 않는다는 다른 이점을 제공한다.

[0039] 인코딩 힌트들은 추출된 움직임 및 특징 정보로부터 도출된다(160). 움직임 및 특징 정보는 모든 이용가능한 정보를 고려하는 원래의 비디오 스트림으로부터 추출되지만, 인코딩 힌트 도출은 이하에 더 상세히 기술되는 인코딩 성능을 개선하기 위한 것과 관련된 이들 특징들을 선택한다.

[0040] 인코더는, 선택된 부분에 속하는 화상 요소 정보(140) 및 일반적으로 주변부로부터 도출된 인코딩 힌트들(170)을 획득한다. 이들 입력들에 기초하여, 인코더는 비디오 스트림의 선택된 부분을 인코딩한다(180).

[0041] 인코딩 힌트들은 후자의 인코딩을 개선하기 위해 선택된 부분에 관한 정보에 추가로 사용된 원래의 비디오 스트림의 선택되지 않은 부분으로부터 도출된 정보의 비트들이다. 이를 달성하기 위하여, 본 발명의 실시예들에 따른 방법들은 클라이언트로부터 내비게이션 및 줌 명령들을 추적하고(단계 120에서), 이후 줌 명령들에 응답하여 전체 움직임 데이터를 크기 조정하고 패닝 또는 틸팅 명령들에 응답하여 ROI 움직임을 추가한다(단계 160에서). 인코딩(단계 180에서)은 힌트들로서 또는 실제 움직임 벡터들로서 사용되는 획득된 움직임 벡터들의 세트에 기초하여 발생한다: 기준 프레임 영역이 인코더에서 이용가능하지 않은 경우, 예측된 움직임 벡터 및 검색 윈도 범위는 인코더 또는 인트라 예측 모드로 전송된다.

[0042] 전술된 방법은 일반적으로 원래의 비디오 스트림의 선택된 공간 부분을 독립적인 비디오 스트림으로서 인코딩하기 위한 장치에 의해 수행될 수 있고, 상기 장치는, 선택된 공간 부분에 속하는 화상 요소 정보를 획득하기 위한 수단(140); 선택된 공간 부분 주변의 원래의 비디오 스트림의 상호 보완적 공간 부분으로부터 도출된 인코딩 힌트들을 획득하기 위한 수단(170); 및 인코딩 힌트들을 사용하여 선택된 공간 부분을 인코딩하기 위한 수단(180)을 포함한다.

[0043] 장치는 원래의 비디오 스트림을 획득하기 위한 수단(100); 독립적인 비디오 스트림의 선택된 공간 부분을 나타내는 선택 정보를 획득하기 위한 수단(120); 및 선택 정보에 따라 원래의 비디오 스트림으로부터 선택된 공간 부분에 속하는 화상 요소 정보를 추출하기 위한 수단(130)을 추가로 포함할 수 있다.

[0044] 특히, 장치는 원래의 비디오 스트림에 속하는 움직임 및 특징 정보를 획득하기 위한 수단(150); 움직임 및 특징 정보에서 선택된 공간 부분 주변의 영역에 속하는 관련 특징들을 식별하기 위한 수단; 및 식별된 관련 특징들로부터 인코딩 힌트들을 도출하기 위한 수단(160)을 추가로 포함할 수 있다.

[0045] 인코딩을 위한 수단(180)은 크기 조정 가능한 비디오 코덱과 함께 동작할 수 있다. 특히, 인코딩을 위한 수단(180)은 원래의 비디오 스트림을 인코딩을 위한 기저층으로서 제공할 수 있다. 또한, 더 구체적으로, 크기 조정 가능한 비디오 코덱은 H.264 SVC 코덱일 수 있다.

[0046] 선택된 공간 부분에 속하는 화상 요소 정보를 획득하기 위한 수단(140)은 인코딩된 비디오 스트림을 획득하도록 적응될 수 있고, 인코딩 수단(180)은 인코딩된 비디오 스트림을 트랜스코딩하도록 적응될 수 있다.

[0047] 도 2는 본 발명의 실시예들에 따른 장치 및 시스템의 개략적인 도시를 제공한다. 장치(200)는 상기 원래의 비디오 스트림(199)을 수신하기 위한 비디오 입력 인터페이스(211), 상기 선택된 공간 부분을 나타내는 선택 정보를 수신하기 위한 선택 정보 입력 인터페이스(232), 상기 원래의 비디오 스트림으로부터 도출된 특징 정보를 수신하기 위한 움직임 및 특징 정보 입력 인터페이스(221), 상기 비디오 입력 인터페이스(211) 및 상기 선택 정보 입력 인터페이스(232)에 선택적으로 결합된, 파노라마 리프레이머(211)로서, 상기 선택 정보에 따라 상기 원래의 비디오 스트림으로부터 상기 선택된 공간 부분에 속하는 화상 요소 정보를 추출하도록 구성되는, 상기 파노라마 리프레이머(211), 상기 움직임 및 특징 정보 입력 인터페이스(221) 및 상기 선택 정보 입력 인터페이스(232)에 선택적으로 결합된 움직임 어댑터(220)로서, 상기 움직임 및 특징 정보의 상기 선택된 공간 부분 주변의 영역에 속하는 관련 특징들을 식별하고 상기 식별된 관련 특징들로부터 인코딩 힌트들을 도출하도록 구성되는, 상기 움직임 어댑터(220), 및 상기 파노라마 리프레이머(210) 및 상기 움직임 어댑터(220)에 동작가능하게

결합된 힌트 인코더(230)로서, 상기 인코딩 힌트들을 사용하여 상기 선택된 공간 부분을 인코딩하도록 구성되는, 상기 힌트 인코더(230)를 포함한다.

[0048] 장치(200)는 네트워크(250)를 통해 인코딩되어 선택된 비디오 스트림을 클라이언트(300)로 전송하기 위한 출력 인터페이스(231)를 추가로 갖는다. 도시된 네트워크(250)는 하나 이상의 네트워크 링크들로 구성될 수 있다. 이는 일반적으로 클라이언트측상의 액세스 링크를 포함한다.

[0049] 용어 "인터페이스"는 본 기술의 숙련자에게 잘 알려진 다수의 계층들의 프로토콜 스택에 걸쳐 데이터 통신 접속을 확립하기 위해 요구되는 필수적인 하드웨어 및 소프트웨어를 나타낸다. 바람직하게는, 표준화된 프로토콜들이 사용된다. 액세스 인터페이스는, 예를 들면, xDSL, xPON, WMAN, 또는 3G 링크와 같은 액세스 링크용 인터페이스를 포함할 수 있다. LAN 인터페이스는, 예를 들면, IEEE 802.3 "이더넷" 링크, IEEE 802.11 "무선 LAN" 링크 중 하나 이상에 대한 인터페이스를 포함할 수 있다. PAN 인터페이스는, 예를 들면, USB 인터페이스 또는 블루투스 인터페이스를 포함할 수 있다. 수 개의 네트워크 세그먼트들에 걸친 통신을 위한 상위 계층 프로토콜들은 바람직하게 TCP/IP 프로토콜 모음으로부터의 프로토콜들일 수 있다.

[0050] 클라이언트(300)는 수신된 비디오 스트림을 디코딩하기 위한 표준 디코더(310)를 포함한다. 클라이언트(300)는 최종 사용자가 관심 영역을 선택하고, 결과의 선택 정보를 장치(200)로 피드백하게 하기 위한 수단(320)을 추가로 포함한다.

[0051] 기술된 장치에 더하여, 본 발명에 따른 시스템은 원래의 비디오 스트림(199)상에 동작하는 움직임 및 특징 분석자(240)를 포함할 수 있다. 바람직하게는, 이러한 움직임 및 특징 분석자(240)는 중앙 집중되거나, 또는 "클라우드에" 위치되고, 본 발명에 따른 하나 또는 여러 장치(200)에 그의 출력을 제공한다. 그러나, 동일한 디바이스 또는 제품에 움직임 및 특징 분석자(240) 및 단일 장치(200)의 기능들을 포함하는 것이 또한 가능하다.

[0052] 명확성 및 간결함 때문에, 특정한 특징들 및 이점들이 방법 실시예들 또는 장치/시스템 실시예들에 관련하여 단지 명시적으로 기술되었다. 숙련자는 특징들 및 이점들이 다른 카테고리의 대응하는 실시예들에 동등하게 적용한다는 것을 이해할 것이다.

[0053] 일 예시적인 배치에서, 각각이 클라이언트(300)를 서빙하는, N 개의 적응식 ROI 인코더들(200)에 연결된 파노라마 비디오 스트림 분석자(240)로 구성된 시스템을 고려한다. 각각의 적응식 ROI 인코더(200)는 도 2에 도시된 움직임 적응 모듈(220), 파노라마 비디오 리프레이밍 모듈(210), 및 클라이언트(300)를 전달하는 힌트 인코더(230)를 포함한다.

[0054] 파노라마 움직임 분석자(240)는 원래의 또는 인코딩된 파노라마 비디오 스트림(199)을 수신하고 움직임 정보를 N 개의 적응식 ROI 인코더들(200)의 움직임 적응 모듈(220)로 전달한다. 파노라마 분석기(240)가 압축된 도메인에서 그의 테스크를 수행하는 경우, 리프레이밍 모듈(210)은 필요한 부분 디코딩을 허용한다. 인코더(230)가 MPEG-7 준수 트랜스코더인 경우, 힌팅 정보는 MPEG-7 트랜스코딩 힌트들로서 인코더(230)로 제공된다. 이러한 트랜스코더는, Peter M.Kuhn, Teruhiko Suzuki 및 Anthony Vetro에 의한 "MPEG-7 Transcoding Hints for Reduced Complexity and Improved Quality"(패킷비디오 회의록, 2001년 1월)로부터 알려진다.

[0055] 각각의 움직임 적응 모듈(220)은 파노라마 움직임 분석자(240)로부터의 움직임 및 특징 정보 및 그의 클라이언트(300)로부터의 ROI 위치 변경 요청들을 수신한다. 이러한 정보에 기초하여, 모듈(220)은 그의 접속된 힌트 인코더(230)에 힌트들을 출력한다. 파노라마 리프레이밍 모듈(210)은 또한 ROI 요청을 판독하고 파노라마의 요청된 영역을 재샘플링하여 크롭하고 이를 요청된 위치에서 요청된 해상도로 힌트 인코더(230)에 대하여 준비한다. 파노라마 영역이 2D(예를 들면, 구면 비디오를 위한 원통형 맵)로 매핑되는 것이 요구되는 경우, 이는 또한 파노라마 리프레이밍 모듈(210)에서 행해진다.

[0056] 각각의 힌트 인코더(230)는 요청된 해상도 및 위치의 원래의 크롭된 비디오 스트림뿐만 아니라 그의 움직임 적응 모듈(220)로부터 힌트들을 수신한다. 힌트들은 압축 성능에서 불이익 없이 인코더(230)의 계산 복잡성을 감소시키는 것을 돋는 정보로 구성된다. H.264의 경우에, 힌트들은 직접 움직임 벡터, 모드 결정(움직임 분할), 검색 윈도가 집중되는 예측된 움직임 벡터뿐만 아니라 검색 윈도의 크기, 인트라 모드 예측, 스kip 결정 등으로 구성될 수 있다. 그의 예는, Ralph A. Braspenning, Gerard de Haan에 의한 "True-motion estimation using feature correspondences"(화상 통신 및 이미지 처리 2004, Vol. 5308, No.1. (2004), pp. 396-407)에서 찾을 수 있다.

[0057] 파노라마 분석자(240) 및 움직임 적응 모듈(220)에 의해 계산된 이들 힌트들을 사용하는 주요 이점은 움직임 예측의 테스크가 높은 정확성으로 전체 파노라마에 대하여 딱 한번 수행되고 모든 인코더들(200)에 의해 재생성되

지 않는다는 것이다.

[0058] 이러한 테스크의 집중화의 결과로서, 시스템의 전체 계산 복잡도는 감소한다. 이는 다음의 간략화된 계산에 의해 예시적인 방식으로 도시될 수 있다. C 가 움직임 추정의 복잡성(C_m) 및 나머지 압축 동작들의 복잡성(C_r)을 갖는 비디오 인코더의 계산 복잡성인 경우, $C = C_r + C_m$ 이고, 여기서 $C_m > C_r$ 이다(일반적으로 $C_m \gg C_r$). 파노라마 분석의 계산 복잡성(도 1의 단계(150); 도 2의 분석자(240)의 함수)은 C_p 로서 표기하고, $C_p > C$ 이다. 이러한 시스템에 따라, N 개의 클라이언트들을 서빙하는 것은 $(C_p + C_r)$ 의 함수의 고전적 선형성 대신에 단지 $C_p + N \times C_r$ 의 복잡성을 초래한다. 하나의 클라이언트를 추가하는 것은 단지 종래의 경우의 항 $C_m + C_r$ 보다 훨씬 작은 C_r 의 증분을 추가한다.

[0059] 파노라마 움직임 및 특징 분석자(240)에 의해 계산된 움직임 정보는 움직임 적응 모듈(220)에 의한 클라이언트 요청에 적응될 필요가 있다. 이는 이용가능한 움직임 벡터들이 힌트 인코더(230)에서 기준 프레임들로서 이용가능한 파노라마 영역들을 지시하고 있는지의 여부를 검출하는 것을 요구한다. 이러한 검출은 도 3에 나타내어지고, 파노라마 원 비디오 프레임들 움직임 및 특징 정보는 M 개의 프레임들에 대한 ROI 요청의 일 예와 함께 도시된다. 시간 t 에서 각각의 파노라마 프레임 $\Pi(t)$ 은 파노라마 움직임 및 특징 분석 모듈(240)에서 각각의 픽셀 특징 정보 및 움직임 정보에 대해 생성하기 위하여 분석된다. 특징 정보는, 예를 들면, 애지 위치 및 배향 정보, SIFT 디스크립터들과 함께 파노라마의 세분화로 구성되고, 움직임 정보는 M 개의 이전 프레임들($\Pi(t-1)$... $\Pi(t-M)$)의 각각에 대한 하나의 움직임 벡터로서 나타내어 기준 프레임들의 역할을 한다.

[0060] 움직임 적응 모듈(220)은 ROI 요청을 수신하고 힌트 인코더(230)에 의해 사용된 이전의 M 개의 기준 프레임들에 대하여 파노라마에서 ROI 위치 변경들을 계산한다. ROI 움직임 변경들은 도 3에서 r 로서 표시된 벡터들로 나타내고, 현재 ROI 매크로 블록들은 정규 격자로 나타내지고 이전 ROI 프레임 위치들은 대시로 그려진 직사각형 영역들(폭 $w(t)$ 및 높이 $h(t)$ 를 갖는)로서 나타내진다.

[0061] 움직임 적응 모듈은 이후 힌트 인코더(230)에 의해 인코딩될 모든 매크로 블록에 대하여, 파노라마 분석자(240)에 의해 파노라마에서 각각의 픽셀에 대해 계산된 각각의 움직임 벡터들이 힌트 인코더의 기준 프레임들에서 이용가능한지의 여부를 검사한다. 움직임 벡터들이 이용가능한 경우, 최상의 모드 결정을 계산하고 파노라마에서 ROI 윈도 위치 변경을 보상하여 이러한 정보를 힌트로서 힌트 인코더(230)에 전송하기 위해 움직임 벡터들을 적응 및 집속시킨다. 어떠한 움직임 벡터도 매크로블록에 대하여 이용가능하지 않은 경우, 분석자(240)로부터의 특징들은 현재 매크로 블록의 특징들이 파노라마 움직임 및 특징 분석 또는 이들 특징들에 의해 또한 안내된 인트라 예측 모드에 의해 행해진 분할(예를 들면, 4×4 또는 16×16 모드들에서, 매크로블록에서 검출된 애지들을 따른 인트라 예측 방향)에 의해 제공된 주어진 근접 영역의 특징들과 매칭하지를 검색 윈도에 알려주기 위하여 사용된다.

[0062] 클라이언트(300)로부터의 ROI 요청은 파노라마에서 ROI 해상도 및 위치 기술이다. 이러한 정보는 움직임 적응 모듈(220)에 의해 및 파노라마 리프레이밍 모듈(210)에 의해 사용된다. 이러한 모듈(210)은 파노라마의 정확한 영역을 선택하고, 이를 서브샘플링하거나 또는 삽입하거나 또는 요청된 해상도로 재매핑한다. 이를 동작들의 출력은 이후 크롭핑되어 힌트 인코더(230)로 전송된다.

[0063] 선택적으로, 수 개의 클라이언트들(300)이 상이한 해상도 능력들을 갖는 동일한 ROI를 요청하는 경우, 이들 클라이언트들(300)은 가상의 슈퍼 클라이언트들로서 합계될 수 있다. 이는 예를 들면, ROI 변위가 스크립트에 의해 가이드되는 경우 가능하다(예를 들면, 객체 또는 사람의 추적 또는 가상 또는 인간 생산 관리자에 의해 결정된 파노라마 공간에서 변위들). 힌트 인코더는 H.264, SVC의 크기 조정 가능한 확장과 같은 크기 조정 가능한 인코딩 기술을 사용한다. 이는 클라이언트(300)의 요구들에 대하여 스트림의 적응을 가능하게 하면서 여전히 인코딩의 계산 복잡성을 감소시킨다. 이후, 힌트들은 바람직하게 움직임 및 특징 분석에 기초하여 계층내 및 계층간 예측 힌트들을 또한 통합한다. 이후, 파노라마 움직임 분석자(240)는 특히 이를 가능하게 하도록 다중 해상도 움직임 및 특징 정보를 출력한다.

[0064] 선택적으로, 본 발명에 따른 방법은 힌트들 없이 각각의 ROI를 개별적으로 인코딩하기 위한 것보다 전체 파노라마상의 움직임 분석을 사용하는 것이 더 효율적인지를 추정하기 위한 추가의 단계를 포함한다. 상기에 규정된 개념들을 재사용하면, 계산은 파노라마 움직임 분석(단계 150) 비용이 더 높은 크기 조정 가능성에 의해 보상되는 클라이언트들의 최소수 N_p 를 결정하기 위해 수행된다. 상기 제공된 분석에 따라, N_p 는 다음의 수식에 의해 주어진다: $N_p \times (C_m + C_r) = C_p + N_p * C_r$. 이는 $N_p = C_p/C_m$ 을 따른다. 따라서, 일 바람직한 실시예에서, 본 발명에

따른 방법은 서빙될 클라이언트들의 수가 적어도 $N_p = C_p/C_m$ 와 같다는 것을 검증한 후에 상기에 기술된 힌트 인코딩을 사용한다.

[0065] 방법들 및 장치들이 개별적인 실시예들로서 상기에 기술되었지만, 이는 단지 명확성의 목적들을 위해 행해지고, 방법 실시예들과 관련되어서만 기술된 특징들은 동일한 기술적 효과들 및 이점들을 획득하기 위해 본 발명에 따른 장치에서 적용될 수 있고 그 반대도 마찬가지라는 것이 주의되어야 한다.

[0066] "프로세서들"이라고 라벨링된 임의의 기능 블록들을 포함하는 도면들에 도시된 다수의 요소들의 기능들은 적절한 소프트웨어와 연관되어 소프트웨어를 실행할 수 있는 하드웨어뿐만 아니라 전용 하드웨어의 사용을 통해 제공될 수 있다. 프로세서에 의해 제공될 때, 기능들은 단일의 전용 프로세서에 의해, 단일의 공유 프로세서에 의해, 또는 그의 일부가 공유될 수 있는 복수의 개별적인 프로세서들에 의해 제공될 수 있다. 더욱이, 용어 "프로세서" 또는 "제어기"의 명시적인 사용은 소프트웨어를 실행할 수 있는 하드웨어를 베타적으로 말하는 것으로 해석되지 않아야 하고, 디지털 신호 프로세서(DSP) 하드웨어, 네트워크 프로세서, 주문형 집적 회로(ASIC), 필드 프로그램 가능 게이트 어레이(FPGA), 소프트웨어를 저장하기 위한 판독 전용 메모리(ROM), 랜덤 액세스 메모리(RAM), 및 비휘발성 저장 장치를 제한 없이 암시적으로 포함할 수 있다. 관습형 및/또는 주문형의 다른 하드웨어가 또한 포함될 수 있다. 유사하게, 도면들에 도시된 임의의 스위치들은 단지 개념적이다. 그들의 기능은 프로그램 로직의 동작을 통해, 전용 로직을 통해, 프로그램 제어 및 전용 로직의 상호 작용을 통해, 또는 수동으로도 실행될 수 있고, 특정 기술은 문맥으로부터 더 구체적으로 이해될 수 있는 바와 같이 실행자에 의해 선택 가능하다.

[0067] 본 기술의 숙련자는 다수의 상술된 방법들의 단계들이 프로그래밍된 컴퓨터들에 의해 수행될 수 있다는 것을 쉽게 인식할 것이다. 여기서, 몇몇 실시예들은 또한 머신 또는 컴퓨터 판독 가능하고, 명령들의 머신 실행 가능하거나 컴퓨터 실행 가능한 프로그램들을 인코딩하는 프로그램 저장 디바이스들, 예를 들면, 디지털 데이터 저장 매체들을 포함하는 것이 의도되고, 여기서 상기 명령들은 상기 상술된 방법들의 단계들 일부 또는 모두를 수행한다. 프로그램 저장 디바이스들은, 예를 들면, 디지털 메모리들, 자기 디스크들 및 자기 테이프들과 같은 자기 저장 매체들, 하드 드라이브들, 또는 광학적 판독 가능 디지털 데이터 저장 매체들일 수 있다. 실시예들은 또한 상술된 방법들의 상기 단계들을 수행하도록 프로그래밍된 컴퓨터들을 포함하도록 의도된다.

부호의 설명

[0068]

210 : 파노라마 리프레이머

211 : 비디오 입력 인터페이스

220 : 움직임 어댑터

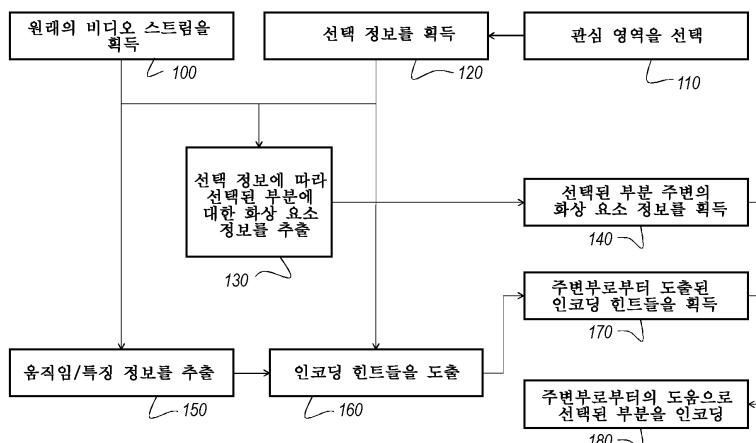
221 : 움직임 및 특징 정보 입력 인터페이스

230 : 힌트 인코더

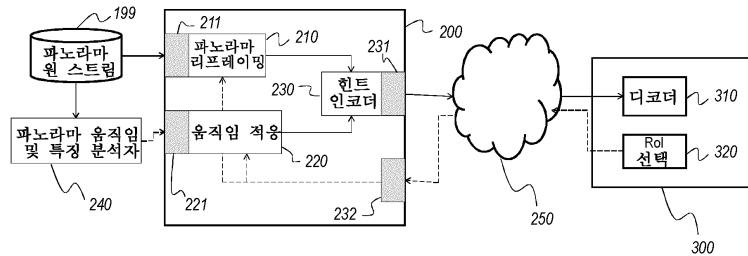
232 : 선택 정보 입력 인터페이스

도면

도면1



도면2



도면3

