



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0001963
(43) 공개일자 2019년01월08일

(51) 국제특허분류(Int. Cl.)
G06T 7/223 (2017.01) G06T 7/11 (2017.01)
G06T 7/30 (2017.01)
(52) CPC특허분류
G06T 7/223 (2017.01)
G06T 7/11 (2017.01)
(21) 출원번호 10-2017-0081609
(22) 출원일자 2017년06월28일
심사청구일자 2017년06월28일

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 104동 904호(창전동, 서강쌍용예가아파트)
김나영
서울특별시 성북구 오패산로3길 17, 104동101호(하월곡동, 동신아파트)
(74) 대리인
특허법인(유한)아이시스

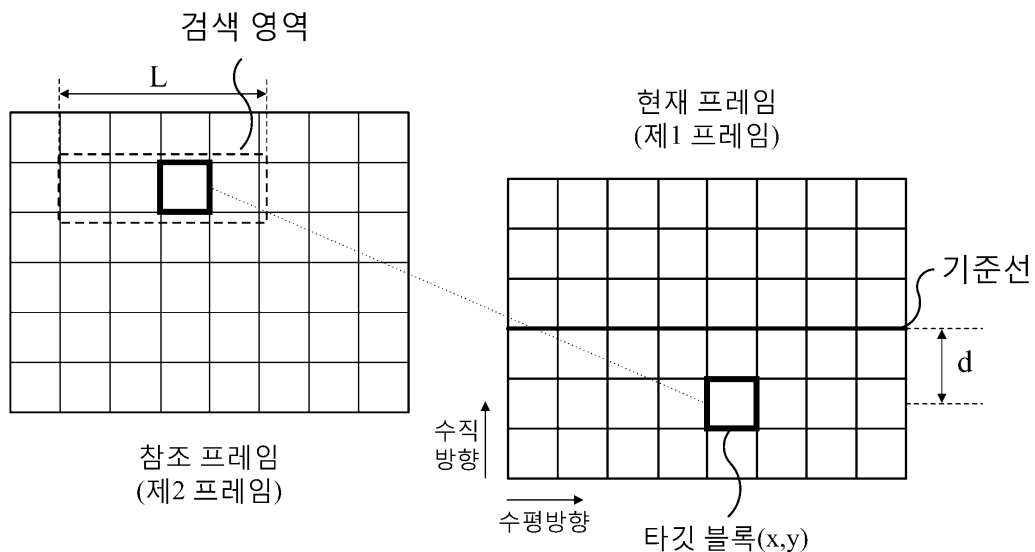
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법

(57) 요약

360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법은 360도 영상의 제1 프레임에서 타깃 블록을 결정하는 단계, 상기 제1 프레임을 기준으로 움직임 추정을 하는 제2 프레임에서 탐색 영역을 결정하는 단계 및 상기 탐색 영역에서 상기 타깃 블록과 차이가 최소인 블록을 결정하는 단계를 포함한다. 상기 탐색 영역의 크기는 상기 타깃 블록의 위치에 따라 결정된다.

대표도 - 도4



(52) CPC특허분류

G06T 7/30 (2017.01)

G06T 2207/30244 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 2016R1D1A1B03932994

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 기초연구사업(학술진흥)-기본연구지원사업

연구과제명 머신러닝 기반의 360도 비디오의 효율적 부복호 및 처리 알고리즘 개발

기 여 율 1/1

주관기관 이화여자대학교 산학협력단

연구기간 2016.11.01 ~ 2019.10.31

명세서

청구범위

청구항 1

360도 영상의 제1 프레임에서 타깃 블록을 결정하는 단계: 및

상기 제1 프레임을 기준으로 움직임 추정을 하는 제2 프레임에서 탐색 영역을 결정하는 단계를 포함하되,

상기 탐색 영역의 크기는 상기 타깃 블록의 위치에 따라 결정되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 2

제1항에 있어서,

상기 탐색 영역의 수평 방향 크기가 상기 타깃 블록의 수직 방향의 위치에 따라 결정되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 3

제1항에 있어서,

상기 탐색 영역의 수평 방향 크기가 상기 제1 프레임의 수직 방향 중심과 상기 타깃 블록의 거리에 따라 결정되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 4

제1항에 있어서,

상기 탐색 영역의 수평 방향 크기가 $\cos \theta$ 의 역수에 비례하는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

(여기서, $\theta = (0.5 - \frac{y}{H})\pi$ 이고, y 는 상기 타깃 블록의 수직 방향 위치이고, H 는 상기 제1 프레임의 수직 방향 길이임)

청구항 5

제1항에 있어서,

상기 360도 영상은 ERP(Equirectangular projection) 모델로 생성되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 6

제1항에 있어서,

상기 탐색 영역의 크기에 따라 블록 매칭을 위한 탐색점의 개수를 조절하여 상기 타깃 블록과 차이가 최소인 블록을 결정하는 단계를 더 포함하는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 7

360도 영상의 제1 프레임에서 타깃 블록을 결정하는 단계: 및

상기 제1 프레임을 기준으로 움직임 추정을 하는 제2 프레임에서 탐색 영역을 결정하는 단계를 포함하되,

상기 360도 영상을 복수의 개별 영상을 이용하여 생성되고, 상기 탐색 영역의 크기는 상기 타깃 블록을 구성하는 개별 영상의 식별자에 따라 크기가 결정되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방

법.

청구항 8

제7항에 있어서,

상기 식별자는 상기 제1 프레임에서 상기 타깃 블록의 수직 방향 위치를 나타내는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 9

제7항에 있어서,

상기 식별자는 상기 타깃 블록을 구성하는 영상의 방향을 나타내는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 10

제7항에 있어서,

상기 식별자는 상기 타깃 블록을 구성하는 영상을 촬영한 카메라를 나타내는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 11

제7항에 있어서,

상기 식별자는 상기 제1 프레임의 수직 방향 중심과 상기 타깃 블록 사이의 거리를 나타내는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 12

제7항에 있어서,

상기 탐색 영역의 수평 방향 크기가 상기 식별자에 따라 결정되는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

청구항 13

제7항에 있어서,

상기 탐색 영역의 수평 방향 크기가 상기 제1 프레임의 수직 방향 중심과 상기 타깃 블록 사이의 거리에 비례하는 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 360도 영상에 대한 움직임 추정 기법에 관한 것이다. 보다 구체적으로 이하 설명하는 기술은 360도 영상의 움직임 추정을 위한 탐색 영역 설정 기법에 관한 것이다.

배경 기술

[0002] 최근 VR(Virtual Reality) 장치 및 이를 이용한 콘텐츠가 주목받고 있다. VR 장치는 사용자에게 360도(전방위) 영상을 제공한다. 360도 영상은 다양한 모델을 통해 생성될 수 있다.

[0003] 한편 최신 비디오 코덱 표준인 HEVC(High Efficiency Video Coding)/H.265 보다 2배 이상의 부호화 성능을 갖는 차세대 비디오 부호화(Future Video Coding) 표준이 준비되고 있다. 해당 표준화 기구는 360도 영상에 대해서도 표준화를 논의하고 있다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국공개특허 US 2008/0137746

발명의 내용

해결하려는 과제

[0005] 이하 설명하는 기술은 360도 영상에 대한 움직임 추정 기법을 제공하고자 한다. 이하 설명하는 기술은 360도 영상에 대한 움직임 추정을 위하여 적응적으로 탐색 영역을 설정하는 기법을 제공하고자 한다.

과제의 해결 수단

[0006] 360도 영상을 위한 움직임 추정에서 적응적 탐색 영역 결정 방법은 360도 영상의 제1 프레임에서 타깃 블록을 결정하는 단계 및 상기 제1 프레임을 기준으로 움직임 추정을 하는 제2 프레임에서 탐색 영역을 결정하는 단계를 포함한다. 상기 탐색 영역의 크기는 상기 타깃 블록의 위치에 따라 결정된다. 또는 상기 탐색 영역의 크기는 상기 타깃 블록을 구성하는 개별 영상의 식별자에 따라 크기가 결정된다.

발명의 효과

[0007] 이하 설명하는 기술은 360도 영상에서 발생하는 왜곡을 고려하여 정확한 움직임 추정을 수행하게 한다.

도면의 간단한 설명

[0008] 도 1은 360도 영상을 생성하는 예이다.

도 2는 블록 매칭에 기반한 움직임 추정의 예이다.

도 3은 ERP 모델에서 발생하는 왜곡의 예이다.

도 4는 360도 영상에 대한 움직임 추정을 위한 탐색 영역의 예이다.

도 5는 360도 영상을 전달하는 시스템에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

[0009] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시례를 가질 수 있는 바, 특정 실시례들을 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.

[0010] 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.

[0011] 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.

[0012] 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 의

해 전달되어 수행될 수도 있음은 물론이다.

- [0013] 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.
- [0015] 최근 VR(가상 현실) 서비스가 주목 받고 있다. VR 서비스는 기본적으로 360도 영상(전방위 영상)을 사용한다. 360도 영상(전방위 영상)은 전방위 장면을 촬영하는 카메라(이하 360도 카메라라고 함)가 획득한 영상을 일정하게 결합한 영상이다. 스마트폰 또는 VR 전용 장치는 360도 영상을 이용하여 장치의 방향에 따라 특정 시점의 영상을 제공한다. 먼저 360도 영상을 생성하는 과정에 대하여 설명한다.
- [0016] 도 1은 360도 영상을 생성하는 예이다. 360도 영상을 생성하는 과정은 (a) 영상 획득, (b) 복수의 영상 결합(stitching) 및 (c) 결합된 영상의 맵핑으로 구성된다. 컴퓨터 장치는 복수의 영상을 입력받아 360도 영상을 생성한다. 경우에 따라서는 인코더가 360도 영상을 생성하고, 영상을 인코딩할 수도 있다. 360도 영상 생성은 컴퓨터 장치가 수행한다고 가정한다.
- [0017] 컴퓨터 장치는 360도 카메라를 이용하여 복수의 영상을 획득한다(a 과정). 360도 카메라는 특정 지점을 기준으로 서로 다른 방향(시점)의 영상을 획득하는 장치이다. 360도 카메라는 기준 지점을 기준으로 서로 다른 방향을 향하는 복수의 카메라 유닛을 포함한다. 도 1에서는 구 형태의 360도 카메라를 도시하였으나 360도 카메라는 다양한 형태를 가질 수 있다.
- [0018] 컴퓨터 장치는 360도 카메라가 획득한 복수의 영상을 일정한 기준으로 결합한다(b 과정). 결합 과정은 복수의 영상을 하나의 영상으로 결합하는 과정에 해당한다. 컴퓨터 장치는 인접한 영상을 서로 이어서(stitching) 하나의 결합된 영상을 생성한다.
- [0019] 마지막으로 컴퓨터 장치는 결합된 영상을 일정한 모델을 사용하여 맵핑을 수행한다(c 과정). 맵핑은 360도 방위 영상을 평면 영상으로 변환하는 과정에 해당한다. 도 1은 ERP(Equirectangular projection) 모델을 사용하여 맵핑한 예를 도시한다. ERP 모델은 360도 카메라를 기준으로 360도 구에 맵핑되는 영상을 2D로 투사하는 방식이다. 나아가 컴퓨터 장치는 다른 다양한 모델을 사용하여 영상을 맵핑할 수 있다. ERP 외에도 360도 영상을 맵핑하는 기법은 CMP(cube map projection), ISP(icosahedral projection), OCP(octahedron projection) 등이 있다.
- [0021] 동영상 코딩에서 인코더는 프레임 사이에서 객체에 대한 움직임 추정(motion estimation)을 한다. 인코더는 동일한 객체에 대한 움직임 추정을 통해 전달하는 데이터의 양을 줄인다. 움직임 추정의 결과는 움직임 벡터(motion vector)이다. 먼저 일반적인 움직임 추정에 대해 설명한다. 도 2는 블록 매칭에 기반한 움직임 추정의 예이다. 일반적으로 인코더는 블록 매칭 기반의 움직임 추정을 수행한다. 인코더는 두 개의 프레임에서 블록 단위로 가장 유사한 블록을 결정하여 움직임 추정을 수행한다. 현재 프레임(제1 프레임)에서 움직임 추정을 수행하는 대상을 타깃 블록이라고 명명한다. 인코더는 현재 프레임의 타깃 블록과 가장 유사한(차이가 가장 적은) 블록을 참조 프레임(제2 프레임)에서 찾는다. 이때 인코더는 일반적으로 참조 프레임에서 일정한 영역(탐색 영역)에 존재하는 블록을 대상으로 타깃 블록과의 유사성을 비교한다. 인코더는 탐색 영역에서 픽셀 단위 또는 서브 픽셀 단위로 블록을 움직이면서 타깃 블록과 차이가 가장 적은 블록을 선택한다. 움직임 추정 과정에 대한 자세한 설명은 생략한다.
- [0023] 이제 전술한 360도 영상의 특징에 대해 설명한다. 이하 ERP 모델을 중심으로 설명한다.
- [0024] 도 3은 ERP 모델에서 발생하는 왜곡의 예이다. 도 3(a)는 ERP 모델로 생성한 360도 영상에 대한 예이다. 도 3(a)에서 기준선은 영상에서 수직방향의 중심에 해당하는 선이다. ERP 모델로 생성된 영상은 기준선을 기준으로 상,하로 이동하게 되면 물체가 좌,우로 늘어지는 왜곡 현상이 나타난다. 전술한 바와 같이 ERP는 지구와 같은 구를 기준으로 2차원 평면을 생성하는 과정이다. 따라서 3차원 지구좌표계를 기준으로 설명하면 기준선은 적도에 해당하는 선이라고 할 수 있다. 왜곡 정도는 기준선을 기준으로 $1/\cos\phi$ 만큼 늘어난다. 여기서 ϕ 는 위도(latitude)에 해당한다. 즉 ϕ 는 지구좌표계에서 영상의 특정 영역에 대한 세로 방향 위치를 나타낸다.

[0025] 도 3(b)는 ERP 모델로 생성한 영상에서 발생한 왜곡에 대한 예이다. 영상에서 일정한 객체가 이동하여 참조 프레임에서의 위치와 현재 프레임에서의 위치가 다르다고 가정한다. 예컨대, 현재 프레임(제1 프레임)에서 제1 블록에 있는 객체가 참조 프레임에서 제2 블록에 있는 객체가 동일하다고 가정한다. 도 3(b)에 도시한 바와 같이 동일한 객체이지만 360도 영상에서 발생한 왜곡에 의해 동일한 객체가 차지하는 영역의 크기가 달라진다. ERP 모델로 생성한 영상은 적도를 기준으로 상하 이동시 좌우로 $1/\cos \phi$ 만큼 영상이 늘어지게 된다. 따라서 EPR 모델로 생성한 영상에 일반적인 블록 매칭 알고리즘을 사용하면 움직임 추정의 정확도가 떨어질 수 있다.

[0027] 이제 360도 영상을 위한 움직임 추정 과정을 설명한다. 블록 매칭을 위한 새로운 탐색 영역 설정 방법을 제안한다. 도 4는 360도 영상에 대한 움직임 추정을 위한 탐색 영역의 예이다. 인코더가 현재 프레임(제1 프레임)에서 타깃 블록을 기준으로 블록 매칭을 수행한다고 가정한다. 인코더는 참조 프레임(제2 프레임)의 일정한 탐색 영역에서 타깃 블록과 매칭되는 블록을 찾는다. 인코더는 타깃 블록의 위치를 기준으로 탐색 영역의 크기를 결정한다. 즉, 타깃 블록의 위치가 ERP 맵핑에 따라 왜곡이 심하게 발생하는 위치라면, 인코더는 탐색 영역의 크기를 더 넓게 설정한다. 탐색 영역의 위치(중심점의 위치)는 기본적으로 종래 블록 매칭에서 사용하는 위치라고 가정한다. 인코더는 탐색 영역의 가로 방향 크기(L)만을 타깃 블록의 위치에 따라 결정할 수 있다.

[0028] (i) 인코더는 타깃 블록이 기준선에서 떨어진 거리(d)에 비례하게 탐색 영역의 크기(L)를 결정할 수 있다. (ii) 블록의 크기가 사전에 결정된다면, 블록의 좌표 위치에 따라 기준선과의 거리가 결정될 수 있다. 따라서 인코더는 타깃 블록의 위치(x, y)에 따라 탐색 영역의 크기(L)를 결정할 수도 있다. 이때 블록의 좌표는 블록의 중심점을 기준으로 설정할 수 있다. 실시예에 따라 좌표는 블록의 다른 지점을 기준으로 설정될 수도 있다. (iii) 사전에 블록의 식별자와 블록의 위치를 맵핑한 정보가 있다면, 인코더는 타깃 블록의 식별자(ID)로 타깃 블록의 위치를 파악할 수도 있다.

[0029] 블록의 좌표를 사용하는 경우 인코더는 아래의 수학식 1에 따라 탐색 영역의 크기를 결정할 수 있다.

수학식 1

[0031]
$$\text{검색 영역의 크기} = \frac{1}{\cos \theta}$$

[0033] 이때 블록의 좌표에 따라 θ 는 아래의 수학식 2와 같이 결정될 수 있다.

수학식 2

[0035]
$$\theta = (0.5 - \frac{y}{H})\pi$$

[0036] 여기서, 프레임은 W(수평 방향 길이) × H(수직 방향 길이)의 크기를 갖는다고 가정한다. y는 타깃 블록의 수직 방향 좌표이다.

[0038] 다만 탐색 영역이 늘어나면 블록 매칭에 대한 복잡도가 증가할 수 있다. 이 경우 탐색 영역이 늘어나는 정도에 따라 매칭을 위한 탐색점의 개수를 조절하여 복잡도 문제를 해결할 수 있다.

[0040] 나아가 360도 영상에서 발생하는 왜곡은 ERP 모델이 아닌 다른 모델에서도 발생할 수 있다. 인코더는 다른 모델

을 사용하여 생성한 360도 영상에서도 타깃 블록의 위치에 따라 탐색 영역을 설정할 수 있다. 인코더는 모델에 따라 발생하는 왜곡을 고려하여 탐색 영역의 크기를 적응적으로 설정할 수 있다.

[0042] 도 4에서는 타깃 블록의 위치를 기준으로 설명하였다. 타깃 블록의 위치는 360도 영상을 구성하는 영상의 종류와 관련이 있다. 360도 영상은 도 1에서 설명한 바와 같이 복수의 영상으로 구성된다. 복수의 영상은 각기 서로 다른 시점(방향)에 대한 영상이다. 복수의 영상은 서로 다른 카메라(내지 카메라 유닛)가 획득한 영상이다. 360도 영상은 복수의 영상 각각이 일정한 위치에 배치된다. 맵핑 기법에 따라서 하나의 평면 영상을 구성하는 복수의 영상의 위치가 사전에 결정될 수 있다.

[0043] 따라서 인코더는 타깃 블록을 구성하는 영상의 종류에 따라 탐색 영역의 크기를 결정할 수도 있다. 다른 말로 표현하면, 인코더는 타깃 블록을 구성하는 영상의 방향(시점)에 따라 탐색 영역의 크기를 결정할 수도 있다. 또 다른 말로 하면, 인코더는 타깃 블록을 구성하는 영상을 촬영한 카메라의 종류에 따라 탐색 영역의 크기를 결정할 수도 있다. 이 경우 인코더는 복수의 영상에 대한 식별자, 카메라의 식별자 또는 타깃 블록을 구성하는 영상의 방향(시점) 중 어느 하나를 기준으로 탐색 영역의 크기를 결정할 수 있다.

[0045] 도 5는 360도 영상을 전달하는 시스템(100)에 대한 예이다. 시스템(100)은 인코더(110)와 디코더(120)를 포함한다. 인코더(110)는 360도 영상을 입력받는다. 도 5에 도시하지 않았지만 360도 카메라(50)가 획득한 영상을 별도의 컴퓨터 장치 또는 서버가 360도 영상으로 변환한다. 경우에 따라서는 인코더(110)가 360도 카메라(50)로부터 전달받은 복수의 영상을 360도 영상으로 변환할 수도 있다.

[0046] 인코더(110)는 영상에 대한 움직임 추정을 수행한다. 이때 인코더(110)는 전술한 바와 같이 타깃 블록의 위치에 따라 탐색 영역을 적응적으로 설정하여 블록 매칭을 수행한다. 인코더(110)는 360도 영상을 인코딩한다. 이때 인코더(110)는 움직임 추정 결과인 움직임 벡터를 이용하여 영상을 인코딩한다.

[0047] 디코더(120)는 인코더(110)가 전달하는 비트 스트림에서 영상 스트림을 인코딩의 역순으로 디코딩한다. 디코더(120)는 영상 데이터와 움직임 벡터 등을 이용하여 360도 영상을 복호한다.

[0049] 본 실시례 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 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

부호의 설명

[0050] 50 : 360도 카메라

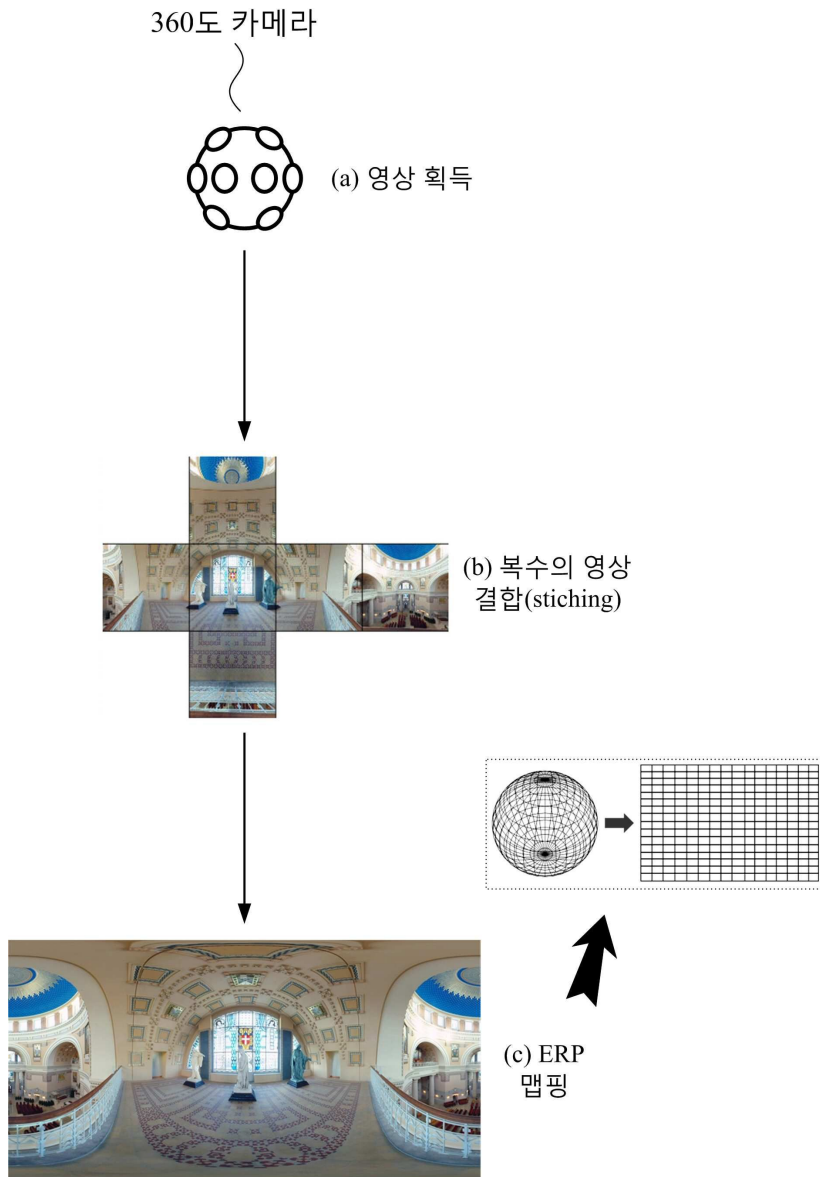
100 : 360도 영상을 전달하는 시스템

110 : 인코더

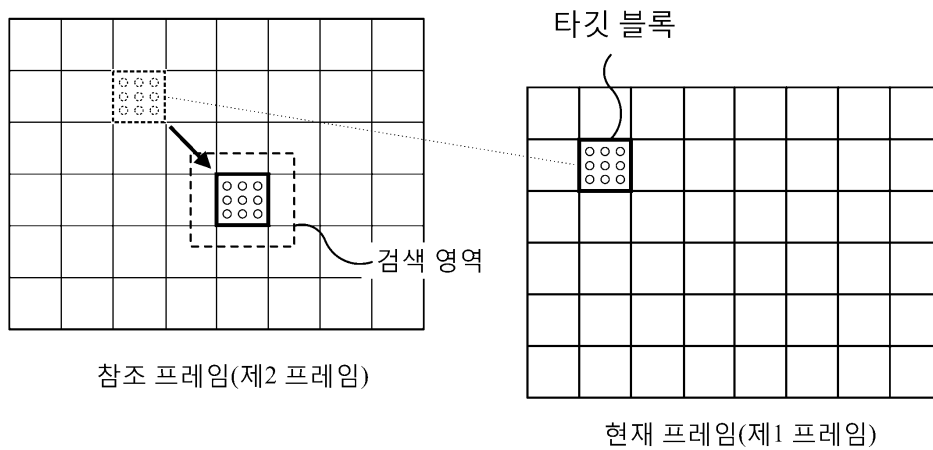
120 : 디코더

도면

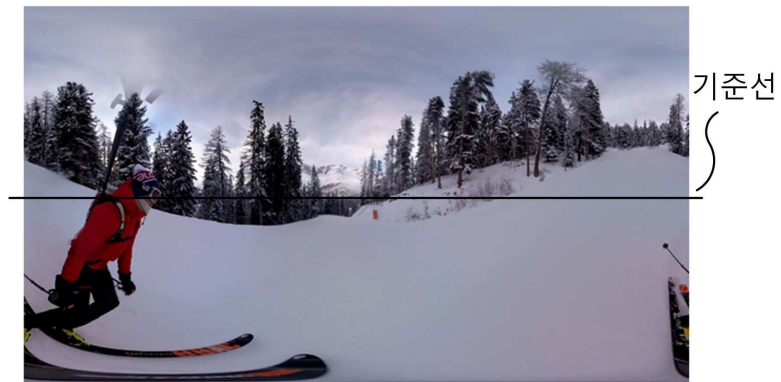
도면1



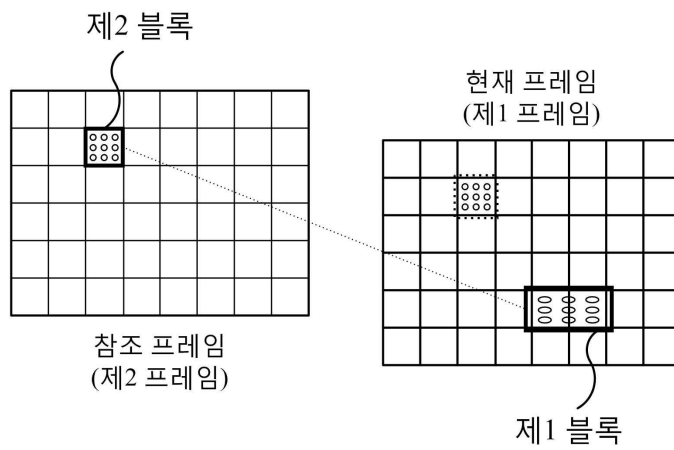
도면2



도면3

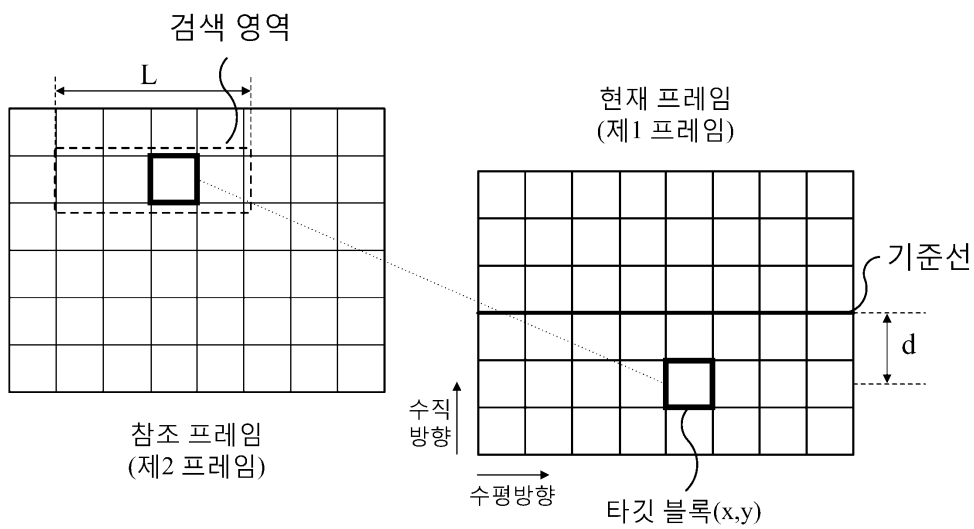


(a)

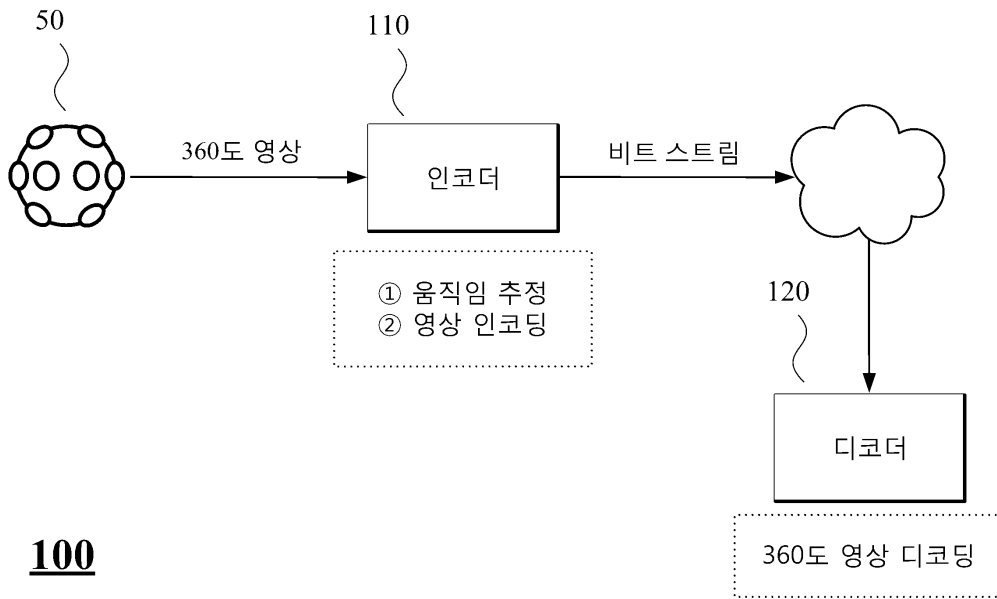


(b)

도면4



도면5



100