(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0096396
(43) 공개일자 2022년07월07일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) *G06T 7/30* (2017.01)
H04N 13/128 (2018.01) *H04N 13/156* (2018.01)
(52) CPC특허분류
G06T 3/4053 (2013.01)
G06T 3/4046 (2013.01)
(21) 출원번호 10-2020-0188790
(22) 출원일자 2020년12월31일
심사청구일자 2020년12월31일

(71) 출원인
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
(72) 발명자
강제원
서울특별시 마포구 독막로 145, 104동 904호(창전동,서강쌍용예가아파트)
김희재
경기도 성남시 분당구 동판교로 155, 702동 2003호(삼평동, 봇들마을7단지아파트)
이병욱
서울특별시 마포구 상암산로1길 24, 404동 1202호(상암동, 상암 월드컵파크 4단지)
(74) 대리인
특허법인(유한)아이시스

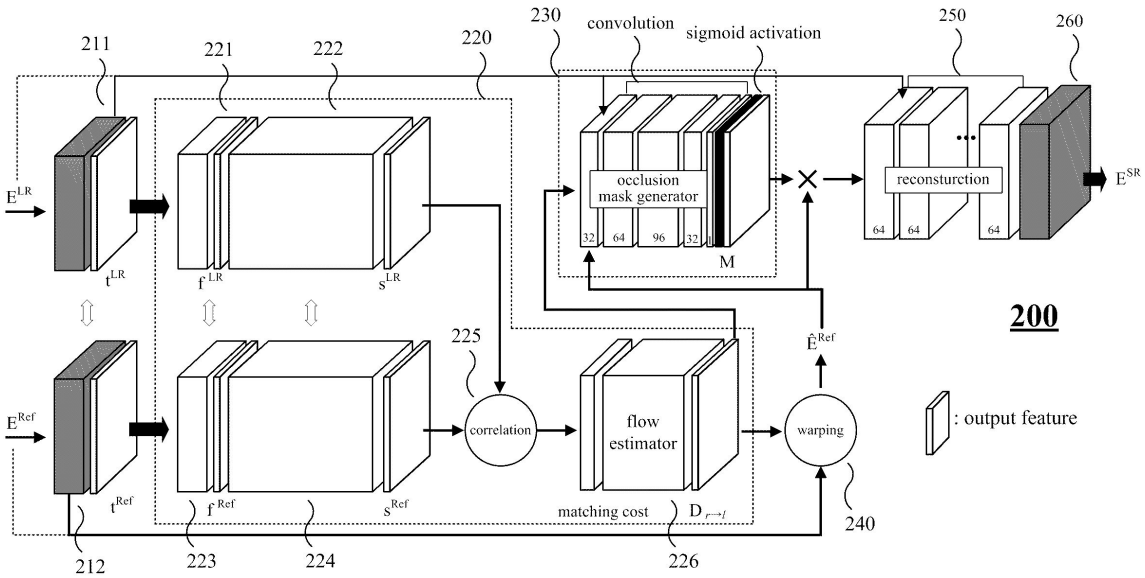
전체 청구항 수 : 총 12 항

(54) 발명의 명칭 **ERP 기반 다시점 360도 영상의 초해상화 방법 및 영상처리장치**

(57) 요 약

ERP 기반 다시점 360도 영상의 초해상화 방법은 영상처리장치가 다시점을 제공하는 복수의 360도 영상 중 저해상도인 타깃 영상 및 고해상도인 참조 영상을 입력받는 단계, 상기 영상처리장치가 상기 타깃 영상 및 상기 참조 영상을 디스패리티 추정 모델에 입력하여 상기 타깃 영상과 상기 참조 영상의 디스패리티 정보를 출력하는 단계, 상기 영상처리장치가 상기 디스패리티 정보를 기준으로 상기 참조 영상을 정렬하는 단계, 상기 영상처리장치가 상기 타깃 영상 및 상기 정렬된 참조 영상을 잔차 블록으로 구성된 재구성 계층에 입력하여 상기 타깃 영상에 대한 고해상도 영상을 출력하는 단계를 포함한다. 상기 타깃 영상과 상기 참조 영상은 ERP(Equirectangular projection) 기반하여 360도 영상으로 모델링된 영상이다.

대 표 도



(52) CPC특허분류
G06T 7/30 (2017.01)
H04N 13/128 (2021.08)
H04N 13/156 (2018.05)
G06T 2207/20016 (2013.01)
G06T 2207/20084 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711103153
과제번호	2018-0-00765-003
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	차세대(UHD)방송서비스활성화기술개발(R&D)
연구과제명	6DoF지원 초고화질 몰입형 비디오의 압축 및 전송 핵심 기술 개발
기 여 율	1/1
과제수행기관명	이화여자대학교산학협력단
연구기간	2020.01.01 ~ 2020.12.31

명 세 서

청구범위

청구항 1

영상처리장치가 다시점을 제공하는 복수의 360도 영상 중 저해상도인 타깃 영상 및 고해상도인 참조 영상을 입력받는 단계;

상기 영상처리장치가 상기 타깃 영상 및 상기 참조 영상을 디스패리티 추정 모델에 입력하여 상기 타깃 영상과 상기 참조 영상의 디스패리티 정보를 출력하는 단계;

상기 영상처리장치가 상기 디스패리티 정보를 기준으로 상기 참조 영상을 정렬하는 단계; 및

상기 영상처리장치가 상기 타깃 영상 및 상기 정렬된 참조 영상을 잔차 블록으로 구성된 재구성 계층에 입력하여 상기 타깃 영상에 대한 고해상도 영상을 출력하는 단계를 포함하되,

상기 타깃 영상과 상기 참조 영상은 ERP(Equirectangular projection) 기반하여 360도 영상으로 모델링된 영상인 ERP 기반 다시점 360도 영상의 초해상화 방법.

청구항 2

제1항에 있어서,

상기 타깃 영상 및 상기 참조 영상은 합성 360도 영상이고,

상기 영상처리장치가 실사 360도 영상인 타깃 영상 및 참조 영상 각각을 전이 계층에 입력하여 합성 360도 영상인 상기 타깃 영상 및 상기 참조 영상으로 변환하는 단계를 더 포함하는 ERP 기반 다시점 360도 영상의 초해상화 방법.

청구항 3

제1항에 있어서,

상기 고해상도 영상은 합성 360도 영상이고,

상기 영상처리장치가 상기 고해상도 영상을 전이 계층에 입력하여 실사 360도 영상으로 변환하는 단계를 더 포함하는 ERP 기반 다시점 360도 영상의 초해상화 방법.

청구항 4

제1항에 있어서,

상기 디스패리티 추정 모델은 상기 타깃 영상과 상기 참조 영상 각각에 대하여 복수의 계층 구조로 특징값을 추출하고, 동일한 계층에서 추출한 상기 타깃 영상의 특징과 상기 참조 영상의 특징에 대한 매칭 정보를 플로우 추정기에 입력하여 상기 디스패리티 정보를 출력하는 ERP 기반 다시점 360도 영상의 초해상화 방법.

청구항 5

제1항에 있어서,

상기 디스패리티 추정 모델은 상기 타깃 영상과 상기 참조 영상 각각에 대하여 아래 수식과 같이 컨볼루션 연산을 수행하여 ERP 왜곡을 반영한 특징을 추출하는 ERP 기반 다시점 360도 영상의 초해상화 방법.

$$s(i,j) = \sum_{m=-K}^{K} \sum_{n=-K}^{K} w(m,n) \cdot f(i+m, j+a_i n)$$

(s(i,j)는 위치 (i,j)에 대한 상기 컨볼루션 연산, s는 출력 특징값, w는 커널, K는 커널의 크기를 나타내는

값, $a_i = 1/(\cos((i + 0.5 - H/2)/H)\pi)$, H는 프레임의 높이)

청구항 6

제1항에 있어서,

상기 영상처리장치가 상기 타깃 영상, 상기 정렬된 참조 영상 및 상기 디스패리티 정보를 마스크 생성 모델에 입력하여 마스크를 생성하는 단계를 더 포함하고,

상기 영상처리장치가 상기 정렬된 참조 영상과 상기 마스크를 곱한 값 및 상기 참조 영상을 상기 재구성 계층에 입력하여 상기 고해상도 영상을 출력하는 ERP 기반 다시점 360도 영상의 초해상화 방법.

청구항 7

다시점을 제공하는 복수의 360도 영상 중 저해상도인 타깃 영상 및 고해상도인 참조 영상을 입력받는 입력장치;

ERP(Equirectangular projection) 영상 간 왜곡을 고려하여 360도 영상을 초해상화하는 신경망 모델을 저장하는 저장장치; 및

상기 타깃 영상 및 상기 참조 영상을 상기 신경망 모델에 입력하여 상기 타깃 영상과 상기 참조 영상의 디스패리티 정보를 생성하고, 상기 타깃 영상 및 상기 디스패리티 정보를 기준으로 정렬된 상기 참조 영상을 이용하여 상기 타깃 영상에 대한 고해상도 영상을 생성하는 연산장치를 포함하되,

상기 타깃 영상과 상기 참조 영상은 ERP 기반하여 360도 영상으로 모델링된 영상인 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

청구항 8

제7항에 있어서,

상기 타깃 영상 및 상기 참조 영상은 합성 360도 영상이고,

상기 신경망 모델은 입력단의 전이 계층을 포함하고,

상기 연산장치는 실사 360도 영상인 타깃 영상 및 참조 영상을 각각 상기 입력단의 전이 계층에 입력하여 합성 360도 영상인 상기 타깃 영상 및 상기 참조 영상을 생성하는 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

청구항 9

제7항에 있어서,

상기 고해상도 영상은 합성 360도 영상이고,

상기 신경망 모델은 출력단의 전이 계층을 포함하고,

상기 연산장치는 상기 고해상도 영상을 상기 출력단의 전이 계층에 입력하여 실사 360도 영상으로 변환하는 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

청구항 10

제7항에 있어서,

상기 신경망 모델은 디스패리티 추정 계층을 포함하고,

상기 연산장치는 상기 타깃 영상과 상기 참조 영상을 상기 디스패리티 추정 계층에 입력하여 복수의 계층 구조로 특징값을 추출하고, 동일한 계층에서 추출한 상기 타깃 영상의 특징과 상기 참조 영상의 특징에 대한 매칭 정보에서 플로우를 추정하여 상기 디스패리티 정보를 생성하는 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

청구항 11

제7항에 있어서,

상기 신경망 모델은 디스패리티 추정 계층을 포함하고,

상기 디스패리티 추정 계층은 상기 타깃 영상과 상기 참조 영상 각각에 대하여 아래 수식과 같이 컨볼루션 연산을 수행하여 ERP 왜곡을 반영한 특징을 추출하는 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

$$s(i,j) = \sum_{m=-K}^{K} \sum_{n=-K}^{K} w(m,n) \cdot f(i+m, j+a_i n)$$

(s(i,j)는 위치 (i,j)에 대한 상기 컨볼루션 연산, s는 출력 특징값, w는 커널, K는 커널의 크기를 나타내는 값, $a_i = 1/(\cos((i+0.5-H/2)/H)\pi)$, H는 프레임의 높이)

청구항 12

제7항에 있어서,

상기 신경망 모델은 상기 타깃 영상, 상기 정렬된 참조 영상 및 상기 디스패리티 정보를 입력하여 마스크를 생성하는 마스크 생성 계층 및 상기 정렬된 참조 영상을 이용하여 상기 타깃 영상의 고해상도 영상을 출력하는 재구성 계층을 포함하고,

상기 연산장치는 상기 정렬된 참조 영상과 상기 마스크를 곱한 값 및 상기 참조 영상을 상기 재구성 계층에 입력하여 상기 고해상도 영상을 출력하는 ERP 기반 360도 영상을 초해상화하는 영상처리장치.

발명의 설명

기 술 분 야

[0001] 이하 설명하는 기술은 ERP 기반의 다시점 360도 영상에 대한 초해상화 기법에 관한 것이다.

배 경 기 술

[0002] 멀티미디어 기술의 발달로 최근 실감 미디어에 대한 관심이 높아지고 있다. 360도 영상은 특정 시점을 기준으로 실제 물리 환경과 같은 전방위 영상을 제공한다. 그리고, 다시점 영상은 시점이 다른 복수의 카메라로 획득한 영상을 말한다. 다시점 360도 영상은 360도 영상을 캡쳐하는 복수의 카메라로 획득한 영상을 말한다. 2차원 영상을 360도 영상으로 맵핑하는 다양한 기법이 있다. 원통도법을 응용한 ERP(Equi-rectangular projection)가 대표적인 맵핑 기법이다.

[0003] 영상 초해상화는 다양한 기법이 연구되고 있다. 참조 영상을 이용한 초해상화 기법도 존재한다. 360도 영상 내지 다시점 360도 영상과 같이 특정 시점의 영상과 연관된 영상이 있는 경우, 참조 영상을 이용한 초해상화가 가능하다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국공개특허 US 2013-0258048호

발명의 내용

해결하려는 과제

[0005] 참조 영상 기반의 초해상화 기법은 입력 영상 간의 대응 관계를 탐색하여 참조 영상에서 저해상도 영상이 참조할 수 있는 정보를 추출하는 것이 중요하다. 그러나 종래 딥러닝 모델은 시점 차이가 클 경우에 영상 간의 디스패리티(disparity)를 보완하는데 어려움이 있고 ERP 영상의 위도 따른 비선형적 왜곡을 다루지 못한다.

[0006] 이하 설명하는 기술은 ERP 기반의 다시점 360도 영상에 대한 초해상화 기법을 제공하고자 한다.

과제의 해결 수단

[0007] ERP 기반 다시점 360도 영상의 초해상화 방법은 영상처리장치가 다시점을 제공하는 복수의 360도 영상 중 저해상도인 타깃 영상 및 고해상도인 참조 영상을 입력받는 단계, 상기 영상처리장치가 상기 타깃 영상 및 상기 참조 영상을 디스패리티 추정 모델에 입력하여 상기 타깃 영상과 상기 참조 영상의 디스패리티 정보를 출력하는 단계, 상기 영상처리장치가 상기 디스패리티 정보를 기준으로 상기 참조 영상을 정렬하는 단계 및 상기 영상처리장치가 상기 타깃 영상 및 상기 정렬된 참조 영상을 잔차 블록으로 구성된 재구성 계층에 입력하여 상기 타깃 영상에 대한 고해상도 영상을 출력하는 단계를 포함한다.

[0008] RP 기반 360도 영상을 초해상화하는 영상처리장치는 다시점을 제공하는 복수의 360도 영상 중 저해상도인 타깃 영상 및 고해상도인 참조 영상을 입력받는 입력장치, ERP 영상 간 왜곡을 고려하여 360도 영상을 초해상화하는 신경망 모델을 저장하는 저장장치 및 상기 타깃 영상 및 상기 참조 영상을 상기 신경망 모델에 입력하여 상기 타깃 영상과 상기 참조 영상의 디스패리티 정보를 생성하고, 상기 타깃 영상 및 상기 디스패리티 정보를 기준으로 정렬된 상기 참조 영상을 이용하여 상기 타깃 영상에 대한 고해상도 영상을 생성하는 연산장치를 포함한다.

발명의 효과

[0009] 이하 설명하는 기술은 입력 영상 간의 디스패리티를 기준으로 참조 영상을 초해상화 목표(target) 영상에 정렬하여 ERP 왜곡에 강인한 초해상화를 수행한다.

도면의 간단한 설명

[0010] 도 1은 다시점 360도 영상 시스템에 대한 예이다.

도 2는 초해상화를 수행하는 신경망 모델에 대한 예이다.

도 3은 초해상화를 수행하는 신경망 모델에 대한 다른 예이다.

도 4는 피라미드 구조의 360도 디스패리티 추정기에 대한 예이다.

도 5는 초해상화를 수행하는 영상처리장치에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

[0011] 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.

[0012] 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.

[0013] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 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.

[0014] 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.

[0015] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기

재하지 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] 360도 영상은 하나의 지점을 기준으로 360도 시점의 영상을 제공하는 영상 콘텐츠를 의미한다. 하나의 360도 영상은 한 지점에 배치된 360도 카메라를 통해 획득한 영상을 통해 제공된다. 2차원 영상을 360도 영상으로 맵핑하는 다양한 기법이 있다. 원통도법을 응용한 ERP가 대표적인 맵핑 기법이다.

[0017] 다시점 360도 영상은 일정 영역에 배치된 복수의 360도 카메라들을 통해 획득한 영상을 통해 제공된다. 다시점 360도 영상은 일정 영역에서 획득한 영상들을 정합하고 합성하여 이동하는 사용자에게 자유로운 시점의 영상을 제공한다.

[0018] 저해상도 영상은 해상도가 일정 기준값 미만으로 해상도가 낮은 영상을 말한다. 360도 저해상도 영상은 360도 카메라가 획득한 영상으로 해상도가 낮은 영상이다. 360도 저해상도 영상은 360도 카메라가 획득한 영상을 다운스케일링한 영상일 수 있다. 이하 저해상도 영상은 LR(low resolution) 영상이라고 표현한다.

[0019] 고해상도 영상은 해상도가 기준값 이상으로 해상도가 높은 영상을 말한다. 360도 고해상도 영상은 360도 카메라가 획득한 영상으로 해상도가 높은 영상이다. 360도 고해상도 영상은 360도 카메라가 획득한 원본 영상 또는 해상도가 기준값 이상인 영상을 의미한다. 이하 고해상도 영상은 HR(high resolution) 영상이라고 표현한다.

[0020] 초해상화(super-resolution)는 저해상도 영상을 고해상도 영상으로 변환하는 기법을 말한다. 종래 초해상화는 픽셀 보간과 같은 기법을 이용하여 수행되었다. 이하 설명하는 기술은 전술한 바와 같이 다시점 360도 영상에 대한 초해상화 기법이다. 이하 다시점 360도 영상에 대한 초해상화 기법을 보다 단순하게 MV-SR(multi-view super-resolution)이라고 표현한다.

[0021] 이하 설명하는 기술은 신경망 모델을 사용하여 초해상화를 한다. 신경망 모델은 RNN(Recurrent Neural Networks), FFNN(feedforward neural network), CNN(convolutional neural network) 등 다양한 모델이 있다. 이하 설명에서 CNN을 중심으로 설명하지만, 초해상화 기술이 특정 신경망 모델로만 구현되는 것은 아니다.

[0023] 도 1은 다시점 360도 영상 시스템에 대한 예이다. 도 1은 디코더에서 초해상화를 수행하는 시스템에 대한 예이다.

[0024] 인코더(30)는 복수의 카메라(11 내지 15)로부터 영상을 수신한다. 인코더(30)는 360도 영상 포맷에 따라 개별 영상을 인코딩할 수 있다. 이때, 인코더(30)는 복수의 360도 영상 중 일부 영상을 저해상도 영상으로 다운 스케일링할 수 있다. 저장장치(50)는 복수의 카메라(11 내지 15)가 획득한 영상으로 구성된 360도 영상들(video stream 11 ~ 15)을 저장한다. video stream(13)은 저해상도 영상이라고 가정한다.

[0025] 360도 영상들(video stream 11 ~ 15)은 네트워크를 통해 수신단으로 전송된다. 디코더(70)는 인코딩된 영상을 디코딩한다. 영상처리장치(100)는 디코딩된 영상을 이용하여 초해상화를 한다. 영상처리장치(100)는 저해상도 영상 video stream(13) 및 인접한 다른 영상(video stream 11, 12, 14 및 15 중 적어도 하나)를 이용하여 video stream(13)을 고해상도 영상으로 변환한다. 저장장치(90)는 모두 고해상도 영상인 360도 영상들(video stream 11 ~ 15)을 저장할 수 있다.

[0026] 수신단에서 디코더(70)와 영상처리장치(100)를 구분하여 표시하였다. 다만, 하나의장치가 디코딩과 초해상화를 수행할 수도 있다. 한편, 인코딩 내지 디코딩은 이하 설명하는 초해상화 과정과 연관이 없다. 영상처리장치는 저해상도 영상과 참조할 고해상도 영상만을 기준으로 초해상화를 한다.

[0027] 이하 다시점 360도 영상에 대한 초해상도를 수행하는 장치를 영상처리장치라고 가정한다. 영상처리장치는 물리적으로 다양한 형태일 수 있다. 예컨대, 영상처리장치는 VR장치, PC, 서버, 프로그램이 임베디드된 칩셋 등일 수 있다. 영상처리장치는 복수의 360도 영상을 입력받아 초해상화를 수행한다.

[0029] ERP 영상은 영상 중 특정 영역에는 일정한 왜곡을 포함한다. 다시점 영상은 서로 다른 위치에 배치된 복수의 카메라로부터 획득된다. 동일한 지점 내지 영역을 촬영한 복수의 영상이라고, 카메라의 위치에 따라 ERP 영상의 왜곡 정도가 달라진다. 따라서, 어느 하나의 360도 카메라가 획득한 영상을 다른 카메라가 획득한 영상을 기준으로 초해상화하는 것이 쉽지 않다. 따라서, 다음과 같은 초해상화 과정을 제안한다.

[0030] 복수의 360도 영상들 중 초해상화 대상인 영상을 타깃 영상이라고 명명한다. 타깃 영상은 저해상도 영상이다. 타깃 영상은 저해상도 타깃 영상과 타깃 영상을 촬영한 카메라에 인접한 카메라가 획득한 고해상도 영상을 사용하여 초해상화된다. 이때 인접한 카메라기 획득한 고해상도 영상을 참조 영상이라고 명명한다. 참조 영상은 타깃 영상이 캡쳐한 영역 전체 또는 일부를 포함한다. 도 2를 기준으로 설명하면, 타깃 영상이 카메라 13이 획득한 영상이라면, 참조 영상은 카메라 11, 12, 14 및 15 중 어느 하나일 수 있다.

[0033] 이하 타깃 영상을 초해상화하는 장치를 영상처리장치라고 명명한다. 영상처리장치는 영상 데이터 처리가 가능한 컴퓨터 장치이다. 예컨대, 영상처리장치는 PC, 스마트기기, 네트워크 상의 서버, TV 셋업박스, VR 장치 등일 수 있다.

[0035] 도 2는 초해상화를 수행하는 신경망 모델(200)에 대한 예이다. 도 2의 신경망 모델(200)은 360도 입력 영상을 입력받아 360도 고해상도 영상을 출력한다. 입력 영상은 저해상도 타깃 영상 E^{LR} 및 참조 영상 E^{Ref}이다. 출력 영상은 타깃 영상 E^{LR}이 초해상화된 결과물은 고해상도 영상 E^{SR}이다.

[0036] 도 2는 전이 계층(transfer layer)을 사용하여 실사(real) 360도 영상을 초해상화하는 모델이다. 제1 전이 계층(211)은 E^{LR}을 입력받아 합성 360도 영상 특징 t^{LR}을 출력한다. 제2 전이 계층(212)은 E^{Ref}를 입력받아 합성 360도 영상 특징 t^{Ref}을 출력한다.

[0037] 디스패리티 추정기(220)는 E^{LR}와 E^{Ref} 사이의 상관관계를 추출한다. 상관관계는 각 영상에서 추출한 특징들 중 연관된 특징들을 정렬한 정보를 포함한다.

[0038] 디스패리티 추정기(220)는 상관관계 $D_{r\rightarrow l}$를 연산하여 플로우(flow) 감독없이 참조 영상의 특징을 저해상도 영상과 함께 전달한다.

[0039] 디스패리티 추정기(220)는 타깃 영상의 t^{LR}에서 제1 특징을 추출하는 구성, 참조 영상의 t^{Ref}에서 제2 특징을 추출하는 구성, 제1 특징과 제2 특징의 상관 관계에서 플로우를 추정하는 구성을 포함한다.

[0040] 간략하게 설명하면, 제1 인코더(221)는 t^{LR}을 입력받아 기본적인 특징을 f^{LR}을 추출하고, 제1 위치 인식 컨볼루션 계층(latitude-aware convolution, 222)은 f^{LR}에서 ERP 영상 사이의 디스패리티 차이를 줄이는 역할을 한다. 제1 위치 인식 컨볼루션 계층(222)는 s^{LR}을 출력한다. 제2 인코더(223)는 t^{Ref}을 입력받아 기본적인 특징을 f^{Ref}을 추출하고, 제2 위치 인식 컨볼루션 계층(224)은 f^{Ref}에서 ERP 영상 사이의 디스패리티 차이를 줄이는 역할을 한다. 제2 위치 인식 컨볼루션 계층(224)은 s^{Ref}을 출력한다. 제1 인코더(221)와 제2 인코더(223), 제1 위치 인식 컨볼루션 계층(222)과 제2 위치 인식 컨볼루션 계층(224) 사이의 양방향 화살표는 커널 파라미터를 공유하는 관계를 나타낸다.

[0041] 영상처리장치는 s^{LR}와 s^{Ref}의 상관 연산(225)을 하고, 플로우 추정기(226)는 상관 연산한 정보를 입력받아 최종적으로 E^{LR}과 E^{Ref} 사이의 상관관계 $D_{r\rightarrow l}$을 출력한다. 영상처리장치는 $D_{r\rightarrow l}$을 이용하여 E^{Ref}를 $\hat{E}^{Ref}$로 와핑(warping)한다(240). $\hat{E}^{Ref}$는 타깃 영상 E^{LR}과 참조 영상 E^{Ref}의 ERP 왜곡을 고려하여 참조 영상을 타깃 영상에 정렬한 결과에 해당한다.

[0042] $\hat{E}^{Ref}$가 타깃과 관련하여 부정적인 특징(성능 저하 특징)을 갖는 경우, 초해상화 성능이 저하된다. 신경망 모델(200)은 이와 같은 성능 저하 특징을 제거하기 위한 구성을 더 포함할 수 있다. 마스크 생성기(230)는 t^{LR} 및 $D_{r\rightarrow l}$을 입력받아 $\hat{E}^{Ref}$에 포함되는 성능 저하 특성을 제거하기 위한 마스크 M을 생성한다.

[0043] 영상처리장치는 M과 $\hat{E}^{Ref}$을 요소별 곱셈(elementwise multiplication)하여 초해상화를 위한 최종적인 특징 데

이터를 마련할 수 있다.

[0044] 재구성 계층(250)은 (i) 저해상도 타깃 영상 E^{LR}(또는 t^{LR}) 및 (ii) $\hat{E}^{Ref}$ 또는 M과 $\hat{E}^{Ref}$을 곱한 결과를 입력받아 영상을 재구성하여 초해상화를 수행한다. 재구성 계층(250)은 잔차 블록(residual block)으로 구성될 수 있다. 잔차 블록은 입력 특징이 들어와 일정 값이 출력되는 과정에서 학습 효율에 따라 컨볼루션을 수행하거나 수행하지 않는 과정을 선택적으로 제공한다.

[0045] 제3 전이 계층(260)은 재구성 계층(250)의 출력인 초해상화된 합성 360도 영상을 실사 합성 영상으로 변환한다. 제3 전이 계층(260)은 초해상화 360 영상 E^{LR}을 출력한다. 신경망 모델(200)에서 전이 계층(211, 212 ,260)은 복잡도가 높은 실사 영상을 합성 영상 차원에서 낮은 복잡도로 초해상화를 수행하게 하다.

[0047] 도 3은 초해상화를 수행하는 신경망 모델(300)에 대한 다른 예이다. 도 3은 합성 360도 영상에 대한 초해상화 모델이다.

[0048] 입력 영상은 합성 영상이고, 저해상도 타깃 영상 t^{LR} 및 참조 영상 t^{Ref}이다. 출력 영상은 타깃 영상 t^{LR}이 초해상화된 결과물은 고해상도 영상 t^{SR}이다.

[0049] 디스패리티 추정기(310)는 t^{LR}와 t^{Ref} 사이의 상관관계를 추출한다. 상관관계는 각 영상에서 추출한 특징들 중 연관된 특징들을 정렬한 정보를 포함한다.

[0050] 디스패리티 추정기(310)는 상관관계 $D_{r\rightarrow l}$를 연산하여 플로우(flow) 감독없이 참조 영상의 특징을 저해상도 영상과 함께 전달한다.

[0051] 디스패리티 추정기(310)는 타깃 영상의 t^{LR}에서 제1 특징을 추출하는 구성, 참조 영상의 t^{Ref}에서 제2 특징을 추출하는 구성, 제1 특징과 제2 특징의 상관 관계에서 플로우를 추정하는 구성을 포함한다.

[0052] 간략하게 설명하면, 제1 인코더(311)는 t^{LR}을 입력받아 기본적인 특징을 f^{LR}을 추출하고, 제1 위치 인식 컨볼루션 계층(312)은 f^{LR}에서 ERP 영상 사이의 디스패리티 차이를 줄이는 역할을 한다. 제1 위치 인식 컨볼루션 계층(222)는 s^{LR}을 출력한다. 제2 인코더(313)는 t^{Ref}을 입력받아 기본적인 특징을 f^{Ref}을 추출하고, 제2 위치 인식 컨볼루션 계층(314)은 f^{Ref}에서 ERP 영상 사이의 디스패리티 차이를 줄이는 역할을 한다. 제2 위치 인식 컨볼루션 계층(314)는 s^{Ref}을 출력한다. 제1 인코더(311)와 제2 인코더(313), 제1 위치 인식 컨볼루션 계층(312)과 제2 위치 인식 컨볼루션 계층(314) 사이의 양방향 화살표는 커널 파라미터를 공유하는 관계를 나타낸다.

[0053] 영상처리장치는 s^{LR}와 s^{Ref}의 상관 연산(315) 하고, 플로우 추정기(326)는 상관 연산한 정보를 입력받아 최종적으로 t^{LR}과 t^{Ref} 사이의 상관관계 $D_{r\rightarrow l}$을 출력한다. 영상처리장치는 $D_{r\rightarrow l}$을 이용하여 t^{Ref}를 $\hat{t}^{Ref}$로 와핑(warping)한다(320). $\hat{t}^{Ref}$는 타깃 영상 t^{LR}과 참조 영상 t^{Ref}의 ERP 왜곡을 고려하여 참조 영상을 타깃 영상에 정렬한 결과에 해당한다.

[0054] $\hat{t}^{Ref}$가 타깃과 관련하여 부정적인 특징(성능 저하 특징)을 갖는 경우, 초해상화 성능이 저하된다. 신경망 모델(300)은 이와 같은 성능 저하 특징을 제거하기 위한 구성을 더 포함할 수 있다. 마스크 생성기(330)는 t^{LR} 및 $D_{r\rightarrow l}$을 입력받아 $\hat{t}^{Ref}$에 포함되는 성능 저하 특성을 제거하기 위한 마스크 M을 생성한다.

[0055] 영상처리장치는 M과 $\hat{t}^{Ref}$을 요소별 곱셈하여 초해상화를 위한 최종적인 특징 데이터를 마련할 수 있다.

[0056] 재구성 계층(340)은 (i) t^{LR} 및 (ii) $\hat{t}^{Ref}$ 또는 M과 $\hat{t}^{Ref}$을 곱한 결과를 입력받아 영상을 재구성하여 초해상화를 수행한다. 재구성 계층(340)은 잔차 블록으로 구성될 수 있다. 잔차 블록은 입력 특징이 들어와 일정 값이 출력되는 과정에서 학습 효율에 따라 컨볼루션을 수행하거나 수행하지 않는 과정을 선택적으로 제공한다. 재구성 계

층(340)은 초해상화된 합성 360도 영상 t^{SR}을 출력한다.

[0058] 도 2의 신경망 모델(200) 및 도 3의 신경망 모델(300)에서 ERP 영상들 사이의 왜곡을 고려하여 마련한 구성이 디스패리티 추정기이다. 디스패리티 추정기에 대하여 설명한다. 위치 인식 컨볼루션 계층(LatConv)이 ERP 왜곡에 강인한 초해상화를 위한 핵심적 구성이다.

[0059] ERP 영상에서 구형 이미지(spherical image) 기준으로 높은 위도(latitude)에 위치한 픽셀들은 경도(longitudinal) 방향으로 밀집된다. 따라서, 픽셀의 밀집도를 평준화하기 위해서는 적도 영역에서 극 영역으로 갈수록 픽셀들을 수평적 확장할 필요가 있다. LatConv이 ERP 영상의 특징 추출에서 이와 같은 동작을 수행한다.

[0060] LatConv는 위도에 따라 경도의 샘플 간격을 적응적으로 설정한다. 프레임 H(height)×W(width)의 i 번째 행(row)에서 샘플링 간격 a_i은 아래의 수학식 1과 같이 정의될 수 있다.

수학식 1

[0062] $$a_i = 1/(\cos((i + 0.5 - H/2)/H)\pi)$$

[0064] 영상처리장치는 a_i에 따라 입력 특징 맵 f를 위도에 따른 수평적 확장(scaling)할 수 있다. 영상처리장치는 각 위치 (i,j)에 대하여 아래 수학식 2와 같이 (2K + 1)×(2K + 1) LatConv 연산을 할 수 있다.

수학식 2

[0066] $$s(i,j) = \sum_{m=-K}^{K} \sum_{n=-K}^{K} w(m,n) \cdot f(i+m, j+a_i n)$$

[0067] 여기서, s는 출력 특징값이고, w는 커널이다. K는 제안된 네트워크에서 1로 설정될 수 있다. LatConv은 입력 특징 f의 차원과 관계없이 채널에 걸쳐 적용될 수 있다. LatConv은 표준적인 역전사로 훈련될 수 있다. 영상처리장치가 영상의 경계 영역에 커널을 적용할 때 수평 방향으로 순환 패딩(circular padding)하고, 수직 방향으로 거울 대칭 스킴(mirror symmetry scheme)을 사용할 수 있다.

[0069] 디스패리티 추정기에 대하여 설명한다. 일반적인 컨볼루션 계층인 인코더(221, 223, 311, 313)는 f^{LR} 및 f^{Ref}을 추출한다. 두 개의 ERP 사이의 왜곡 때문에, f^{LR} 및 f^{Ref} 사이의 상호 관계를 찾기 어렵다. LatConv을 통과하여 생산되는 s^{LR} 및 s^{Ref} 는 ERP 왜곡의 정도가 유사해진다.

[0071] 도 4는 피라미드 구조의 360도 디스패리티 추정기(400)에 대한 예이다.

[0072] 디스패리티 추정기(400)는 인코더(410), 위치 인식 컨볼루션 계층(LatConv, 420) 및 플로우 추정기(440)를 포함한다.

[0073] 인코더(410)는 각각의 입력 영상에서 특징을 추출한다. 인코더(410)는 타깃 영상과 참조 영상 각각에 대한 특징을 추출하는 개별 계층으로 구성된다. 인코더(410)의 동일 계층에서 출력되는 특징은 ERP 왜곡이 서로 다르다. 예컨대, A와 B는 구형 왜곡(spherical distortion)이 서로 다르다.

[0074] 인코더(410)는 복수의 계층에서 마치 피라미드와 같은 특징들을 추출할 수 있다. 영상처리장치는 타깃 영상과 참조 영상 각각에 대하여 인코더(410)의 동일 계층의 특징을 비교하여 디스패리티를 추정할 수 있다.

[0075] LatConv(420)는 피라미드의 각 계층에서의 특징에 대한 컨볼루션 연산을 한다. LatConv(420)는 f_l^{LR} 및 f_l^{Ref}을 입력받아 s_l^{LR} 및 s_l^{Ref}을 출력한다. l은 계층의 레벨을 의미한다. LatConv(420)는 참조 영상의 특징을 타깃 영상의 특징에 정렬한다. LatConv(420)는 위도에 따른 경도상의 특징 밀집도의 왜곡을 줄여 참조 영상을 타깃 영상에 매칭한다. s_l^{LR} 및 s_l^{Ref}은 구형 왜곡이 매칭될 수 있다. LatConv(420)의 왜곡 보정은 수학식 2와 같이 수행된다.

[0076] 영상처리장치는 각 레벨 l에서 s_l^{LR} 및 s_l^{Ref} 사이의 텍스처 특징을 매칭하고 연산한 매칭 비용(430)을 플로우 추정기(440)에 입력하여 플로우를 추정한다. 영상처리장치는 상위 피라미드 레벨에서의 특징 볼륨(volume)과 추정된 플로우를 현재 레벨의 플로우를 추정하기 위하여 결합한다. 플로우 추정기(440)는 6 계층 CNN으로 디스패리티를 추정한다. 영상처리장치는 픽셀을 대응되는 좌표에 재맵핑하여 E^{Ref}의 추정된 디스패리티를 $\hat{E}^{Ref}$에 정렬한다.

[0078] 마스크 생성기(230, 330)는 E^{LR}, $\hat{E}^{Ref}$, $D_{r\to l}$ 및 E^{LR}과 $\hat{E}^{Ref}$의 차이 절대값을 결합(concatenation)하여 입력하고, 결합마스크 M을 출력할 수 있다. 도 2 및 도 3에 도시한 바와 같이 마스크 생성기(230, 330)는 5개의 컨볼루션 계층 및 시그모이드 활성 계층으로 구성될 수 있다. 마스크 생성기(230, 330)는 영상 재구성(초해상화)을 위하여 어떤 영역이 사용되어야 하는지 결정한다. 참조 영상에서 바람직하지 않은 텍스처는 아래 수학식 3과 같이 억제(필터링)될 수 있다. 상기 수학식 3과 같은 데이터가 재구성 계층(250, 340)에 입력된다.

수학식 3

[0080) $$M\hat{E}^{Ref} + (1 - M)E^{LR}$$

[0082] 재구성 계층(250, 340)은 64개 필터를 갖는 복수의 잔차 블록으로 구성될 수 있다.

[0084] 학습 과정에서 사용되는 손실함수에 대하여 설명한다. 아래 수학식 4와 수학식 5는 각각 재구성 계층의 손실함수와 참조 영상을 타깃에 정렬하는 와핑 과정의 손실함수에 해당한다.

수학식 4

[0086] $$\mathcal{L}_{rec} = \sqrt{\|E^{GT} - E^{SR}\|^2 + \rho^2}$$

수학식 5

[0087] $$\mathcal{L}_{warp} = \sqrt{\left\|E^{GT} - \hat{E}^{Ref}\right\|^2 + \rho^2}$$

[0089] 여기서 E^{GT}는 진성값(ground truth)이고, E^{SR}은 초해상화 결과값이다. ρ는 0.01로 설정될 수 있다. 전체 손실함수는 아래 수학식 6과 같이 표현될 수 있다.

수학식 6

[0091] $$\mathcal{L} = \mathcal{L}_{rec} + \mathcal{L}_{warp}$$

[0093] 도 5는 초해상화를 수행하는 영상처리장치(500)에 대한 예이다. 영상처리장치(500)는 VR장치, PC, 스마트기기, 네트워크 서버 등과 같은 형태일 수 있다.

[0094] 영상처리장치(500)는 저장장치(510), 메모리(520), 연산장치(530), 인터페이스장치(540) 및 통신장치(550)를 포함할 수 있다.

[0095] 저장장치(510)는 영상처리장치(500)의 동작을 위한 프로그램 내지 코드를 저장할 수 있다. 저장장치(510)는 전술한 신경망 모델(200 또는 300)을 저장할 수 있다. 저장장치(510)는 신경망 모델(200 또는 300) 학습을 위한 프로그램 내지 코드를 저장할 수도 있다. 저장장치(510)는 신경망 모델이 생성한 고해상도 타깃 영상을 저장할 수 있다.

[0096] 메모리(520)는 영상처리장치(500)의 동작 과정에서 생성되는 데이터 및 정보 등을 임시 저장할 수 있다.

[0097] 인터페이스장치(540)는 외부로부터 일정한 명령 및 데이터를 입력받는 장치이다. 인터페이스장치(540)는 물리적으로 연결된 입력장치 또는 물리적인 인터페이스(키패드, 터치 패널 등)로부터 일정한 정보를 입력받을 수 있다. 인터페이스장치(540)는 신경망 모델, 신경망 모델 학습을 위한 정보, 학습 데이터 등을 입력받을 수 있다. 인터페이스장치(540)는 신경망 모델 업데이트를 위한 파라미터값을 입력받을 수도 있다. 인터페이스장치(540)는 초해상화를 위한 복수의 360도 영상을 입력받을 수 있다. 인터페이스장치(540)는 전술한 타깃 영상 및 참조 영상을 입력받을 수 있다.

[0098] 통신장치(550)는 무선 네트워크를 통해 일정한 정보를 송수신한다. 통신장치(550)는 신경망 모델, 신경망 모델 학습을 위한 정보, 학습 데이터 등을 입력받을 수 있다. 통신장치(550)는 신경망 모델 업데이트를 위한 파라미터값을 수신할 수 있다. 통신장치(550)는 신경망 모델 입력하기 위한 타깃 영상 및 참조 영상을 수신할 수 있다. 통신장치(550)는 신경망 모델이 생성한 고해상도 타깃 영상을 외부 객체에 전송할 수 있다.

[0099] 인터페이스장치(540) 및 통신장치(550)는 사용자 또는 외부 객체로부터 일정한 정보 및 데이터를 입력받을 수 있다. 따라서 인터페이스장치(540) 및 통신장치(550)를 포괄하여 입력장치라고 명명할 수 있다.

[0100] 연산장치(530)는 저장장치(510)에 저장된 프로그램 내지 코드를 이용하여 영상처리장치(500)의 동작을 제어한다. 연산장치(530)는 신경망 모델을 이용하여 초해상화를 수행한다.

[0101] 연산장치(530)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드된 칩과 같은 장치일 수 있다.

[0103] 도 2의 신경망 모델(200)을 기준으로 연산장치(530)의 동작을 설명한다.

[0104] 연산장치(530)는 실사 360도 영상인 타깃 영상 E^{LR}을 제1 전이 계층(211)에 입력하여 합성 360도 영상 t^{LR}을 생성한다. 연산장치(530)는 실사 360도 영상인 참조 영상 E^{Ref}을 제2 전이 계층(212)에 입력하여 합성 360도 영상 t^{Ref}을 생성한다.

[0105] 연산장치(530)는 제1 인코더(221)에 t^{LR}을 입력하여 특징값 f^{LR}을 출력하고, f^{LR}을 제1 위치 인식 컨볼루션 계층(222)에 입력하여 s^{LR}을 출력한다. 연산장치(530)는 제2 인코더(223)에 t^{Ref}을 입력하여 특징값 f^{Ref}을 출력하고, f^{Ref}을 제2 위치 인식 컨볼루션 계층(224)에 입력하여 s^{Ref}을 출력한다.

[0106] 연산장치(530)는 s^{LR} 및 s^{Ref}을의 특징값을 도 4와 같이 피라미드 구조와 같은 개별 계층에서 비교하고, 플로우 추정 모델(226)을 이용하여 상관관계 $D_{r\to l}$을 출력한다.

[0107] 연산장치(530)는 상관관계 $D_{r\to l}$을 기준으로 참조 영상 E^{Ref}을 재맵핑(와핑)하여 $\hat{E}^{Ref}$를 생성한다.

[0108] 연산장치(530)는 마스크 생성 모델 내지 마스크 생성 계층(230)에 타깃 영상 E^{LR}, 상관관계 $D_{r\to l}$을 및 $\hat{E}^{Ref}$을 입력하여 마스크 M을 생성한다.

[0109] 연산장치(530)는 $\hat{E}^{Ref}$과 마스크 M을 요소별로 곱셈한 결과를 재구성 계층(250)에 입력하여 합성 360도 고해상도 영상을 생성한다.

[0110] 연산장치(530)는 재구성 계층(250)의 출력값을 제3 전이 계층(260)에 입력하여 최종적인 실사 360도 영상인 E^{SR}을 생성한다.

[0112] 도 3의 신경망 모델(300)을 기준으로 연산장치(530)의 동작을 설명한다.

[0113] 연산장치(530)는 제1 인코더(311)에 합성 360도 영상인 타깃 영상 t^{LR}을 입력하여 특징값 f^{LR}을 출력하고, f^{LR}을 제1 위치 인식 컨볼루션 계층(312)에 입력하여 s^{LR}을 출력한다. 연산장치(530)는 제2 인코더(313)에 합성 360도 영상인 참조 영상 t^{Ref}을 입력하여 특징값 f^{Ref}을 출력하고, f^{Ref}을 제2 위치 인식 컨볼루션 계층(314)에 입력하여 s^{Ref}을 출력한다.

[0114] 연산장치(530)는 s^{LR} 및 s^{Ref}을의 특징값을 도 4와 같이 피라미드 구조와 같은 개별 계층에서 비교하고, 플로우 추정 모델(326)을 이용하여 상관관계 $D_{r\to l}$을 출력한다.

[0115] 연산장치(530)는 상관관계 $D_{r\to l}$을 기준으로 참조 영상 E^{Ref}을 재맵핑(와핑)하여 $\hat{E}^{Ref}$를 생성한다.

[0116] 연산장치(530)는 마스크 생성 모델 내지 마스크 생성 계층(330)에 타깃 영상 E^{LR}, 상관관계 $D_{r\to l}$을 및 $\hat{E}^{Ref}$을 입력하여 마스크 M을 생성한다.

[0117] 연산장치(530)는 $\hat{E}^{Ref}$과 마스크 M을 요소별로 곱셈한 결과를 재구성 계층(340)에 입력하여 합성 360도 고해상도 영상인 t^{SR}을 생성한다.

[0118] 한편, 연산장치(530)는 360도 고해상도 영상을 이용하여 특정 지점에 위치한 사용자에게 사용자 시점의 360도 영상을 합성할 수 있다.

[0119] 출력장치(560)는 초해상화 과정의 인터페이스 화면을 출력할 수 있다. 출력장치(560)는 초해상화된 결과물인 고해상도 영상을 출력할 수 있다.

[0121] 또한, 상술한 바와 같은 초해상화 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.

[0122] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니

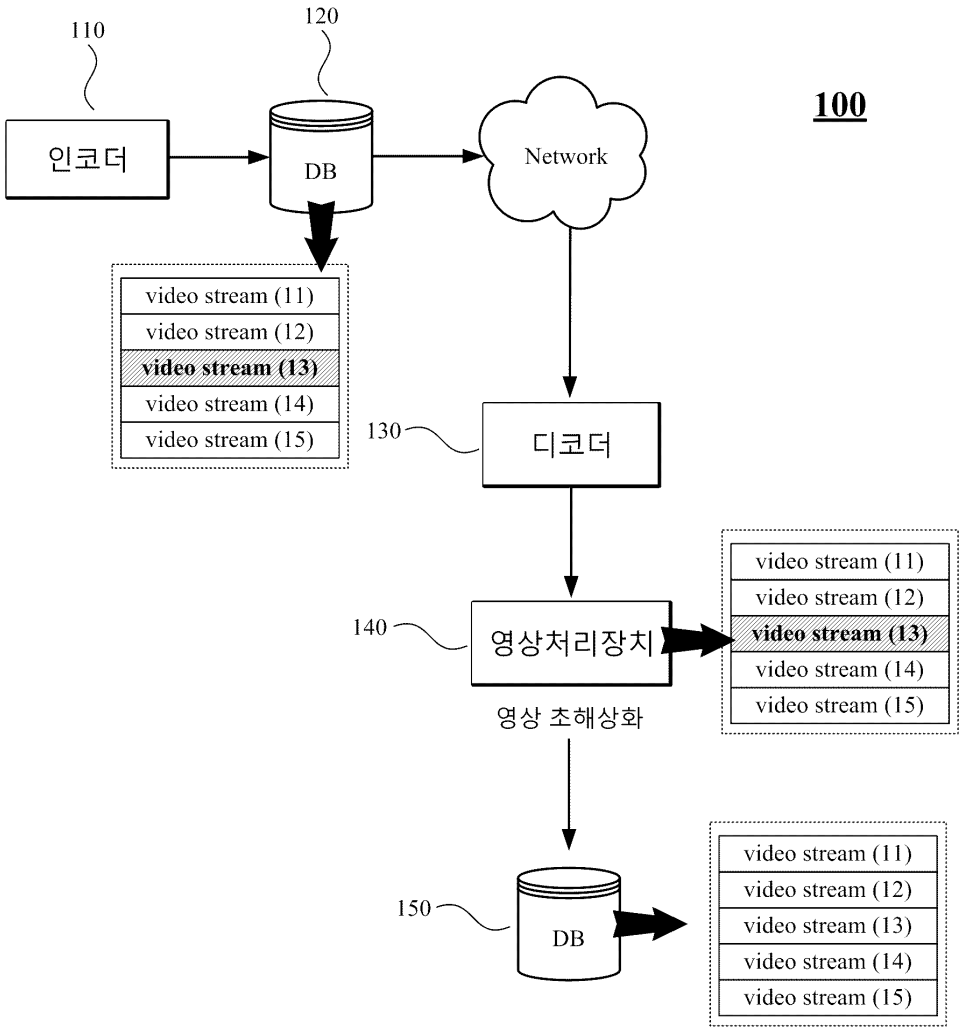
라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM (read-only memory), PROM (programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.

[0123] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.

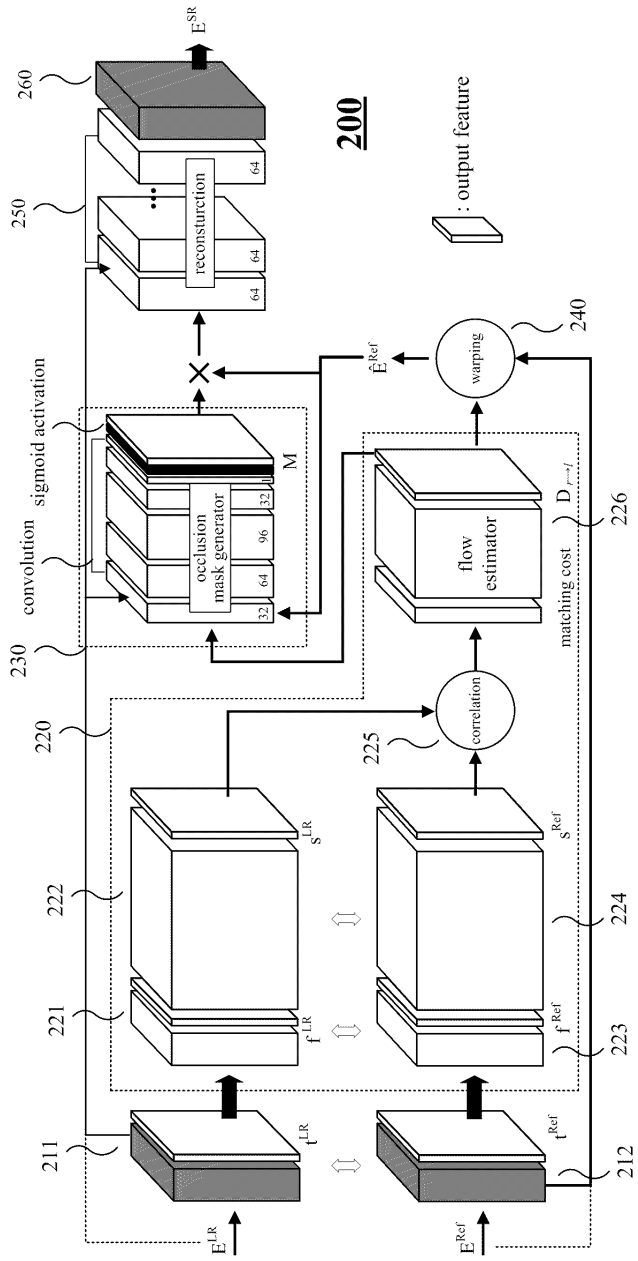
[0124] 본 실시례 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

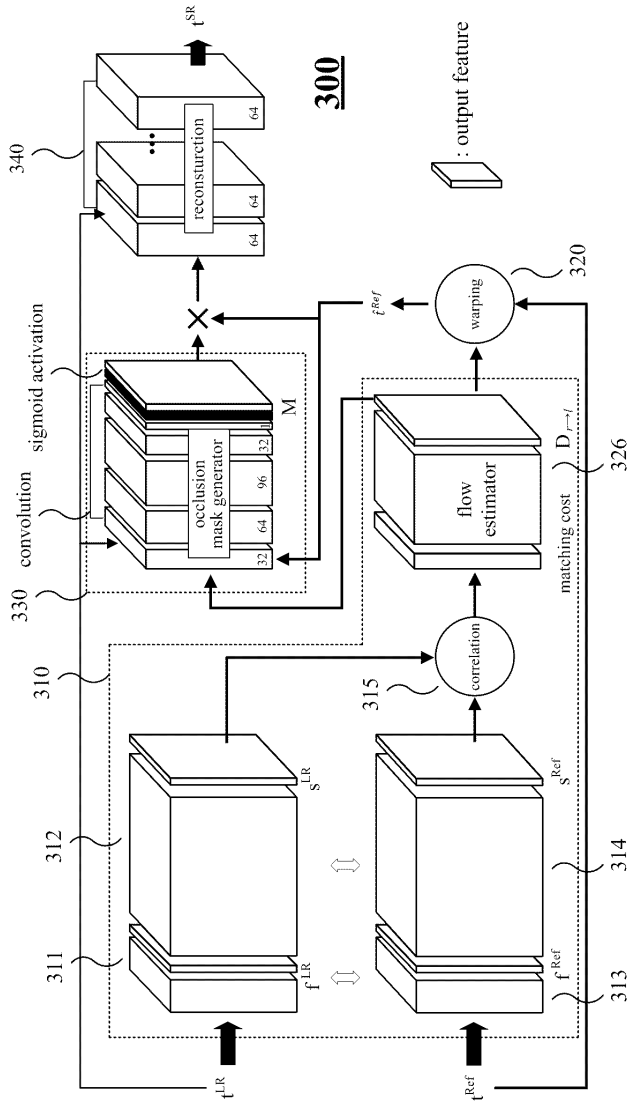
도면1



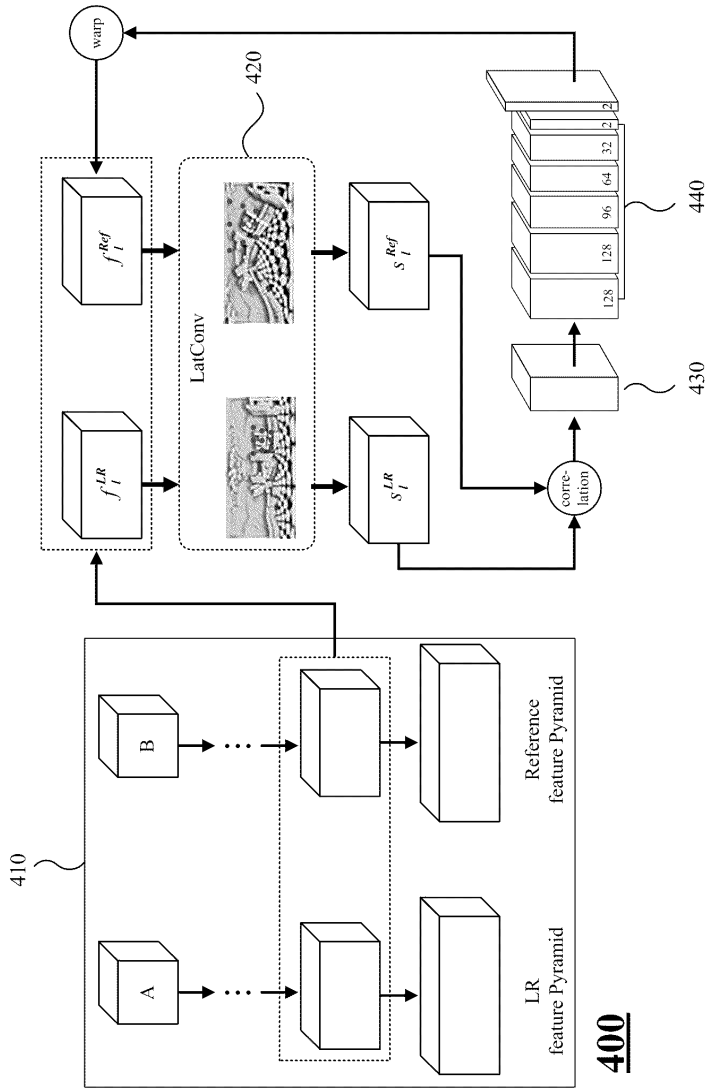
도면2


200
: output feature
E^{SR}
260
250
reconsturction
64
64
64
240
warping
$\hat{E}^{Ref}$
M
sigmoid activation
convolution
occlusion mask generator
32
96
64
32
230
220
225
correlation
flow estimator
$D_{t\to f}$
matching cost
226
222
221
s^{LR}
f^{LR}
s^{Ref}
f^{Ref}
224
223
211
212
t^{LR}
t^{Ref}
E^{LR}
E^{Ref}

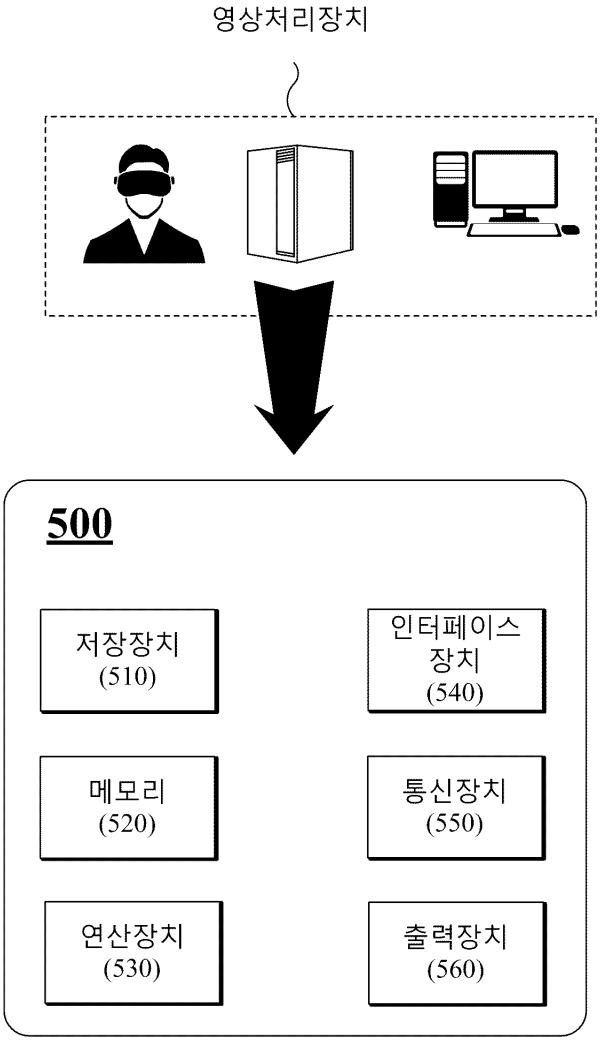
도면3


300
: output feature
340
reconsturction
64
64
64
I^{SR}
sigmoid activation
convolution
330
occlusion mask generator
32
64
96
32
M
310
312
311
S^{LR}
f^{LR}
I^{LR}
315
correlation
warping
320
I^{Ref}
flow estimator
matching cost
$D_{I\to I}$
326
S^{Ref}
f^{Ref}
314
313
I^{Ref}

도면4


warp
420
LatConv
410
B
A
Reference feature Pyramid
LR feature Pyramid
400
corre-lation
430
440
128
128
96
64
32
2
2

도면5


영상처리장치
500
저장장치
(510)
인터페이스
장치
(540)
메모리
(520)
통신장치
(550)
연산장치
(530)
출력장치
(560)