



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0043668
(43) 공개일자 2023년03월31일

- (51) 국제특허분류(Int. Cl.)
G06T 5/50 (2006.01) G06N 3/04 (2023.01)
G06N 3/08 (2023.01) G06T 3/00 (2019.01)
G06T 5/00 (2019.01)
- (52) CPC특허분류
G06T 5/50 (2023.01)
G06N 3/045 (2023.01)
- (21) 출원번호 10-2022-0065357
(22) 출원일자 2022년05월27일
심사청구일자 2022년09월01일
- (30) 우선권주장
1020210126502 2021년09월24일 대한민국(KR)
- (71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
- (72) 발명자
이희경
대전광역시 유성구 가정로 218
엄기문
대전광역시 유성구 가정로 218
(뒷면에 계속)
- (74) 대리인
(유)한양특허법인

전체 청구항 수 : 총 20 항

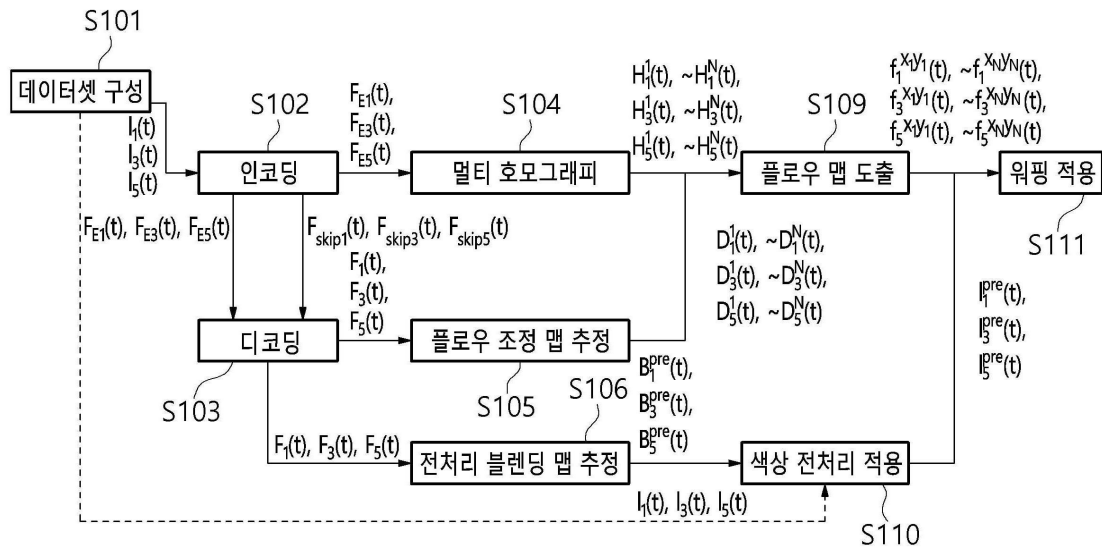
(54) 발명의 명칭 딥러닝 네트워크 기반의 파노라마 영상 생성 방법 및 장치

(57) 요약

본 발명의 일실시예에 따른 파노라마 영상 생성 방법은 복수개의 카메라에서 동시에 얻어진 K개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하는 단계와, 상기 영상 중 K/2 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하는 단계와, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피, 플로우 조정 맵 추정, 전처리 블렌딩 맵 추정, 색상 전처리 적용, 워핑 적용을 수행하는 단계와, 상기 인코딩 특징 맵, 스킵 커넥션 특징 맵, 디코딩 특징 맵, 멀티 호모그래피, 플로우 조정 맵, 전처리 블렌딩 맵, 색상 전처리 적용, 워핑 적용의 결과를 결합하여 파노라마 영상을 생성하는 단계를 포함한다.

(뒷면에 계속)

대표도



래피를 추정하는 단계와, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하는 단계와, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하는 단계와, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하는 단계와, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하는 단계와, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하는 단계와, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 단계와, 상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 $K/2$ 개의 영상을 기초로 손실 함수를 계산하는 단계와, 상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트하는 단계를 포함할 수 있다.

(52) CPC특허분류

조동현

대전광역시 유성구 대학로 99

G06N 3/08 (2023.01)

G06T 3/0068 (2013.01)

G06T 3/0093 (2013.01)

G06T 5/001 (2013.01)

G06T 5/006 (2023.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

(72) 발명자

송대영

대전광역시 유성구 대학로 99

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126210
과제번호	2018-0-00207
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원(IITP)
연구사업명	한국전자통신연구원연구개발지원사업
연구과제명	[전문연구실] 이머시브 미디어 전문연구실
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

복수개의 카메라에서 동시에 얻어진 K 개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하는 단계;

상기 영상 중 $K/2$ 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하는 단계;

상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하는 단계;

상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하는 단계;

상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하는 단계;

상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하는 단계;

상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하는 단계;

상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하는 단계; 및

상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 단계;

를 포함하는 파노라마 영상 생성 방법.

청구항 2

제1항에 있어서,

상기 데이터 셋을 구성하는 단계는,

상기 영상을 캘리브레이션, 비네팅 영역 제거, 영상의 RGB 색 보정 중 적어도 하나를 수행하여 상기 데이터 셋을 구성하는 파노라마 영상 생성 방법.

청구항 3

제1항에 있어서,

상기 인코딩 특징 맵, 상기 스킵 커넥션 특징 맵 및 상기 디코딩 특징 맵을 추출하는 단계는,

상기 영상 중 $K/2$ 개의 영상을 제1 딥러닝 네트워크의 입력으로 하여 상기 인코딩 특징 맵 및 상기 스킵 커넥션 특징 맵을 추출하는 단계; 및

상기 인코딩 특징 맵 및 상기 스킵 커넥션 특징 맵을 제2 딥러닝 네트워크의 입력으로 하여 상기 디코딩 특징 맵을 추출하는 단계;

를 포함하는 파노라마 영상 생성 방법.

청구항 4

제1항에 있어서,

상기 멀티 호모그래피를 추정하는 단계는,

상기 인코딩 특징 맵을 제3 딥러닝 네트워크의 입력으로 하여, 상기 $K/2$ 개의 영상의 방향 당 사용자가 초기 설정한 깊이 구간의 N 개의 호모그래피를 생성하여 멀티 호모그래피를 추정하는 단계;

를 포함하는 파노라마 영상 생성 방법.

청구항 5

제1항에 있어서,

상기 플로우 조정 맵을 추정하는 단계는,

상기 디코딩 특징 맵을 제4 딥러닝 네트워크의 입력으로 하여 플로우 맵을 미세 조정하기 위한 플로우 조정 맵을 추정하는 파노라마 영상 생성 방법.

청구항 6

제1항에 있어서,

상기 전처리 블렌딩 맵을 추정하는 단계는,

상기 디코딩 특징 맵을 제5 딥러닝 네트워크의 입력으로 하여 각 영상의 RGB 채널 별 색상 블렌딩 전처리를 실시하기 위한 전처리 블렌딩 맵을 추정하는 파노라마 영상 생성 방법.

청구항 7

제1항에 있어서,

상기 가중치 맵을 추정하는 단계는,

상기 디코딩 특징 맵을 제6 딥러닝 네트워크의 입력으로 하여 워핑 영상 가중합에 사용될 복수개의 가중치 맵을 추정하는 파노라마 영상 생성 방법.

청구항 8

제1항에 있어서,

상기 후처리 블렌딩 맵을 추정하는 단계는,

상기 디코딩 특징 맵을 제7 딥러닝 네트워크의 입력으로 하여 정합된 파노라마 영상의 RGB 채널 별로 적용할 후처리 블렌딩 맵을 추정하는 파노라마 영상 생성 방법.

청구항 9

제1항에 있어서,

상기 플로우 맵을 도출하는 단계는,

상기 멀티 호모그래피를 제8 딥러닝 네트워크의 입력으로 하여 최초 플로우 맵을 생성하여 최초 플로우 맵을 생성하고, 상기 최초 플로우 맵에 상기 플로우 조정 맵을 더하여 최종 플로우 맵을 도출하는 파노라마 영상 생성 방법.

청구항 10

제1항에 있어서,

상기 전처리된 영상을 생성하는 단계는,

상기 영상 및 상기 전처리 블렌딩 맵을 제9 딥러닝 네트워크의 입력으로 하여 상기 전처리된 영상을 생성하는 파노라마 영상 생성 방법.

청구항 11

제1항에 있어서,

상기 워핑 영상을 생성하는 단계는,

상기 전처리된 영상 및 플로우 맵을 제10 딥러닝 네트워크의 입력으로 하여 상기 전처리된 영상의 각 픽셀을 이동시켜 워핑 영상을 생성하는 파노라마 영상 생성 방법.

청구항 12

제1항에 있어서,

상기 정합 영상을 생성하는 단계는,

상기 복수의 워핑 영상 및 상기 가중치 맵을 제11 딥러닝 네트워크의 입력으로 하여, 상기 복수개의 워핑 영상과 같은 개수를 가지는 가중치 맵 간의 아다마르 곱 연산을 수행한 뒤, 이를 모두 가중합하여 정합 영상을 생성하는 파노라마 영상 생성 방법.

청구항 13

제1항에 있어서,

상기 출력 영상을 생성하는 단계는,

상기 정합 영상 및 상기 후처리 블렌딩 맵을 제12 딥러닝 네트워크의 입력으로 하여, 상기 정합 영상을 상기 후처리 블렌딩 맵으로 보정하여 출력 영상을 생성하는 파노라마 영상 생성 방법.

청구항 14

제1항에 있어서,

상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 $K/2$ 개의 영상을 기초로 손실 함수를 계산하는 단계를 더 포함하는 파노라마 영상 생성 방법.

청구항 15

제14항에 있어서,

상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트하는 단계를 더 포함하는 파노라마 영상 생성 방법.

청구항 16

파노라마 영상 생성을 위한 제어 프로그램이 저장된 메모리; 및

상기 메모리에 저장된 제어 프로그램을 실행하는 프로세서를 포함하고,

상기 프로세서는,

복수개의 카메라에서 동시에 얻어진 K 개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하고, 상기 영상 중 $K/2$ 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하고, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하고, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하고, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하고, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하고, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하고, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하고, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 파노라마 영상 생성 장치.

청구항 17

제16항에 있어서,

상기 프로세서는,

상기 복수의 워핑 영상 및 상기 가중치 맵을 제11 딥러닝 네트워크의 입력으로 하여, 상기 복수개의 워핑 영상과 같은 개수를 가지는 가중치 맵 간의 아다마르 곱 연산을 수행한 뒤, 이를 모두 가중합하여 정합 영상을 생성하는 파노라마 영상 생성 장치.

청구항 18

제16항에 있어서,

상기 프로세서는,

상기 정합 영상 및 상기 후처리 블렌딩 맵을 제12 덤핑 네트워크의 입력으로 하여, 상기 정합 영상을 상기 후처리 블렌딩 맵으로 보정하여 출력 영상을 생성하는 파노라마 영상 생성 장치.

청구항 19

제16항에 있어서,

상기 프로세서는,

상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 $K/2$ 개의 영상을 기초로 손실 함수를 계산하는 파노라마 영상 생성 장치.

청구항 20

제16항에 있어서,

상기 프로세서는,

상기 손실 함수를 기초로 상기 덤핑 네트워크를 업데이트하는 파노라마 영상 생성 장치.

발명의 설명

기술 분야

[0001] 본 발명은 신경망을 이용하여 파노라마 영상을 생성하기 위한 파노라마 영상 생성 방법 및 장치에 관한 것이다.

배경 기술

[0002] 일반적으로, 영상을 정합하기 위해서는 두 대 이상의 카메라에서 얻어진 영상을 이용하여 등장방형도법(Equirectangular projection, ERP) 등의 공간 변환을 수행한 후, 시점선을 기준으로 영상들을 교차하여 정합시키는 방법이 주로 사용되고 있다.

[0003] 그러나, 덤핑에 기반하지 않은 종래의 영상 정합 방법들은 대부분 매 정합마다 특징점 매칭(Feature Matching)과 에너지 최적화(Energy Function Optimization)에 의존한다. 특징점 매칭 방법은 덤핑 기반의 특징맵(Feature Map) 추출 방법에 비하여 강인하지 못하며 특징점 매칭의 실패는 알고리즘의 완전한 실패를 야기한다.

[0004] 또한, 종래의 영상 정합 방법들은 매 영상마다 로컬 호모그래피(Local Homography) 추정, 최적 시점선 추정, 2D 변환 선택을 위한 계산을 위하여 에너지 함수 최적화를 시행하는 것은 알고리즘의 계산 복잡도를 크게 하여 실시간으로 고해상도의 파노라마 영상을 정합시키는 데에 문제점이 있다.

[0005] 더불어, 현존하는 덤핑 기반의 영상 정합 방법들은 신경망 학습에 필요한 정답 영상 생성을 위해 하드웨어 고정 방법, 유사 라벨링(Pseudo Labeling), 가상 영상 시뮬레이터를 이용한 정답 영상 생성 방법, 비지도 학습(Unsupervised Learning) 방법 등을 이용하고 있다.

[0006] 하드웨어 고정 방법은 정답 영상의 취득을 위하여 시점을 고정해 놓고 시차(Parallax)가 없는 정답 영상 및 시차가 존재하는 입력 영상을 취득하는 방법으로, 특수한 하드웨어가 필요하고 하드웨어 동작 방식 때문에 영상의 시간 동기화(Temporal Synchronization)를 보정할 수 없으므로 동적인 물체를 영상에 포함할 수 없다.

[0007] 유사 라벨링 방법은 그 자체로 정답 데이터가 아니고, 기존에 존재하는 방법으로 정합하여 만든 것으로 왜곡이 존재할 수 있으며 신경망 훈련에 한계가 있다.

[0008] 가상 영상 시뮬레이터로 신경망을 학습시키는 방법은 현실 영상에 적용하기 어려운 문제가 있다. 비지도 학습 기반 방법은 학습에 들어가는 데이터를 원래 이미지로부터 잘라내어(Crop) 생성하기 때문에 넓은 시야각을 갖는 파노라마를 얻기 어려운 문제가 있다.

[0009] 따라서, 종래 덤핑 기반의 영상 정합 방법들은 실사 영상만을 이용하여 180~360도의 넓은 시야각을 갖는 파노라마 영상을 생성하지 못하게 된다.

발명의 내용

해결하려는 과제

- [0010] 본 발명의 목적은 실사 영상만을 이용하여 넓은 시야각을 갖는 파노라마 영상을 생성하기 위한 파노라마 영상 생성 방법 및 장치를 제공하는 데 있다.
- [0011] 또한, 본 발명의 목적은 물체의 왜곡에도 강인하며 추론 시에는 정합할 복수의 영상을 입력받아 에너지 함수 최적화와 같은 별도의 추가 계산없이 파노라마 영상을 생성하기 위한 파노라마 영상 생성 방법 및 장치를 제공하는 데 있다.

과제의 해결 수단

- [0012] 상기한 목적을 달성하기 위한 본 발명에 따른 파노라마 영상 생성 방법은 복수개의 카메라에서 동시에 얻어진 K개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하는 단계와, 상기 영상 중 K/2 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하는 단계와, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하는 단계와, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하는 단계와, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하는 단계와, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하는 단계와, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하는 단계와, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하는 단계와, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 단계와, 상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 K/2 개의 영상을 기초로 손실 함수를 계산하는 단계와, 상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트하는 단계를 포함할 수 있다.
- [0013] 상기 데이터 셋을 구성하는 단계는, 상기 영상을 캘리브레이션, 비네팅 영역 제거, 영상의 RGB 색 보정 중 적어도 하나를 수행하여 상기 데이터 셋을 구성할 수 있다.
- [0014] 상기 인코딩 특징 맵, 상기 스킵 커넥션 특징 맵 및 상기 디코딩 특징 맵을 추출하는 단계는, 상기 영상 중 K/2 개의 영상을 제1 딥러닝 네트워크의 입력으로 하여 상기 인코딩 특징 맵 및 상기 스킵 커넥션 특징 맵을 추출하는 단계와, 상기 인코딩 특징 맵 및 상기 스킵 커넥션 특징 맵을 제2 딥러닝 네트워크의 입력으로 하여 상기 디코딩 특징 맵을 추출하는 단계를 포함할 수 있다.
- [0015] 상기 멀티 호모그래피를 추정하는 단계는, 상기 인코딩 특징 맵을 제3 딥러닝 네트워크의 입력으로 하여, 상기 K/2 개의 영상의 방향 당 사용자가 초기 설정한 깊이 구간의 N개의 호모그래피를 생성하여 멀티 호모그래피를 추정할 수 있다.
- [0016] 상기 플로우 조정 맵을 추정하는 단계는, 상기 디코딩 특징 맵을 제4 딥러닝 네트워크의 입력으로 하여 플로우 맵을 미세 조정하기 위한 플로우 조정 맵을 추정할 수 있다.
- [0017] 상기 전처리 블렌딩 맵을 추정하는 단계는, 상기 디코딩 특징 맵을 제5 딥러닝 네트워크의 입력으로 하여 각 영상의 RGB 채널 별 색상 블렌딩 전처리를 실시하기 위한 전처리 블렌딩 맵을 추정할 수 있다.
- [0018] 상기 가중치 맵을 추정하는 단계는, 상기 디코딩 특징 맵을 제6 딥러닝 네트워크의 입력으로 하여 워핑 영상 가중합에 사용될 복수개의 가중치 맵을 추정할 수 있다.
- [0019] 상기 후처리 블렌딩 맵을 추정하는 단계는, 상기 디코딩 특징 맵을 제7 딥러닝 네트워크의 입력으로 하여 정합된 파노라마 영상의 RGB 채널 별로 적용할 후처리 블렌딩 맵을 추정할 수 있다.
- [0020] 상기 플로우 맵을 도출하는 단계는, 상기 멀티 호모그래피를 제8 딥러닝 네트워크의 입력으로 하여 최초 플로우 맵을 생성하여 최초 플로우 맵을 생성하고, 상기 최초 플로우 맵에 상기 플로우 조정 맵을 더하여 최종 플로우 맵을 도출할 수 있다.
- [0021] 상기 전처리된 영상을 생성하는 단계는, 상기 영상 및 상기 전처리 블렌딩 맵을 제9 딥러닝 네트워크의 입력으로 하여 상기 전처리된 영상을 생성할 수 있다.
- [0022] 상기 워핑 영상을 생성하는 단계는, 상기 전처리된 영상 및 플로우 맵을 제10 딥러닝 네트워크의 입력으로 하여 상기 전처리된 영상의 각 픽셀을 이동시켜 워핑 영상을 생성할 수 있다.

- [0023] 상기 정합 영상을 생성하는 단계는, 상기 복수의 워핑 영상 및 상기 가중치 맵을 제11 딥러닝 네트워크의 입력으로 하여, 상기 복수개의 워핑 영상과 같은 개수를 가지는 가중치 맵 간의 아다마르 곱 연산을 수행한 뒤, 이를 모두 가중합하여 정합 영상을 생성할 수 있다.
- [0024] 상기 출력 영상을 생성하는 단계는, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 제12 딥러닝 네트워크의 입력으로 하여, 상기 정합 영상을 상기 후처리 블렌딩 맵으로 보정하여 출력 영상을 생성할 수 있다.
- [0025] 또한, 실시예에 따른 파노라마 영상 생성 장치는 파노라마 영상 생성을 위한 제어 프로그램이 저장된 메모리와, 상기 메모리에 저장된 제어 프로그램을 실행하는 프로세서를 포함하고, 상기 프로세서는, 복수개의 카메라에서 동시에 얻어진 K개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하고, 상기 영상 중 K/2개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하고, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하고, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하고, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하고, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하고, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하고, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하고, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성할 수 있다.
- [0026] 상기 프로세서는, 상기 복수의 워핑 영상 및 상기 가중치 맵을 제11 딥러닝 네트워크의 입력으로 하여, 상기 복수개의 워핑 영상과 같은 개수를 가지는 가중치 맵 간의 아다마르 곱 연산을 수행한 뒤, 이를 모두 가중합하여 정합 영상을 생성할 수 있다.
- [0027] 상기 프로세서는, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 제12 딥러닝 네트워크의 입력으로 하여, 상기 정합 영상을 상기 후처리 블렌딩 맵으로 보정하여 출력 영상을 생성할 수 있다.
- [0028] 상기 프로세서는, 상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 K/2 개의 영상을 기초로 손실 함수를 계산할 수 있다.
- [0029] 상기 프로세서는, 상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트할 수 있다.

발명의 효과

- [0030] 본 발명은 딥러닝 네트워크를 이용하여 특징을 추출함으로써, 영상의 밝기, 대비, 색상 변화, 저화질, 적은 겹치는 영역 등의 악조건에 대하여 영상 정합의 성공률이 향상될 수 있다.
- [0031] 또한, 본 발명은 시접선 부근의 색상정보를 SSIM 손실 함수로 학습하여 블렌딩 맵으로 파노라마 영상을 보정하기 때문에 입력 영상별로 각각 다른 광원에 노출되어도 시접선 부근의 부자연스러운 무늬가 없는 영상 생성이 가능하다.
- [0032] 또한, 본 발명은 멀티 호모그래피, 플로우 미세 조정 맵, 가중합의 효과로 영상의 복잡한 깊이 패턴에 강인하여 대응하여 영상을 정합할 수 있다.
- [0033] 또한, 본 발명은 오직 영상 정합 신경망의 훈련 시에만 손실 함수를 계산하고, 훈련된 신경망의 추론에는 정합할 K/2 개의 영상만을 입력받아 별도의 추가 계산없이 신경망을 통과시켜 파노라마 영상을 획득하여 정합 속도를 향상시킬 수 있다.
- [0034] 또한, 본 발명은 약한 지도학습 시의 단점인 입력과 입력 영상 사이의 시차, 입력과 정답 영상 사이의 시차, 정답과 정답 영상 사이의 시차를 사전 훈련된 VGG16 모델을 이용하여 물체는 최대한 보존하고, 정확한 선(Edge)의 위치 정보 전달은 회피하여 인간의 시각에 더 자연스러운 시차 왜곡을 극복할 수 있다.
- [0035] 또한, 본 발명은 종래 딥러닝 기반 기술과 대비하여 정답 파노라마 영상이 필요하지 않고 각각 시차를 가지는 K개의 영상만을 이용함으로써, 카메라에 가깝거나 동적으로 움직이는 물체도 포함하여 넓은 시야각을 갖는 파노라마 영상을 강인하게 정합할 수 있다.

도면의 간단한 설명

- [0036] 도 1 및 도 2는 본 발명의 일실시예에 따른 파노라마 영상 생성 방법을 나타낸 블록도이다.
- 도 3은 본 발명의 일실시예에 따른 파노라마 영상의 생성을 위해 사용되는 U-Net의 구조를 나타낸 도면이다.

도 4는 본 발명의 일실시예에 따른 파노라마 영상의 생성을 위해 사용되는 VGG16 모델의 구조를 나타낸 도면이다.

도 5는 실시예에 따른 컴퓨터 시스템의 구성을 나타낸 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0037] 본 발명의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 상세하게 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 발명은 이하에서 개시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 것이며, 단지 본 실시예들은 본 발명의 개시가 완전하도록 하며, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이며, 본 발명은 청구항의 범주에 의해 정의될 뿐이다. 명세서 전체에 걸쳐 동일 참조 부호는 동일 구성 요소를 지칭한다.
- [0038] 비록 "제1" 또는 "제2" 등이 다양한 구성요소를 서술하기 위해서 사용되나, 이러한 구성요소는 상기와 같은 용어에 의해 제한되지 않는다. 상기와 같은 용어는 단지 하나의 구성요소를 다른 구성요소와 구별하기 위하여 사용될 수 있다. 따라서, 이하에서 언급되는 제1 구성요소는 본 발명의 기술적 사상 내에서 제2 구성요소일 수도 있다.
- [0039] 본 명세서에서 사용된 용어는 실시예를 설명하기 위한 것이며 본 발명을 제한하고자 하는 것은 아니다. 본 명세서에서, 단수형은 문구에서 특별히 언급하지 않는 한 복수형도 포함한다. 명세서에서 사용되는 "포함한다(comprises)" 또는 "포함하는(comprising)"은 언급된 구성요소 또는 단계가 하나 이상의 다른 구성요소 또는 단계의 존재 또는 추가를 배제하지 않는다는 의미를 내포한다.
- [0040] 다른 정의가 없다면, 본 명세서에서 사용되는 모든 용어는 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에게 공통적으로 이해될 수 있는 의미로 해석될 수 있다. 또한, 일반적으로 사용되는 사전에 정의되어 있는 용어들은 명백하게 특별히 정의되어 있지 않는 한 이상적으로 또는 과도하게 해석되지 않는다.
- [0041] 본 문서에서, "A 또는 B", "A 및 B 중 적어도 하나", "A 또는 B 중 적어도 하나", "A, B 또는 C", "A, B 및 C 중 적어도 하나", 및 "A, B, 또는 C 중 적어도 하나"와 같은 문구들 각각은 그 문구들 중 해당하는 문구에 함께 나열된 항목들 중 어느 하나, 또는 그들의 모든 가능한 조합을 포함할 수 있다.
- [0042] 이하, 첨부된 도면을 참조하여 본 발명의 실시예들을 상세히 설명하기로 하며, 도면을 참조하여 설명할 때 동일하거나 대응하는 구성 요소는 동일한 도면 부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.
- [0043] 본 발명은 약한 지도학습(Weak Supervision)을 기반으로 하는 컨볼루션 신경망(Convolutional Neural Network, CNN) 기반의 영상 정합 방법에 관한 것이다.
- [0044] 본 발명은 시차가 없는 정답 영상을 만들기 어렵기 때문에 정합할 영상을 입력으로 하고 정합할 영상 사이를 촬영한 영상을 정답으로 이용할 수 있다.
- [0045] 이와 같은 정답 데이터 생성 비용을 줄이기 위해 물체를 포함한 다른 영상을 활용하는 학습을 약한 지도학습이라 한다. 본 발명에서 사용되는 모든 K장의 영상 간의 시차(Parallax)가 존재하기 때문에 정합하고자 하는 영상이 K/2장 일 때, 각 입력 영상이 겹치는 영역의 온전한 정보를 담은 K/2장의 정답 영상과 각 정답 영상에 대한 K/2장의 마스크를 활용할 수 있다.
- [0046] 실시예에 따른 파노라마 영상 생성 방법은 파노라마 영상 생성 장치에서 수행될 수 있으며, 일례로 파노라마 영상 생성 장치는 도 5의 컴퓨터 시스템으로 구현될 수 있다.
- [0047] 도 1 및 도 2는 본 발명의 일실시예에 따른 파노라마 영상 생성 방법을 나타낸 블록도이다.
- [0048] 도 1에 도시된 바와 같이, 실시예에 따른 파노라마 영상 생성 장치는 데이터 셋을 구성할 수 있다(S101). 데이터 셋의 구성은 딥러닝 네트워크를 학습시키기 위한 제작 과정을 포함할 수 있다.
- [0049] 실시예에 따른 파노라마 영상 생성 장치는 복수의 카메라를 이용한 영상 촬영, 캘리브레이션(Calibration), 비네팅(vinetting) 영역 제거, 영상 중 학습을 위한 정답 영상의 RGB 색보정 등을 수행할 수 있다.
- [0050] 실시예에 따른 파노라마 영상 생성 장치는 복수개의 카메라로부터 K개의 영상을 동시에 얻을 수 있다.
- [0051] 실시예에 따른 파노라마 영상 생성 장치는 수집된 K개의 영상 중 일정한 간격을 가지는 K/2개(e.g. $I_1(t)$),

$I_3(t)$, $I_5(t)$, if $K=6$)의 영상을 제1 딥러닝 네트워크(U-net 구조 신경망의 인코더)에 입력 영상($I_{1,3,\dots,K/2}(t)$)으로 입력할 수 있다.

[0052] 실시예에 따른 파노라마 영상 생성 장치는 제1 딥러닝 네트워크를 통해 인코딩 특징 맵($F_{E\{1,3,\dots,K/2\}}(t)$) 및 스킵 커넥션 특징 맵($F_{skip\{1,3,\dots,K/2\}}(t)$)을 추출하도록 인코딩을 수행할 수 있다(S102). 여기서, t 는 시간에 대한 변수이다.

[0053] 실시예에 따른 파노라마 영상 생성 장치는 인코딩 특징 맵과 스킵 커넥션 특징 맵을 제2 딥러닝 네트워크의 입력으로 사용할 수 있다.

[0054] 실시예에 따른 파노라마 영상 생성 장치는 제2 딥러닝 네트워크(U-net 구조 신경망의 디코더)를 통해 디코딩 특징 맵($F_{D\{1,3,\dots,K/2\}}(t)$)을 추출하도록 디코딩을 수행할 수 있다(S103).

[0055] 추출된 디코딩 특징 맵은 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵, 후처리 블렌딩 맵들을 도출하기 위한 입력으로 사용될 수 있다.

[0056] 도 3은 본 발명의 일실시예에 따른 파노라마 영상의 생성을 위해 사용되는 U-Net의 구조를 나타낸 도면이다.

[0057] 도 3에 도시된 바와 같이, U-Net 이란 인코더-디코더 구조를 가지면서, 인코더(100) 각 레이어의 출력을 디코더(200)의 각 레이어에 추가적으로 연결하여 인코더(100)의 고 레벨 특징 맵과 입력 영상의 정보를 동시에 활용할 수 있도록 만든 신경망 구조이다.

[0058] 스킵 커넥션(300)은 인코더(100)에 있는 각 레이어의 출력을 디코더(200)의 각 레이어의 입력으로 더해주는 것을 의미할 수 있다.

[0059] 도 1로 돌아가서, 실시예에 따른 파노라마 영상 생성 장치는 인코딩 특징 맵을 제3 딥러닝 네트워크의 입력으로 사용할 수 있다.

[0060] 실시예에 따른 파노라마 영상 생성 장치는 제3 딥러닝 네트워크를 통해 멀티호모그래피를 생성할 수 있다(S104).

[0061] 호모그래피는 영상 정합을 위하여 모든 입력 영상을 2D 변환하기 위한 변환 행렬을 의미할 수 있다. 실사 영상을 포함하는 화소들은 촬영하는 카메라의 렌즈로부터 다양한 깊이(Depth)를 갖기 때문에, 한 개의 글로벌 호모그래피로는 이 깊이 변화에 적절히 대응하여 영상을 정합하기 어렵다. 따라서, 복수의 호모그래피를 추정할 수 있다. 호모그래피의 개수는 사용자의 필요에 의해 변경될 수 있지만 입력 영상 당 3~8개 정도를 사용할 수 있다. 여기서, 영상 당 생성하는 호모그래피 개수를 N 이라 칭할 수 있다.

[0062] 실시예에 따른 파노라마 영상 생성 장치는 인코딩 특징 맵($F_{E\{1,3,\dots,K/2\}}(t)$)을 입력받아, 입력된 영상의 $K/2$ 개의 방향 당 사용자가 초기 설정한 깊이 구간의 개수 N 개의 호모그래피($H_{1,3,\dots,K/2}^{1\sim N}(t)$)를 생성할 수 있다. 따라서, 총 생성되는 멀티 호모그래피의 개수는 $(\frac{K}{2} \times N)$ 개가 될 수 있다.

[0063] 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵을 이용하여 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정할 수 있다.

[0064] 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵($F_{D\{1,3,\dots,K/2\}}(t)$)을 제4 딥러닝 네트워크의 입력으로 사용할 수 있다.

[0065] 실시예에 따른 파노라마 영상 생성 장치는 호모그래피를 이용하여 최초 플로우 맵을 만들었을 때, 플로우 맵을 미세 조정하기 위한 플로우 조정 맵($D_{1,3,\dots,K/2}^{1\sim N}(t)$)을 추정할 수 있다(S105).

[0066] N 개의 호모그래피만으로 2D 변환을 할 경우 영상의 연속적인 깊이를 표현할 수 없기 때문에 최초 플로우 맵을

플로우 조정 맵($D_{1,3,\dots,\frac{K}{2}}^{1\sim N}(t)$)으로 가감하여 연속적으로 변하는 값에 맞춰 워핑될 수 있도록 만들 수 있다.

- [0067] 여기서, 워핑은 어떤 플로우 정보에 따라 입력된 영상의 화소가 이동하는 작용을 의미한다. 또한, 플로우 또는 플로우 맵은 워핑을 위해 입력된 영상의 원래 화소의 위치로부터 상하좌우로 얼마나 이동시켜야 하는 지에 대한 정보를 맵(Tensor) 형태로 표현한 것을 의미한다.
- [0068] 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵을 제5 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0069] 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵을 입력받아 워핑을 실시하기 전 각 입력된 영상($I_{1,3,\dots,\frac{K}{2}}(t)$)의 RGB 채널 별 색상 블렌딩 전처리를 실시하기 위한 전처리 블렌딩 맵($B_{1,3,\dots,\frac{K}{2}}^{pre}(t)$)을 추정할 수 있다(S106).
- [0070] 도 2에 도시된 바와 같이, 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵을 제6 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0071] 실시예에 따른 파노라마 영상 생성 장치는 제5 딥러닝 네트워크를 통해 워핑 영상 가중합에 사용될 ($\frac{K}{2} \times N$) 개의 가중치 맵($W_{1,3,\dots,\frac{K}{2}}^{1\sim N}(t)$)을 추정할 수 있다(S107).
- [0072] 실시예에 따른 파노라마 영상 생성 장치는 디코딩 특징 맵을 제7 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0073] 실시예에 따른 파노라마 영상 생성 장치는 제7 딥러닝 네트워크를 통해 정합된 파노라마 영상의 RGB 채널 별로 적용할 후처리 블렌딩 맵($B^{post}(t)$)을 추정할 수 있다(S108).
- [0074] 도 1로 돌아가서, 실시예에 따른 파노라마 영상 생성 장치는 멀티 호모그래피 및 플로우 조정 맵을 제8 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0075] 실시예에 따른 파노라마 영상 생성 장치는 ($\frac{K}{2} \times N$) 개의 멀티 호모그래피($H_{1,3,\dots,\frac{K}{2}}^{1\sim N}(t)$)로 최초 플로우 맵을 생성하고, 이를 플로우 조정 맵($D_{1,3,\dots,\frac{K}{2}}^{1\sim N}(t)$)과 더하여 플로우를 미세 조정할 수 있다. 상기 과정을 통해 최종 플로우 맵($f_{1,3,\dots,\frac{K}{2}}^{x_1\sim N,y_1\sim N}(t)$)이 도출될 수 있다(S109).
- [0076] 실시예에 따른 파노라마 영상 생성 장치는 전처리 블렌딩 맵을 제9 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0077] 실시예에 따른 파노라마 영상 생성 장치는 전처리 블렌딩 맵($B_{1,3,\dots,\frac{K}{2}}^{pre}(t)$)을 이용하여 입력된 영상($I_{1,3,\dots,\frac{K}{2}}(t)$)을 보정할 수 있다. 상기와 같이 색상 전처리를 적용하여 전처리된 영상($I_{1,3,\dots,\frac{K}{2}}^{pre}(t)$)을 생성할 수 있다(S110).
- [0078] 실시예에 따른 파노라마 영상 생성 장치는 전처리된 영상 및 플로우 맵을 제10 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0079] 실시예에 따른 파노라마 영상 생성 장치는 전처리된 영상($I_{1,3,\dots,\frac{K}{2}}^{pre}(t)$)의 각 픽셀을 최종 도출된 플로우 맵($f_{1,3,\dots,\frac{K}{2}}^{x_1\sim N,y_1\sim N}(t)$)을 이용하여 이동시킴으로써, ($\frac{K}{2} \times N$) 개의 워핑 영상($I_{1,3,\dots,\frac{K}{2}}^{W_{1\sim N}}(t)$)을 생성할 수 있다(S111).
- [0080] 도 2에 도시된 바와 같이, 실시예에 따른 파노라마 영상 생성 장치는 워핑 영상 및 가중치 맵을 제11 딥러닝 네트워크의 입력으로 사용할 수 있다.
- [0081] 실시예에 따른 파노라마 영상 생성 장치는 ($\frac{K}{2} \times N$) 개의 워핑 영상($I_{1,3,\dots,\frac{K}{2}}^{W_{1\sim N}}(t)$)과 같은 개수를 가지는 가중치 맵($W_{1,3,\dots,\frac{K}{2}}^{1\sim N}(t)$)간 아다마르 곱(Hadamard Product) 연산을 수행하고, 수행된 결과를 모두 더하도록 가중합을 계산하여

넓은 시야각을 가지는 정합 영상($I_{PANO}(t)$)을 생성할 수 있다(S112).

[0082] 실시예에 따른 파노라마 영상 생성 장치는 정합 영상 및 후처리 블렌딩 맵을 제12 딥러닝 네트워크의 입력으로 사용할 수 있다.

[0083] 실시예에 따른 파노라마 영상 생성 장치는 정합 영상($I_{PANO}(t)$)을 후처리 블렌딩 맵($B^{post}(t)$)으로 보정할 수 있다. 상기와 같이 색상 후처리를 적용하여 최종 출력 영상($I_{PANO}^{post}(t)$)을 생성할 수 있다(S113).

[0084] 실시예에 따른 파노라마 영상 생성 장치는 출력 영상($I_{PANO}^{post}(t)$) 및 정답으로 사용될 영상($I_{2,4,...K}(t)$)을 기초로 약한 지도학습을 실시할 수 있다. 여기서, 정답으로 사용될 영상은 입력된 영상 중 인코딩에 사용되지 않은 영상일 수 있다.

[0085] 실시예에 따른 파노라마 영상 생성 장치는 출력 영상($I_{PANO}^{post}(t)$)을 정답으로 사용될 영상($I_{2,4,...K}(t)$)의 모양에 정확하게 맞는 마스크를 이용하여 해당 부분만 각각 잘라낼 수 있다.

[0086] 각 마스크로 자른 모델의 출력을 $I_{2,4,...K}^{post}(t)$ 라 할 때, 전처리 블렌딩 맵과 후처리 블렌딩 맵의 생성을 위한 학습을 위하여 출력($I_{2,4,...K}^{post}(t)$) 및 정답으로 사용될 영상($I_{2,4,...K}(t)$)에서 각각 겹치지 않는 부분(Non-overlapping region)만을 자를 수 있다. 이로 인해 잘라진 출력($I_{2,4,...K}^{post-non}(t)$) 및 잘라진 영상($I_{2,4,...K}^{non}(t)$)을 생성할 수 있다.

[0087] 상기와 같이 생성된 잘라진 출력($I_{2,4,...K}^{post-non}(t)$) 및 잘라진 영상($I_{2,4,...K}^{non}(t)$)을 SSIM(Structural similarity index measure) 손실 함수($L_{SSIM}(t)$)로 연산할 수 있다.

[0088] SSIM 손실 함수는 수학적 식 1과 같다.

[0089] [수학적 식 1]

$$L_{SSIM}(t) = SSIM(I_{2,4,...K}^{non}(t), I_{2,4,...K}^{post-non}(t))$$

[0091] 또한, 영상의 시차를 극복하고 자연스러운 정합을 학습시키기 위해 물체 분류(Object Classification) 용도로 사전 훈련된 VGG16 모델을 사용할 수 있다.

[0092] 도 4는 본 발명의 일실시예에 따른 파노라마 영상의 생성을 위해 사용되는 VGG16 모델의 구조를 나타낸 도면이다.

[0093] 도 4에 도시된 바와 같이, VGG16 모델의 백본(backbone)은 컨볼루션 레이어(400) 및 maxpool(500)로 이루어지며, max-pooling 계층을 기준으로 5단계로 나누어져 있으며, 백본에 출력($I_{2,4,...K}^{post}(t)$) 및 정답으로 사용될 영상($I_{2,4,...K}(t)$)을 통과시켜 max-pooling 계층 이후에서 특징 맵 $F_{(2,4,...K)P}^{VGG16-post}(t) = VGG16(I_{2,4,...K}^{post}(t))$ 과 $F_{(2,4,...K)P}^{VGG16}(t) = VGG16(I_{2,4,...K}(t))$ 을 얻을 수 있다.

[0094] 여기서 P는 백본의 단계를 의미하며, P=(3~5) 단계에서 High-level 특징 맵만을 사용할 수 있다. 이때, 지각 손실 함수($PL_1(t)_P$)는 수학적 식 2와 같다.

[0095] [수학적 식 2]

$$PL_1(t)_P = |F_{(2,4,...K)P}^{VGG16}(t) - F_{(2,4,...K)P}^{VGG16-post}(t)|, (P = 3 \sim 5)$$

[0097] 지각 손실 함수의 합인 $PL_1(t) = 0.4PL_1(t)_3 + 0.6PL_1(t)_4 + PL_1(t)_5$ 과 $L_{SSIM}(t)$ 를 합산하여 손실 함수의 텐서인 $L(t) = L_{SSIM}(t) + PL_1(t)$ 를 구할 수 있다(S114).

[0098] 실시예에 따른 파노라마 영상 생성 장치는 텐서($L(t)$)를 딥러닝 네트워크에 역전파시켜 딥러닝 네트워크를 업데이트

이트할 수 있다(S115).

- [0099] 실시예에 따른 파노라마 영상 생성 장치는 컴퓨터 판독 가능한 기록매체와 같은 컴퓨터 시스템에서 구현될 수 있다.
- [0100] 도 5는 실시예에 따른 컴퓨터 시스템의 구성을 나타낸 블록도이다.
- [0101] 도 5를 참조하면, 실시예에 따른 컴퓨터 시스템(1000)은 버스(1020)를 통하여 서로 통신하는 하나 이상의 프로세서(1010), 메모리(1030), 사용자 인터페이스 입력 장치(1040), 사용자 인터페이스 출력 장치(1050) 및 스토리지(1060)를 포함할 수 있다. 또한, 컴퓨터 시스템(1000)은 네트워크에 연결되는 네트워크 인터페이스(1070)를 더 포함할 수 있다.
- [0102] 프로세서(1010)는 중앙 처리 장치 또는 메모리나 스토리지에 저장된 프로그램 또는 프로세싱 인스트럭션들을 실행하는 반도체 장치일 수 있다. 프로세서(1010)는 일종의 중앙처리장치로서 파노라마 영상 생성 장치의 전체 동작을 제어할 수 있다.
- [0103] 프로세서(1010)는 데이터를 처리할 수 있는 모든 종류의 장치를 포함할 수 있다. 여기서, '프로세서(processor)'는 예를 들어 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(microporcessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array) 등의 처리 장치를 망라할 수 있으나, 이에 한정되는 것은 아니다.
- [0104] 메모리(1030)는 실시예에 따른 파노라마 영상 생성 방법을 수행하기 위한 제어 프로그램 등 전반적인 동작을 위한 다양한 데이터가 저장될 수 있다. 구체적으로, 메모리(1030)에는 파노라마 영상 생성 장치에서 구동되는 다수의 응용 프로그램, 파노라마 영상 생성 장치의 동작을 위한 데이터 및 명령어가 저장될 수 있다.
- [0105] 메모리(1030) 및 스토리지(1060)는 휘발성 매체, 비휘발성 매체, 분리형 매체, 비분리형 매체, 통신 매체, 또는 정보 전달 매체 중에서 적어도 하나 이상을 포함하는 저장 매체일 수 있다. 예를 들어, 메모리(1030)는 ROM(1031)이나 RAM(1032)을 포함할 수 있다.
- [0106] 일 실시예에 따르면, 컴퓨터 프로그램을 저장하고 있는 컴퓨터 판독 가능한 기록 매체로서, 복수개의 카메라에서 동시에 얻어진 K개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하는 단계와, 상기 영상 중 K/2 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하는 단계와, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하는 단계와, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하는 단계와, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하는 단계와, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하는 단계와, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하는 단계와, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하는 단계와, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 단계와, 상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 K/2 개의 영상을 기초로 손실 함수를 계산하는 단계와, 상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트하는 단계의 동작을 포함하는 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함할 수 있다.
- [0107] 일 실시예에 따르면 컴퓨터 판독 가능한 기록매체에 저장되어 있는 컴퓨터 프로그램으로서, 복수개의 카메라에서 동시에 얻어진 K개의 영상을 기초로 딥러닝 네트워크의 학습을 위한 데이터 셋을 구성하는 단계와, 상기 영상 중 K/2 개의 영상을 기초로 인코딩 특징 맵, 스킵 커넥션 특징 맵 및 디코딩 특징 맵을 추출하는 단계와, 상기 인코딩 특징 맵을 기초로 멀티 호모그래피를 추정하는 단계와, 상기 디코딩 특징 맵을 기초로 플로우 조정 맵, 전처리 블렌딩 맵, 가중치 맵 및 후처리 블렌딩 맵을 추정하는 단계와, 상기 멀티 호모그래피 및 상기 플로우 조정 맵을 기초로 플로우 맵을 도출하는 단계와, 상기 전처리 블렌딩 맵을 기초로 전처리된 영상을 생성하는 단계와, 상기 전처리된 영상 및 상기 플로우 맵을 기초로 복수개의 워핑 영상을 생성하는 단계와, 상기 복수개의 워핑 영상 및 가중치 맵을 기초로 정합 영상을 생성하는 단계와, 상기 정합 영상 및 상기 후처리 블렌딩 맵을 기초로 상기 정합 영상을 보정하여 출력 영상을 생성하는 단계와, 상기 출력 영상 및 상기 데이터 셋 중 정답으로 사용될 나머지 K/2 개의 영상을 기초로 손실 함수를 계산하는 단계와, 상기 손실 함수를 기초로 상기 딥러닝 네트워크를 업데이트하는 단계를 포함하는 방법을 프로세서가 수행하도록 하기 위한 명령어를 포함할 수

있다.

[0108] 본 발명에서 설명하는 특정 실행들은 실시예들로서, 어떠한 방법으로도 본 발명의 범위를 한정하는 것은 아니다. 명세서의 간결함을 위하여, 종래 전자적인 구성들, 제어시스템들, 소프트웨어, 상기 시스템들의 다른 기능적인 측면들의 기재는 생략될 수 있다. 또한, 도면에 도시된 구성 요소들 간의 선들의 연결 또는 연결 부재들은 기능적인 연결 및/또는 물리적 또는 회로적 연결들을 예시적으로 나타낸 것으로서, 실제 장치에서는 대체 가능하거나 추가의 다양한 기능적인 연결, 물리적인 연결, 또는 회로 연결들로서 나타내어질 수 있다. 또한, "필수적인", "중요하게" 등과 같은 구체적인 언급이 없다면 본 발명의 적용을 위하여 반드시 필요한 구성 요소가 아닐 수 있다.

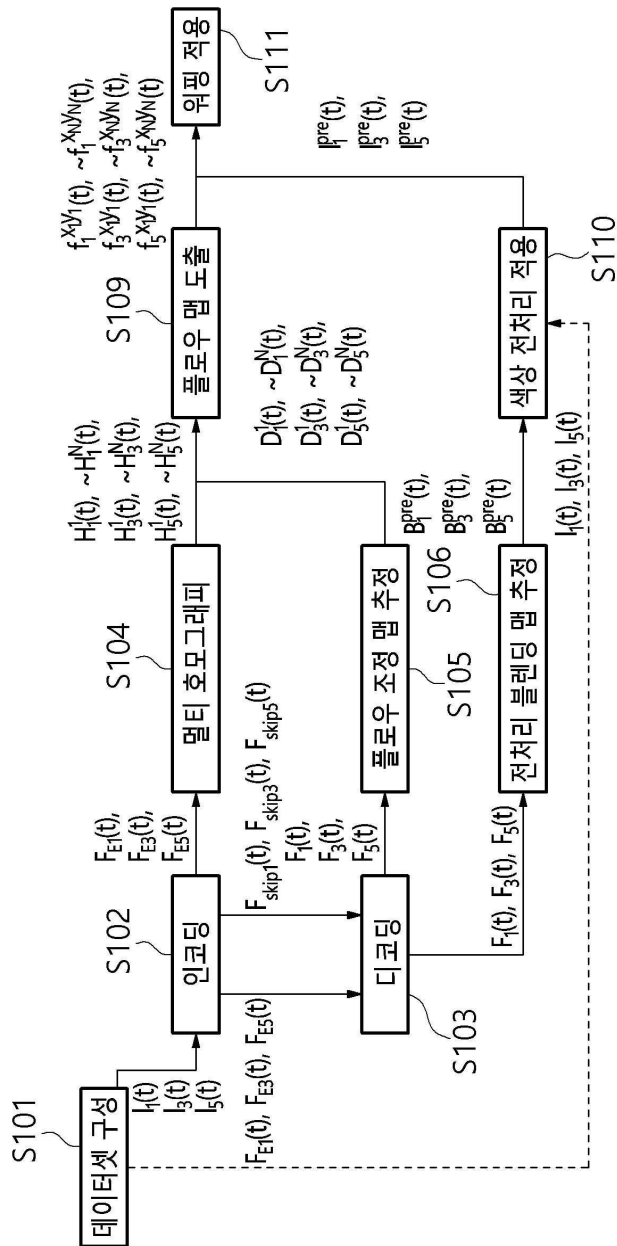
[0109] 따라서, 본 발명의 사상은 상기 설명된 실시예에 국한되어 정해져서는 아니되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 또는 이로부터 등가적으로 변경된 모든 범위는 본 발명의 사상의 범주에 속한다고 할 것이다.

부호의 설명

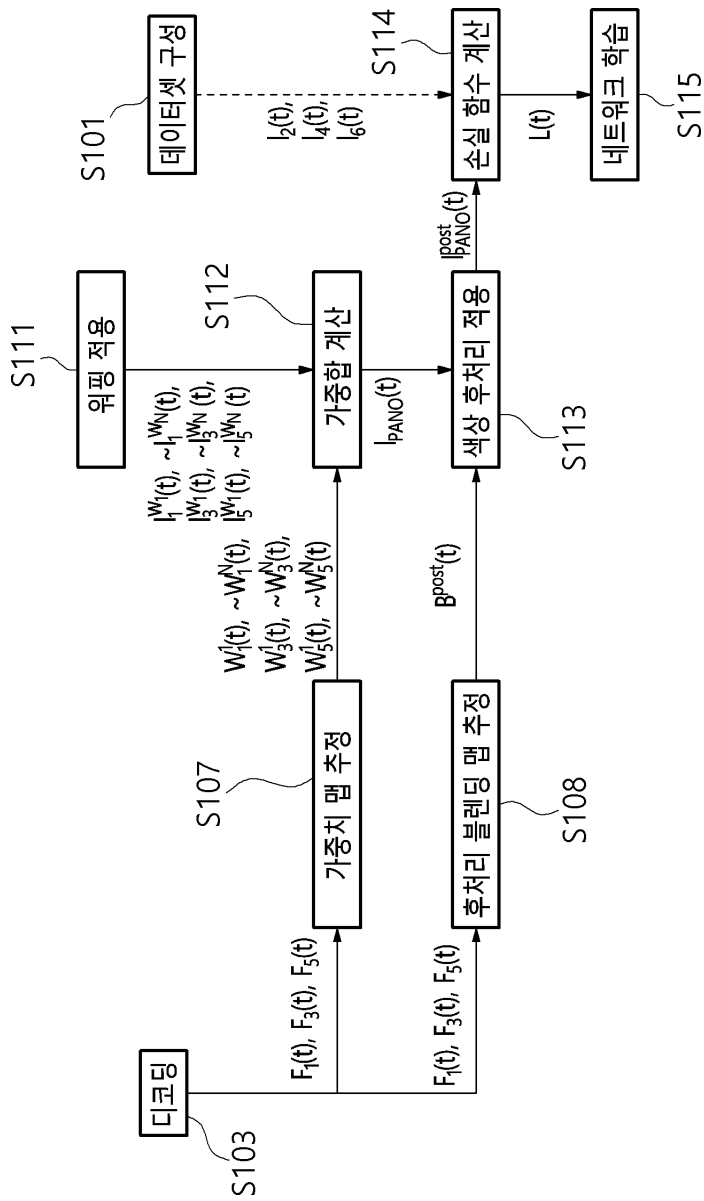
[0110] 1000: 컴퓨터 시스템
1010: 프로세서
1030: 메모리

도면

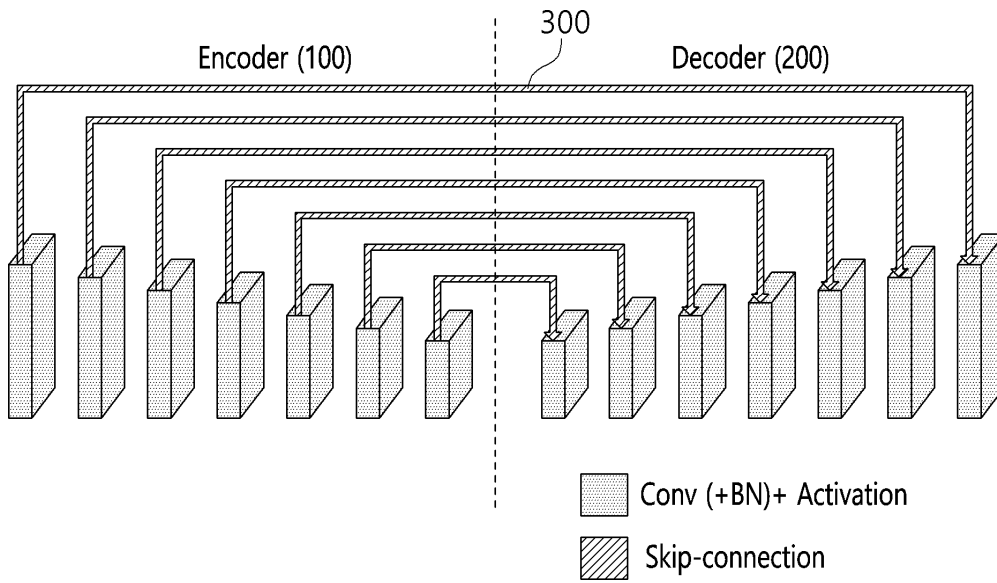
도면1



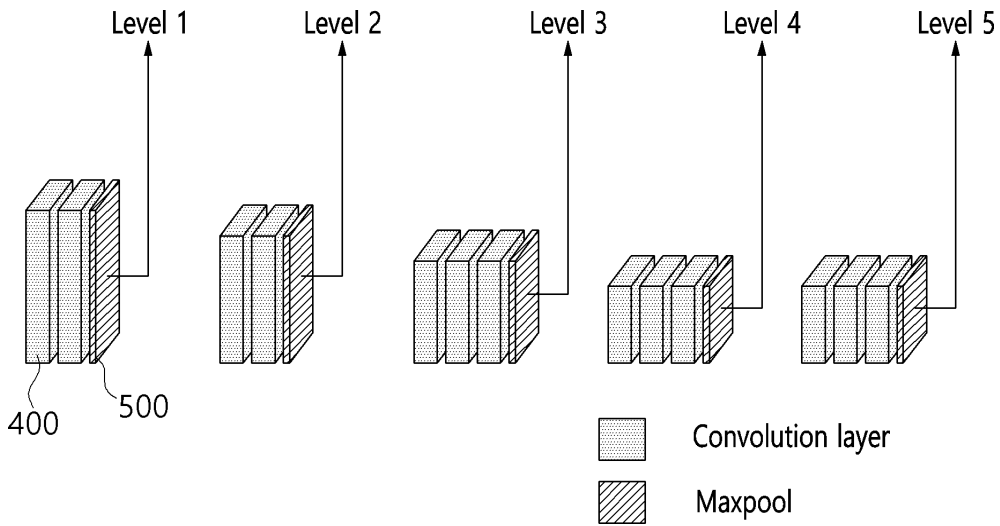
도면2



도면3



도면4



도면5

