



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0079085
(43) 공개일자 2017년07월10일

- | | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)
 <i>H04N 19/17</i> (2014.01) <i>H04N 19/172</i> (2014.01)
 <i>H04N 19/174</i> (2014.01) <i>H04N 19/176</i> (2014.01)
 <i>H04N 19/18</i> (2014.01)</p> <p>(52) CPC특허분류
 <i>H04N 19/17</i> (2015.01)
 <i>H04N 19/172</i> (2015.01)</p> <p>(21) 출원번호 10-2015-0189267
 (22) 출원일자 2015년12월30일
 심사청구일자 2015년12월30일</p> | <p>(71) 출원인
 이화여자대학교 산학협력단
 서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)</p> <p>(72) 발명자
 강제원
 서울특별시 마포구 독막로 145, 110동 502호(창전동, 서강쌍용예가아파트)</p> <p>(74) 대리인
 특허법인(유한)아이시스</p> |
|---|--|

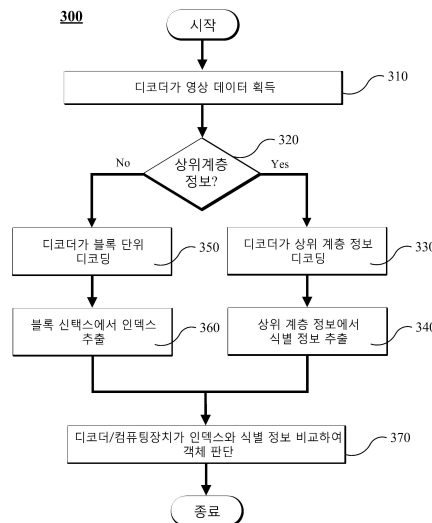
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 **객체를 포함하는 영상에 대한 인코딩 방법 및 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법**

(57) 요약

객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법은 디코더가 디코딩할 영상 데이터를 획득하는 단계, 상기 디코더가 영상 데이터 중 상위 계층 정보를 구분하는 단계, 상기 디코더가 상기 상위 계층 정보를 디코딩하는 단계 및 상기 디코더가 상기 상위 계층 정보에서 상기 영상에 포함된 객체에 대한 식별 정보를 추출하는 단계를 포함한다.

대표도 - 도7



(52) CPC특허분류

H04N 19/174 (2015.01)

H04N 19/176 (2015.01)

H04N 19/18 (2015.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2014R1A1A2056587

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 기초연구사업(학술진흥)-일반연구자지원사업-기본연구

연구과제명 미래 스마트 미디어 이미지/비디오 서비스를 위한 정보 부호화 기술

기 여 율 1/1

주관기관 이화여자대학교 산학협력단

연구기간 2014.11.01 ~ 2017.04.30

명세서

청구범위

청구항 1

인코더가 블록 단위로 영상을 인코딩하는 단계; 및

상기 인코더가 상기 영상의 디코딩에 사용되는 상위 계층 정보에 상기 영상에 포함된 객체에 대한 식별 정보를 저장하고, 상기 상위 계층 정보를 인코딩하는 단계를 포함하는 객체를 포함하는 영상에 대한 인코딩 방법.

청구항 2

제1항에 있어서,

상기 상위 계층 정보는 VPS(Video Parameter Set), SPS(Sequence Parameter Set), PPS(Picture Parameter Set) 및 슬라이스 헤더를 포함하는 집합 중 적어도 어느 하나인 객체를 포함하는 영상에 대한 인코딩 방법.

청구항 3

제1항에 있어서,

상기 상위 계층 정보는 전체 비디오 시퀀스, 하나의 비디오 프레임, 복수의 비디오 프레임, 하나의 슬라이스 및 복수의 슬라이스를 포함하는 단위 집합 중 적어도 하나의 단위마다 포함되는 객체를 포함하는 영상에 대한 인코딩 방법.

청구항 4

제1항에 있어서,

상기 식별 정보는 상기 객체에 대한 인덱스와 상기 인덱스가 나타내는 객체의 정보를 저장한 테이블 또는 상기 객체의 종류를 나타내는 인덱스를 포함하는 영상에 대한 인코딩 방법.

청구항 5

제1항에 있어서,

상기 인코더는 상기 영상을 블록 단위로 인코딩하는 단계에서 상기 블록에 포함된 상기 객체에 대한 정보를 나타내는 인덱스를 포함하여 상기 영상을 블록 단위로 인코딩하는 객체를 포함하는 영상에 대한 인코딩 방법.

청구항 6

제5항에 있어서,

상기 인덱스는 PU(Prediction Unit), CU(Coding Unit) 및 TU(Transform Unit)를 포함하는 집합 중 적어도 하나의 신택스(syntax)에 포함되는 객체를 포함하는 영상에 대한 인코딩 방법.

청구항 7

디코더가 디코딩할 영상 데이터를 획득하는 단계;

상기 디코더가 영상 데이터 중 상위 계층 정보를 구분하는 단계;

상기 디코더가 상기 상위 계층 정보를 디코딩하는 단계; 및

상기 디코더가 상기 상위 계층 정보에서 상기 영상에 포함된 객체에 대한 식별 정보를 추출하는 단계를 포함하는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 8

제7항에 있어서,

상기 상위 계층 정보는 VPS(Video Parameter Set), SPS(Sequence Parameter Set), PPS(Picture Parameter Set) 및 슬라이스 헤더를 포함하는 집합 중 적어도 어느 하나인 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 9

제7항에 있어서,

상기 상위 계층 정보는 전체 비디오 시퀀스, 하나의 비디오 프레임, 복수의 비디오 프레임, 하나의 슬라이스 및 복수의 슬라이스를 포함하는 단위 집합 중 적어도 하나의 단위마다 포함되는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 10

제7항에 있어서,

상기 식별 정보는 상기 객체에 대한 인덱스와 상기 인덱스가 나타내는 객체의 정보를 저장한 테이블 또는 상기 객체의 종류를 나타내는 인덱스를 포함하는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 11

제7항에 있어서,

상기 디코더는 블록 단위로 영상을 디코딩하면서 상기 블록에 포함된 상기 객체에 대한 정보를 나타내는 인덱스를 추출하는 단계를 더 포함하는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 12

제11항에 있어서,

상기 인덱스는 PU(Prediction Unit), CU(Coding Unit) 및 TU(Transform Unit)를 포함하는 집합 중 적어도 하나의 인덱스(syntax)에 포함되는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

청구항 13

제11항에 있어서,

상기 식별 정보는 상기 객체에 대한 인덱스와 상기 인덱스가 나타내는 객체의 정보를 저장한 테이블을 포함하고, 상기 디코더는 상기 테이블과 상기 블록에서 추출한 인덱스를 비교하여 상기 블록에 포함된 상기 객체에 대한 정보를 결정하는 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 영상 인코딩 및 디코딩 기법에 관한 것이다.

배경 기술

[0002] 영상에 포함된 객체를 분석하는 이미지 처리 기술은 다양한 분야에서 이용되고 있다. 예컨대, 이미지 처리 기술은 영상에서 사람의 얼굴 등을 인식하여 인증을 수행하는데 사용되기도 한다. 다만 기본적으로 이미지 처리 기술은 주어진 영상을 일정한 방식으로 처리하고 분석하는 과정이 필요하다.

[0003] 한편 상용 영상/동영상 코덱은 이미지 영상을 작은 블록으로 분할하여 인코딩하고, 블록 단위로 해당 블록의 압축 정보를 디코딩한다. 그러나 모든 블록이 공통으로 공유하는 정보 등은 블록마다 부호화 하여 전송하는 대신 전체 비디오 시퀀스, 비디오 프레임 혹은 슬라이스 단위로 정보를 전송하기도 한다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국공개특허 US2013-0195178호

발명의 내용

해결하려는 과제

[0005] 이하 설명하는 기술은 인코더가 영상에 포함된 객체에 대한 정보를 영상 데이터와 함께 인코딩하고, 디코더가 영상과 함께 객체에 대한 정보를 디코딩하여 객체에 대한 정보를 전송하는 기법을 제공하고자 한다.

과제의 해결 수단

[0006] 영상에 대한 인코딩 방법은 인코더가 블록 단위로 영상을 인코딩하는 단계 및 상기 인코더가 상기 영상의 디코딩에 사용되는 상위 계층 정보에 상기 영상에 포함된 객체에 대한 식별 정보를 저장하고, 상기 상위 계층 정보를 인코딩하는 단계를 포함하는 객체를 포함한다.

[0007] 영상 데이터에 대한 디코딩 방법은 디코더가 디코딩할 영상 데이터를 획득하는 단계, 상기 디코더가 영상 데이터 중 상위 계층 정보를 구분하는 단계, 상기 디코더가 상기 상위 계층 정보를 디코딩하는 단계 및 상기 디코더가 상기 상위 계층 정보에서 상기 영상에 포함된 객체에 대한 식별 정보를 추출하는 단계를 포함한다.

발명의 효과

[0008] 이하 설명하는 기술은 영상에 포함된 객체에 대한 정보를 사전에 인코딩하여 디코더가 영상에 대한 정보를 빠르고 쉽게 파악할 수 있다.

도면의 간단한 설명

[0009] 도 1은 객체를 포함하는 영상에 대한 인코딩 방법에 대한 순서도의 예이다.

도 2는 HEVC의 비트 스트림의 구조를 도시한 예이다.

도 3은 상위 계층 정보에 저장되는 식별 정보에 대한 예이다.

도 4는 객체 정보를 저장하는 테이블 및 영상에 포함되는 객체에 대한 예이다.

도 5는 객체를 포함하는 영상에 대한 인코딩 방법에 대한 순서도의 다른 예이다.

도 6은 상위 계층 정보에 저장되는 식별 정보에 대한 다른 예이다.

도 7은 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법에 대한 순서도의 예이다.

도 8은 영상에 포함된 객체 정보가 전달되는 시스템에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

[0010] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0011] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0012] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 실시된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.

- [0013] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.
- [0014] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0016] 이하 설명하는 기술은 배경 및 객체를 포함하는 영상을 일정하게 코딩하여 전달하는 과정에서 객체에 대한 정보를 추가적으로 전달하는 기법에 관한 것이다. 인코더가 사전에 객체에 대한 정보를 추가적으로 인코딩하고, 디코더는 객체에 대한 정보를 디코딩하여 영상 처리(image processing) 과정 없이도 곧바로 해당 영상에 어떤 객체가 있는지 여부를 알 수 있다.
- [0017] 정지 영상 또는 동영상을 코딩과 관련된 대부분의 표준 기술은 영상 프레임을 복수의 블록으로 구분하고, 블록마다 코딩을 수행한다. 예컨대, H.264, HEVC/H.265 등과 같은 표준은 영상을 블록 단위로 처리하는 기법에 대해 정의하고 있다. 이하 설명의 편의를 위해 HEVC/H.265를 기준으로 설명하고자 한다. 다만 이하 설명하는 기술이 반드시 HEVC/H.265에만 적용되는 것은 아니고, 모든 블록 기반 영상 처리 기법에 적용할 수 있다.
- [0019] 도 1은 객체를 포함하는 영상에 대한 인코딩 방법(200)에 대한 순서도의 예이다. 인코더는 소스 영상 및 객체 정보를 획득한다(110). 소스 영상은 인코더가 인코딩할 데이터이고, 객체 정보는 소스 영상에 포함된 객체에 대한 정보를 의미한다. 따라서 인코더는 사전에 영상에 포함된 객체에 대한 정보를 획득해야 한다. 이하 영상에 포함된 객체에 대한 정보를 객체 정보라고 명명한다.
- [0020] 객체 정보는 객체의 색상, 객체의 형태(모양), 객체의 크기 등과 같은 물리적 정보를 포함할 수 있다. 나아가 객체 정보는 객체의 종류를 식별하는 정보를 포함할 수 있다. 인코더는 소스 영상을 분석하여 객체 정보를 생성할 수 있다. 또는 별도의 컴퓨터 장치가 영상을 분석하여 추출한 객체 정보를 생성할 수도 있다. 또는 영상을 분석하는 관리자가 객체 정보를 입력하여 객체 정보를 생성할 수도 있다. 이 밖에도 해당 분야의 통상의 지식을 가진 자가 주지하고 있는 다양한 방법으로 객체 정보를 마련할 수 있다. 이하 객체 정보는 사전에 생성된 것을 전제로 설명한다.
- [0021] 인코더는 상위 계층 정보에 객체에 대한 식별 정보를 저장하고, 상위 계층 정보를 인코딩 한다(120). 상위 계층 정보에 대해 먼저 설명한다. 전술한 바와 같이 블록 기반 영상 코딩 기법은 실제 영상 데이터(색상값)를 블록 단위로 처리한다. 복수의 블록이 슬라이드를 구성하고, 또 복수의 블록이 프레임을 구성하고, 복수의 프레임이 전체 비디오 시퀀스를 구성한다. 따라서 블록의 상위 계층은 슬라이드, 프레임, 비디오 시퀀스 등을 포함한다. 한편 영상에서 전체 비디오 시퀀스, 하나 이상의 비디오 프레임 또는 하나 상의 슬라이드가 공유하는 정보가 있다. 예컨대, 영상에 포함된 객체를 직접 나타내는 데이터와는 관계 없는 파라미터(parameter) 등이 존재한다. 이 경우 영상 코딩 기법은 모든 블록에 해당 파라미터를 저장하지 않고, 일정한 상위 계층에 해당 정보를 저장한다. 복수의 블록이 공통으로 공유하는 정보는 전체 비디오 시퀀스, 비디오 프레임 혹은 슬라이스 단위로 정보를 전송하는 것이다. 결국 상위 계층 정보는 개별 블록이 아닌 블록의 상위 계층에 저장되는 복수의 블록에 공통되는 특정 정보를 의미한다.
- [0022] 도 2는 HEVC의 비트 스트림의 구조를 도시한 예이다. H.264도 사실 거의 동일한 비트 스트림 구조를 갖는다. 도 2를 기준으로 상위 계층 정보를 설명한다. 도 2는 설명의 편의를 위해 HEVC의 비트 스트림 중 일부를 도시한 예이다.
- [0023] HEVC는 네트워크를 통한 영상 데이터를 효과적으로 전달하기 위해 NAL(Network Abstraction Layer)로 구성된다. NAL은 "NAL unit + Payloads"로 구성된다. Payloads에는 RBSP(Raw Byte Sequence Payload)가 위치한다. RBSP에는 VPS, SPS, PPS 등의 정보를 알려주는 파라미터 셋과, VCL(Video Coding Layer)에 해당하는 Slice 데이터 RBSP가 포함될 수 있다. HEVC는 슬라이스 헤더가 PPS를 PPS가 SPS를, SPS가 VPS를 참조하는 구조를 갖는다.

VPS(Video Parameter Set)는 공간적 스케일러빌리티 (Spatial Scalability)와 화질적 스케일러빌리티 (SNR Scalability)를 지원하는 HEVC 스케일러블 비디오 표준(SHVC)과 멀티 뷰 확장을 지원하는 HEVC 다시점 비디오 부호화 표준을 효과적으로 지원하기 위한 파라미터 셋이다. SPS(Sequence Parameter Set)는 시퀀스 단위의 수준에서 전송되어지는 선택 요소들을 포함한다. 예를 들어, 영상의 너비와 높이 정보와 입력영상의 비트 심도와 같은 정보들은 매 프레임마다 전송되지 않고, SPS 라는 상위 수준에서 전송된다. PPS(Picture Parameter Set)는 여러 장의 픽처 레벨에서의 선택요소들을 표현하는 파라미터 셋이다.

[0024] 도 2에서 상위 계층 정보는 VPS, SPS, PPS를 포함한다. 또한 실제 영상 데이터를 저장하는 VCL NAL인 슬라이스도 슬라이스 헤더라는 상위 계층 정보가 존재한다. 따라서 상위 계층 정보는 VPS, SPS, PPS, 슬라이스 헤더 등을 포함한다. 나아가 향후 등장할 영상 표준에서도 복수의 블록에 공통되는 정보가 저장되는 영역이 있다면, 역시 해당 영역은 상위 계층 정보에 포함된다. 따라서 인코더는 VPS, SPS, PPS, 슬라이스 헤더 등 중 적어도 하나에 식별 정보를 저장할 수 있다.

[0025] 도 1에 대한 설명으로 돌아가면, 인코더는 상위 계층 정보에 객체에 대한 식별 정보를 저장한다(120). 상위 계층 정보에 저장되는 객체에 대한 정보는 다양한 종류를 가질 수 있다. 대표적으로 식별 정보는 객체의 종류를 나타내는 정보(예컨대, 건물, 사람 등)와 해당 정보에 대응하는 인덱스(index)를 포함하는 테이블일 수 있다. 예컨대, 사람은 인덱스 '1'로 나타내고, 건물은 인덱스 '2'로 나타낼 수 있다.

[0026] 인코더는 블록 단위로 영상을 인코딩하는 과정에서 해당 블록에 객체가 존재하는지 여부를 판단할 수 있다. 인코더는 사전에 특정 블록에 포함된 객체의 존재 여부 및 객체의 종류 등에 대한 정보를 획득해야 한다. 인코더가 직접 영상을 분석하여 블록에 포함된 객체를 확인할 수도 있고, 인코더가 별도의 영상 처리 장치가 사전에 분석한 결과를 이용할 수도 있을 것이다. 인코더는 블록에 포함되는 특정 선택요소(syntax)에 인덱스를 추가한다(130). 블록에 포함되는 인덱스는 식별 정보에 테이블에서 정의한 인덱스 정보와 대응된다. 예컨대, 인코더는 해당 블록에 특정 객체가 존재하면 객체에 대한 인덱스를 추가한다. 한편 인코더는 객체가 존재하지 않는다면 객체가 존재하지 않음을 나타내는 인덱스를 추가할 수 있다. 또한 인코더는 해당 블록에 존재하는 배경에 대한 인덱스도 추가할 수도 있다. 이후 인코더는 인덱스를 포함한 블록에 대한 통상적인 블록 단위 인코딩을 수행한다(140).

[0027] 도 1에 도시하지 않았지만 동영상 경우 다음 프레임에서도 동일한 작업이 수행될 수 있다. 나아가 다음 프레임에서도 동일한 위치의 블록에 동일한 객체가 존재하는 경우 인코더는 다음 프레임에서는 블록 단위로 인덱스를 추가하지 않을 수도 있다. 즉, 인코더는 프레임 단위 상위 계층에 객체가 동일하다는 정보만을 추가하여 코딩할 수도 있다. 나아가 객체가 이동하는 경우에도 인코더는 객체의 이동 정보(움직임 벡터)를 참조하고, 객체가 위치하는 새로운 블록에 인덱스를 추가하지 않을 수도 있을 것이다.

[0028]

[0029] 도 3은 상위 계층 정보에 저장되는 식별 정보에 대한 예이다. 도 3은 도 1에서 인코더가 상위 계층 정보에 저장되는 식별 정보에 대한 예이다. 도 3은 인덱스 및 객체 정보를 포함하는 테이블 형태의 식별 정보에 대한 예이다. 도 3(a)를 살펴보면, 인덱스 0은 객체가 없음을 나타내고, 인덱스 1은 객체가 사람인 경우이고, 인덱스 2는 객체가 자동차인 경우이고, 인덱스 3은 객체가 건물인 경우이다. 인코더는 이와 같은 테이블을 상위 계층에 저장하여 인코딩하고, 블록에 대한 코딩을 수행할 때 객체가 존재하는 경우 해당 인덱스를 같이 저장한다.

[0030] 도 3(b)는 차량을 관리하는 시스템에 전달하기 위한 식별 정보에 해당한다. 인덱스 0은 객체가 없는 경우이고, 인덱스 1은 등록된 차량 번호 1을 나타내고, 인덱스 2는 등록된 차량 번호 2를 나타내고, 인덱스 3은 미등록된 차량 번호를 나타낸다. 따라서 차량을 관리하는 시스템은 인코더가 전달한 데이터에서 인덱스 및 개체 정보를 포함하는 테이블 및 블록에 포함된 인덱스 정보를 디코딩하기만 하면, 해당 영상에 등록되지 않은 특정 차량이 있는지 여부를 확인할 수 있다. 즉, 차량을 관리하는 시스템은 영상 처리 과정 없이 매우 빨리 미등록 차량을 찾아낼 수 있다.

[0031] 도 3(c)는 침입자를 관리하는 시스템에 전달하기 위한 식별 정보에 해당한다. 인덱스 0은 객체가 없는 경우이고, 인덱스 1은 사전에 인증된 얼굴 1을 나타내고, 인덱스 2는 사전에 인증된 얼굴 2를 나타내고, 인덱스 3은 등록되지 않은 침입자(후보)를 나타낸다. 예컨대, 건물을 관리하는 시스템은 CCTV가 전송하는 데이터에서 인덱스 및 개체 정보를 포함하는 테이블 및 블록에 포함된 인덱스 정보를 디코딩하기만 하면 침입자를 촬영한 영상을 쉽게 찾을 수 있다.

[0032] 도 3(d)는 사람을 인식하는 시스템에 전달하기 위한 식별 정보에 해당한다. 사람을 인식하는 시스템은 얼굴 등

의 특정 영역(또는 복수의 영역)을 기준으로 사람을 인증할 수 있다. 또는 사람을 인식하는 시스템은 손 영역을 기준으로 특정한 인터페이스 명령을 생성할 수도 있다. 인덱스 0은 원하는 영역(ROI)이 없는 경우를 나타내고, 인덱스 1은 눈 영역을 나타내고, 인덱스 2는 입 영역을 나타내고, 인덱스 3은 손 영역을 나타낸다. 하나의 프레임에 존재하는 복수의 블록은 각각 서로 다른 신체 영역을 포함할 수 있다. 이 경우 사람을 인식하는 시스템은 전체 프레임에 대한 이미지 처리를 하지 않고, 인덱스 및 개체 정보를 포함하는 테이블 및 블록에 포함된 인덱스 정보를 디코딩하기만 하면 특정한 블록에 원하는 영역이 존재하는 것을 쉽게 파악할 수 있다.

[0033] 도 3은 주로 객체의 형태적 특징에 기반한 객체의 종류를 나타내는 객체 정보에 대한 예이다. 나아가 전술한 바와 같이 객체 정보는 객체의 크기, 객체의 색상, 객체의 형태의 종류 등과 같은 물리적 정보를 포함할 수도 있다.

[0035] 도 4는 객체 정보를 저장하는 테이블 및 영상에 포함되는 객체에 대한 예이다. 도 4(a)는 도 1의 인코더가 상위 계층 정보에 저장하는 식별 정보에 대한 예이다. 도 4(b)는 도 1의 인코더가 영상에서 객체가 위치하는 블록을 인코딩하는 예를 설명하기 위한 것이다.

[0036] 도 4(a)는 도심지에서 거리를 촬영하는 CCTV(이 경우 CCTV가 인코더에 해당) 또는 CCTV에 연결된 인코더가 사용하는 식별 정보에 대한 예이다. 인덱스 0은 객체가 없는 경우이고, 인덱스 1은 객체가 건물인 경우이고, 인덱스 2는 객체가 버스인 경우이고, 인덱스 3은 객체가 승용차인 경우이고, 인덱스 4는 객체가 사람인 경우이고, 인덱스 5는 객체가 동물인 경우이다.

[0037] 도 4(b)는 CCTV가 촬영한 하나의 프레임에 대한 예이다. 도 4(b)에서 굵은 사각형으로 표시한 것은 하나의 블록을 나타낸다. 영상 코딩 표준에 따라 다양한 형태의 블록을 사용하지만 도 4(b)는 설명의 편의를 위해 사각형 형태로 블록을 표시하였다. 도 4(b)에서 좌측 상단에 위치한 블록 A는 4개의 서브 블록(A1, A2, A3 및 A4)으로 구분된다. 블록 A는 모두 건물을 포함한다. 따라서 인코더는 블록 A 전체에 대한 선택스에 건물을 나타내는 인덱스 1을 저장하고, 해당 블록을 인코딩할 수 있다. 또는 인코더는 블록 A1, A2, A3 및 A4 각각에 대해 건물을 나타내는 인덱스 1을 저장하고, 해당 블록을 인코딩할 수도 있다. 도 4(b)의 A 블록 아래에 위치한 블록 B는 2개의 서브 블록(B1 및 B2)로 구분된다. 블록 B는 모두 버스를 포함한다. 따라서 인코더는 블록 B 전체 또는 서브 블록(B1 및 B2 각각)에 대한 선택스에 버스를 나타내는 인덱스 2를 저장하고 해당 블록을 인코딩할 수 있다. 도 4(b)의 B블록 우측에 위치한 블록 C는 자동차를 포함한다. 인코더는 블록 C에 대한 선택스에 인덱스 3을 저장하고 해당 블록을 인코딩할 수 있다. 도 4(b)의 B 블록 아래에 위치한 블록 D는 동물(개)을 포함한다. 인코더는 블록 D에 대한 선택스에 인덱스 5를 저장하고 해당 블록을 인코딩할 수 있다. 도 4(b)의 C 블록 아래에 위치한 블록 E는 사람을 포함한다. 인코더는 블록 E에 대한 선택스에 인덱스 4를 저장하고 해당 블록을 인코딩할 수 있다.

[0038] 인코더가 객체에 대한 인덱스를 저장하는 블록 선택스는 다양할 수 있다. 예컨대, PU(Prediction unit), CU(Coding unit), TU(Transform unit) 등에 대한 선택스에 인덱스 정보가 포함될 수 있다. 아래 표 1은 Coding unit syntax에 인덱스 정보가 포함되는 예이다. 아래 표 1은 CU의 선택스 중 앞부분 일부만을 나타낸 것이다. 아래 선택스 중 cu_object_index가 객체에 대한 인덱스를 나타내는 부분이다.

표 1

[0039]

```
coding_unit( x0, y0, log2CbSize ) {
    cu_object_index
    if( transquant_bypass_enabled_flag )
        cu_transquant_bypass_flag
    if( slice_type != I )
        cu_skip_flag[ x0 ][ y0 ]
    nCbS = ( 1 << log2CbSize )
    if( cu_skip_flag[ x0 ][ y0 ] )
        prediction_unit( x0, y0, nCbS, nCbS )
    else {
        if( slice_type != I )
            pred_mode_flag
        if( CuPredMode[ x0 ][ y0 ] != MODE_INTRA || log2CbSize == MinCbLog2SizeY )
            .
            .
            .
    }
```

[0041]

한편 인코더와 디코더가 인덱스에 해당하는 객체 정보를 사전에 공유할 수도 있다. 이 경우 도 1과는 달리 인코더가 객체에 대한 식별 정보를 상위 계층 정보에 저장하지 않을 수도 있다. 인코더는 블록에 대한 인덱스에 객체에 대한 인덱스만을 저장하여 인코딩하고, 디코더는 해당 인덱스만을 디코딩할 수 있다. 사전에 디코더가 해당 인덱스가 어떤 객체 정보를 나타내는지 알고 있기 때문에 특정 객체가 영상에 있는지 여부를 확인할 수 있다.

[0042]

나아가 인덱스와 객체 정보를 매칭한 테이블이 별도의 서비스 서버에 위치할 수 있다. 이 경우 서비스 서버는 디코더로부터 디코딩된 정보를 전송받고, 디코딩된 인덱스를 기준으로 특정 영상에 특정 객체가 포함되었는지 여부를 빠르게 파악할 수 있다.

[0044]

도 5는 객체를 포함하는 영상에 대한 인코딩 방법(200)에 대한 순서도의 다른 예이다. 인코더는 소스 영상 및 객체 정보를 획득한다(210). 인코더는 특정한 단위(내지 영역)를 기준으로 해당 단위 내에 객체가 존재하는지 여부를 판단할 수 있다. 예컨대, 특정한 단위는 하나의 블록, 복수의 블록, 하나의 슬라이드, 복수의 슬라이드, 하나의 프레임, 복수의 프레임 등과 같이 다양할 수 있다. 인코더가 직접 영상을 분석하여 특정한 단위에 포함된 객체를 확인할 수도 있고, 인코더가 별도의 영상 처리 장치가 사전에 분석한 결과를 이용할 수도 있을 것이다.

[0045]

특정한 단위 내에 특정 객체가 존재하는 경우, 인코더는 상위 계층 정보에 해당 객체에 대한 식별 정보를 저장하고, 상위 계층 정보를 인코딩한다(220). 상위 계층 정보는 인코더가 객체가 존재하는지 파악하는 특정한 단위의 종류에 따라 달라질 수 있다. 예컨대, 슬라이드가 단위라면 인코더는 슬라이드 헤더에 식별 정보를 저장할 수도 있다. 물론 인코더는 그 경우에도 VPS, SPS, PPS 등에 식별 정보를 저장할 수도 있다. 특정한 단위 내에 특정 객체가 존재하지 않는 경우 i)인코더는 해당 단위에 특정 객체가 존재하지 않는다는 별도의 식별 정보를 상위 계층 정보에 저장할 수 있다. 또는 ii)인코더가 상위 계층 정보에 식별 정보를 기재하지 않을 수도 있다. 이 경우 사전에 정의된 약속에 따라 디코더가 상위 계층 정보에 식별 정보가 없는 경우 해당 특정한 단위에 특정 객체가 없는 것으로 판단할 수 있다. 마지막으로 인코더는 영상 데이터를 블록 단위로 인코딩한다(230).

[0046]

도 5의 인코딩 방법(200)은 도 1의 인코딩 방법(100)과 달리 인코더가 블록의 인덱스에 객체에 대한 인덱스를 저장하지 않았다. 인코더는 상위 계층 정보에 특정한 식별 정보만을 저장하여 인코딩하였다. 이 경우 디코더는 특정 슬라이드 단위, 프레임 단위 또는 전체 비디오 시퀀스 단위로 해당 단위 내에 특정 객체가 존재하는지 여부를 알 수 있다.

[0047]

도 6은 상위 계층 정보에 저장되는 식별 정보에 대한 다른 예이다. 도 6은 도 5의 인코더가 상위 계층 정보에

저장하는 식별 정보에 대한 예이다.

- [0048] 도 6(a)는 인덱스 1과 3만을 포함하는 식별 정보에 대한 예이다. 도 6(a)의 경우 인코더와 디코더가 특정 인덱스와 객체 정보를 매칭한 정보를 사전에 공유하는 것을 전제로 한다. 예컨대, 디코더는 도 6(a)와 같은 인덱스 정보를 상위 계층 정보로부터 추출하고, 슬라이드 단위, 프레임 단위 또는 전체 비디오 시퀀스 단위에 인덱스 1에 해당하는 자동차와 인덱스 3에 해당하는 사람이라는 객체가 존재하는 것을 파악할 수 있다.
- [0049] 도 6(b)는 인덱스와 인덱스에 매칭되는 테이블 및 인덱스를 포함하는 식별 정보에 대한 예이다. 인코더는 테이블과 인덱스를 상위 계층 정보에 저장하여 인코딩하고, 디코더는 상위 계층 정보에서 테이블과 인덱스를 디코딩한다. 디코더는 슬라이드 단위, 프레임 단위 또는 전체 비디오 시퀀스 단위로 해당 단위 내에 인덱스 1에 해당하는 사람과 인덱스 3에 해당하는 건물이라는 객체가 존재하는 것을 파악할 수 있다. 물론 전술한 바와 같이 디코더는 테이블과 인덱스만을 디코딩하고, 별도의 컴퓨팅 장치가 영상에 특정 객체가 있는 것을 파악할 수도 있다.
- [0051] 도 7은 객체 정보를 포함하는 영상 데이터에 대한 디코딩 방법(300)에 대한 순서도의 예이다. 도 7은 도 1의 인코딩 방법(100)으로 인코딩한 영상 데이터를 디코딩하는 방법에 대한 예이다. 디코더는 압축된 영상 데이터를 획득한다(310). 디코더는 저장 매체(하드디스크, 메모리 카드 등)에 저장된 영상 데이터를 디코딩할 수도 있고, 유선 또는 무선 네트워크로 전달되는 영상 데이터를 디코딩할 수도 있다. 디코더는 영상 데이터에 상위 계층 정보가 있는지 확인한다(320). HEVC 표준 기술에 따라 디코더는 NAL 단위로 데이터를 처리한다. 도 2에 도시한 바와 같이 non VCL NAL 중 상위 계층 정보에 해당하는 VPS, SPS, PPS 등이 존재한다면 디코더는 해당 상위 계층 정보를 디코딩한다(330). 또한 디코더는 또 다른 상위 계층 정보에 해당하는 슬라이스 헤더를 디코딩한다(330). 슬라이스 세그먼트는 VCL NAL에 해당한다. 나아가 표준에 정의되지 않은 새로운 또는 다른 상위 계층 정보를 사용할 수도 있을 것이다.
- [0052] 디코더는 상위 계층 정보를 디코딩하면서 상위 계층 정보에 포함된 식별 정보를 추출할 수 있다(340). 다양한 상위 계층 정보 중 어떤 상위 계층 정보를 사용할지는 사전에 정의될 수 있다. 디코더는 VCL NAL 중 슬라이스 헤더 다음에 오는 실제 영상 데이터를 블록 단위로 디코딩한다(350). 디코더는 블록 선택스에서 인덱스를 추출할 수 있다(360).
- [0053] 이후 디코더 또는 디코더로부터 디코딩한 데이터를 전달받은 컴퓨팅 장치가 블록에 포함된 인덱스와 식별 정보인 테이블(인덱스와 객체 정보 매칭된 정보)를 비교하여 해당 영상에 객체가 존재하는지 또는 어떤 객체가 존재하는지 판단한다(370). 컴퓨팅 장치는 서버, PC, 스마트폰 등과 같이 일정한 연산 능력이 있는 장치를 의미한다.
- [0055] 도 8은 영상에 포함된 객체 정보가 전달되는 시스템(400)에 대한 예이다. 도 8은 CCTV(410)가 서버(450)에 수집한 영상 데이터를 전달하는 예이다. CCTV(410)은 영상을 수집하고 인코딩한다. 물론 도 8과는 달리 별도의 인코더가 CCTV(410)가 수집한 영상을 인코딩할 수도 있다. CCTV(410)는 영상 수집을 하고 영상을 분석하여 영상에 포함된 객체 정보를 영상에 대한 정보와 함께 인코딩할 수 있다. CCTV(410A)는 비디오 시퀀스 단위로 객체에 대한 식별정보를 전달하는 예이고, CCTV(410B)는 비디오 프레임 단위로 객체에 대한 식별 정보를 전달하는 예이고, CCTV(410C)는 슬라이스 또는 복수의 블록 단위로 객체에 대한 식별 정보를 전달하는 예이다. CCTV(410)가 인코딩한 영상 및 객체 정보는 네트워크를 통해 서버(450)에 전달된다. 서버(450)는 전달된 영상 데이터를 디코딩한다. 서버(450)는 블록에 포함된 인덱스 및 상위 계층 정보에 포함된 식별정보(테이블)를 비교하여 영상에 어떤 객체가 존재하는지 쉽게 판단할 수 있다.
- [0056] 예컨대, CCTV와 같이 개별적으로 영상을 수집하는 장치가 매우 많은 환경이라면, 종래 서버는 전달받은 영상을 분석하여 객체를 식별하는데 매우 많은 시간이 소요될 수 있다. 분산된 인코더 또는 CCTV가 개별적으로 영상을 분석하여 객체 정보를 영상 데이터에 함께 인코딩하여 전달하면, 서버는 매우 빠르게 영상에 포함된 객체를 파악할 수 있다.
- [0057] 도 8에서는 CCTV(410)를 예로 도시하였지만, CCTV외에도 영상을 수집하는 다양한 장치가 존재한다. 차량에 설치된 블랙박스 카메라, 개인이 소지하는 스마트폰, IoT 환경에서 광범위한 영역에 배치되는 이미지 센서 등이 영상을 수집할 수 있다. 그리고 영상을 수집하는 장치 또는 별도의 인코더가 객체 정보가 포함된 영상 데이터를

인코딩할 수 있다.

[0058] 나아가 도시하지 않았지만, 동영상 서비스를 하는 시스템(VOD 서비스) 또는 이미지를 제공하는 시스템(앨범 서비스 등)은 기본적으로 일정한 영상 데이터를 관리한다. 예컨대, 앨범 서비스는 특정 객체를 기준으로 사진을 분류하여 서비스할 수도 있다. 이 경우 사전에 인코더가 객체 정보를 포함하는 영상 데이터를 인코딩하면, 앨범 서비스를 제공하는 서버는 별도의 이미지 처리 과정 없이 빠르게 자신이 보유한 영상을 분류하여 제공할 수 있다. 종전 시스템에서는 앨범 서비스가 실시간으로 영상을 분석하여 분류하거나, 서비스 제공 전에 영상을 분석하여 영상을 태깅(tagging)하는 과정이 필요하다. VOD 서비스는 스마트 TV 등장과 함께 보다 다양한 서비스를 제공할 수 있다. 예컨대, 시청자가 영화 또는 드라마를 보면서 배우가 입고 있는 옷, 배우가 소지한 소품, 배경에 나타나는 가구, 자동차 등을 화면에서 선택하면 해당 제품에 대한 소개(내기 구매)를 제공하는 서비스가 예상된다. 이 경우 VOD 서비스를 제공하는 시스템은 영상에 포함된 객체에 대한 정보를 이용하여 매우 빠르고 정확하게 해당 객체에 대한 정보를 시청자에게 제공할 수 있다. 이와 같은 서비스가 상용화되기 위해서는 시스템이 영상을 실시간으로 분석하여 객체를 식별하는 과정만으로는 부족할 수 있다. 매우 다양한 제품이 존재할 수 있기 때문에 정확한 분류를 위해서는 영상 인코딩 과정에서 사전에 정확한 제품에 대한 정보가 함께 인코딩되어야 할 것이다. VOD 서비스를 제공하는 시스템은 영상에 포함된 인덱스와 상위 계층 정보에 포함된 식별 정보를 이용하여 현재 사용자가 선택한 객체에 대한 정보를 정확하고 빠르게 제공할 수 있다.

[0060] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시예는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

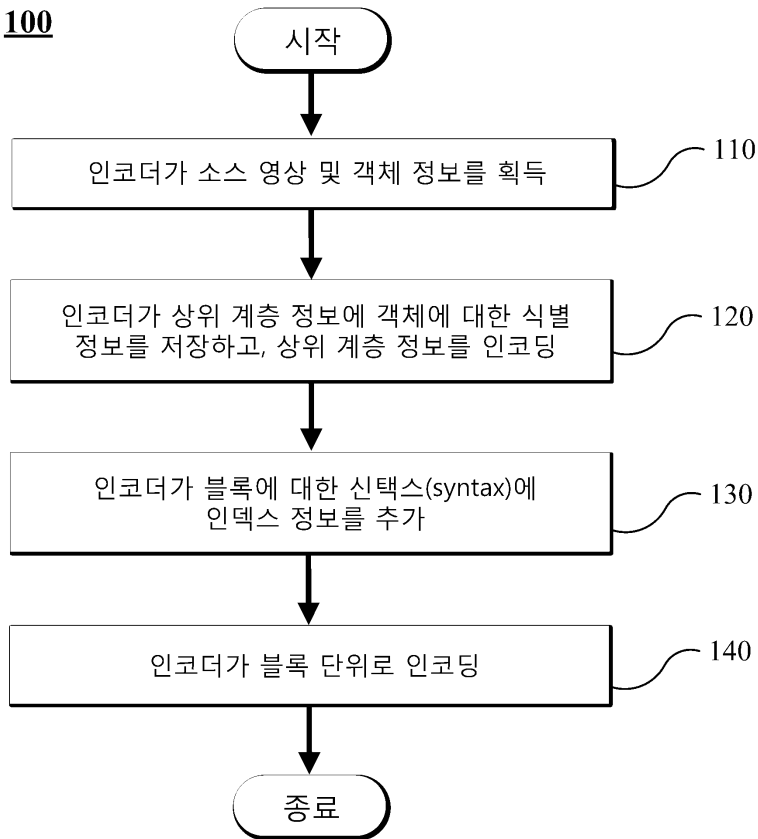
부호의 설명

[0061] 400 : 영상에 포함된 객체 정보가 전달되는 시스템
410 : CCTV
450 : 서버

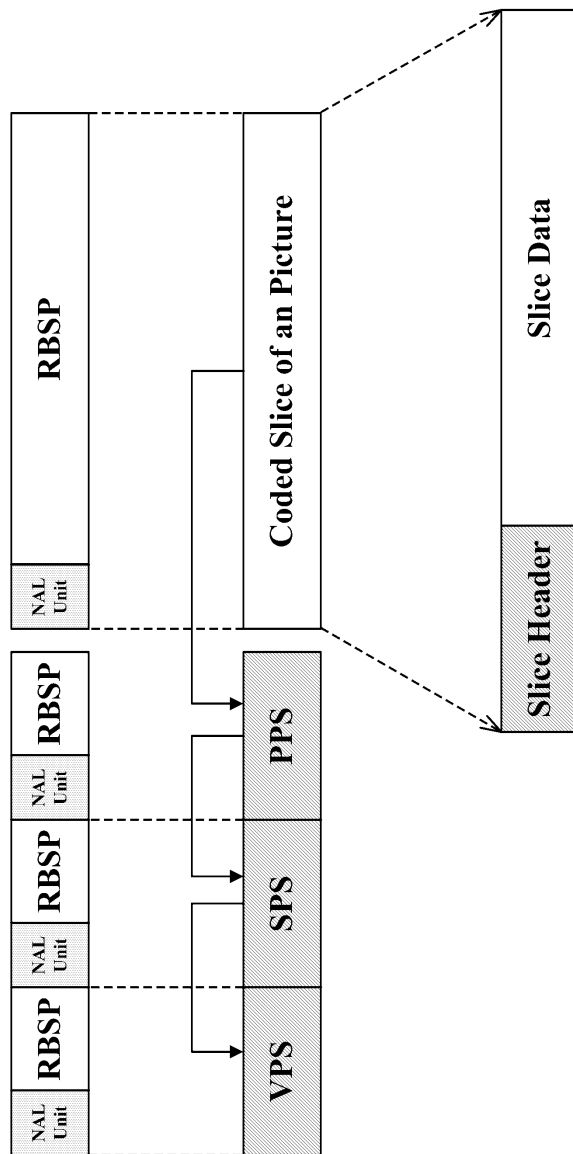
도면

도면1

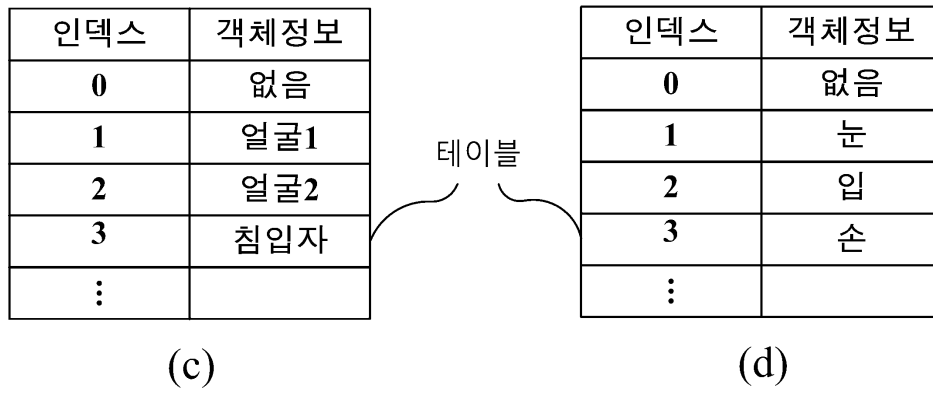
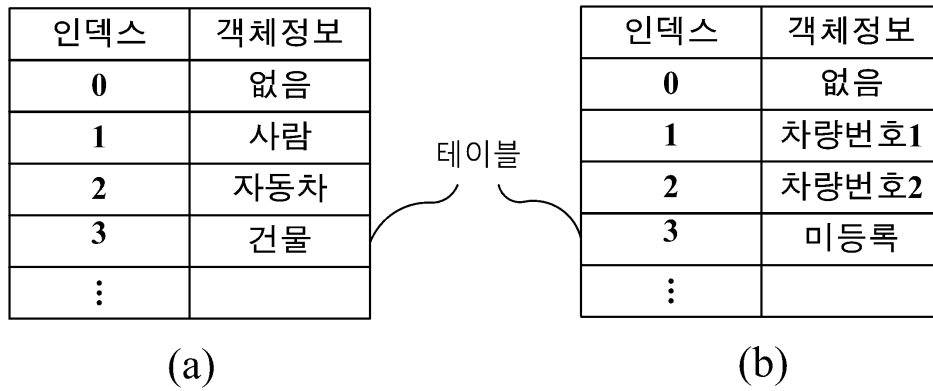
100



도면2



도면3

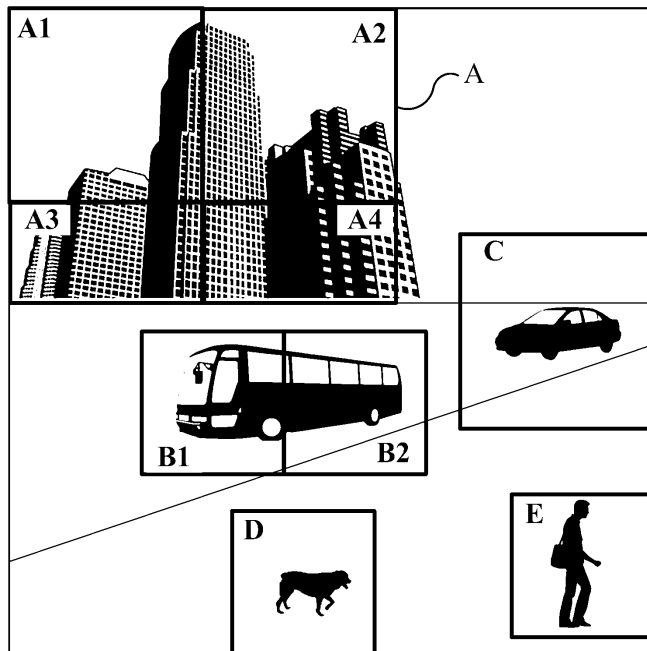


도면4

인덱스	객체정보
0	없음
1	건물
2	버스
3	승용차
4	사람
5	동물

테이블

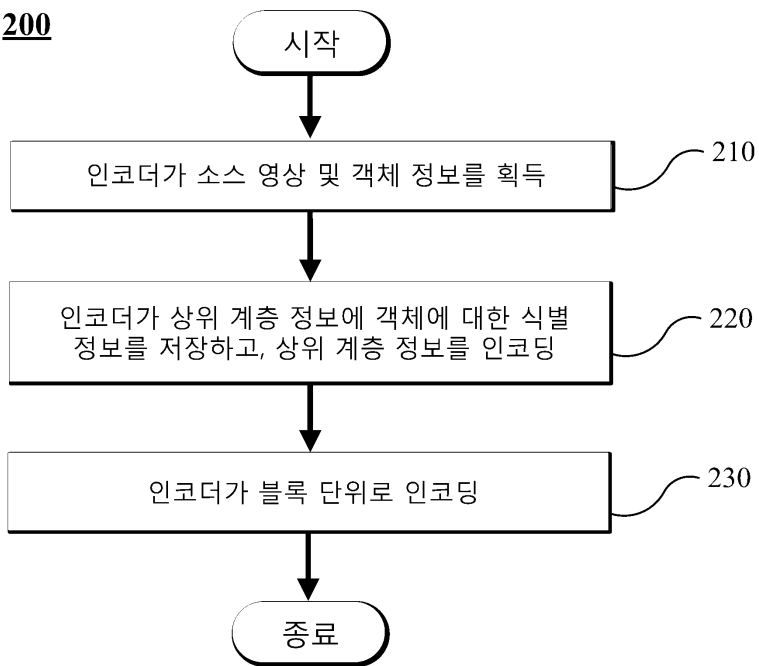
(a)



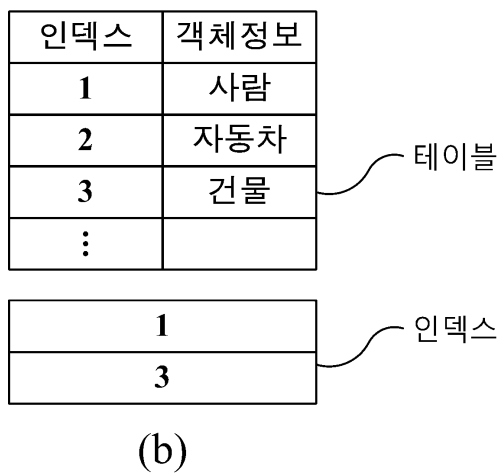
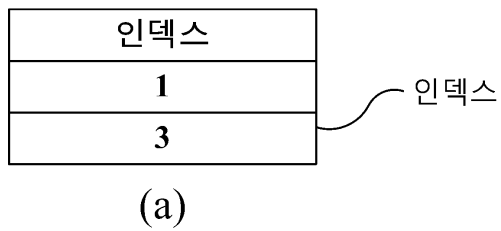
(b)

도면5

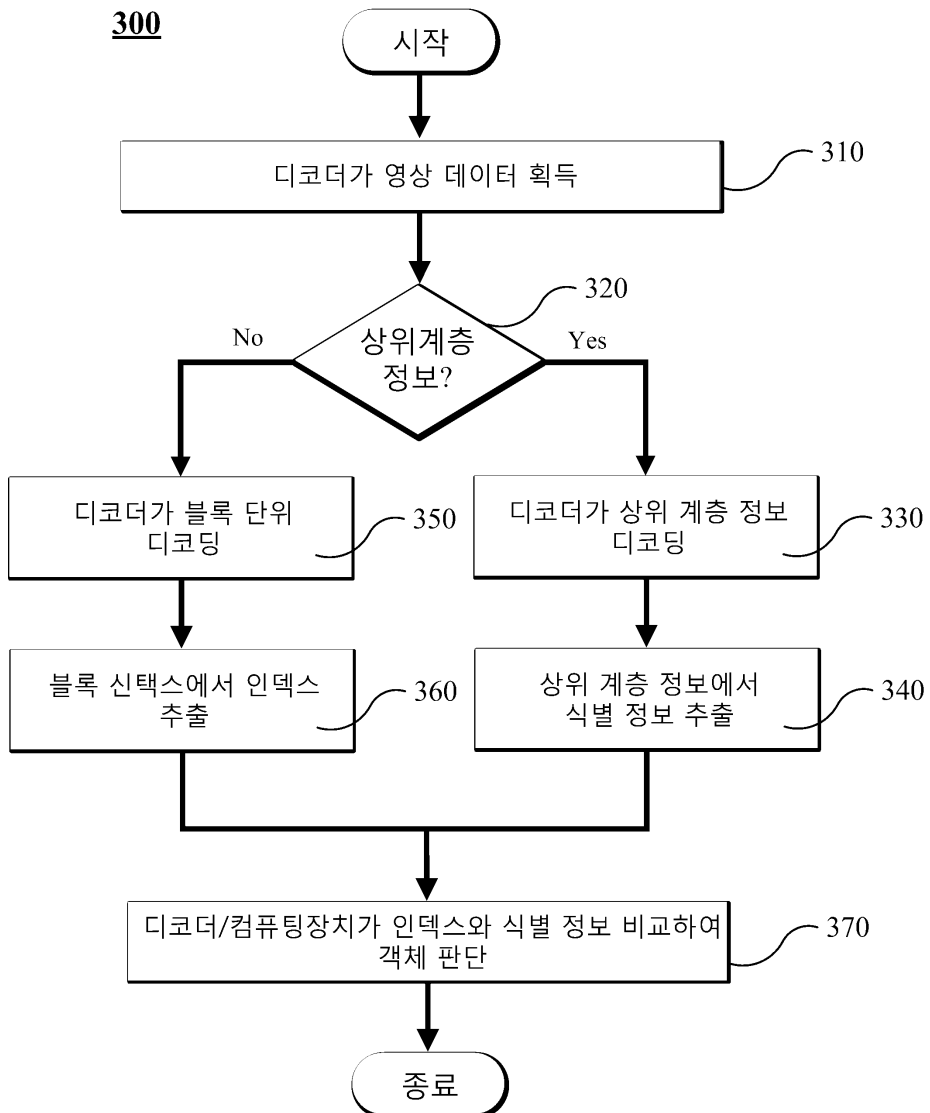
200



도면6



도면7



도면8

