



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0072723
(43) 공개일자 2022년06월02일

(51) 국제특허분류(Int. Cl.)
H04N 13/128 (2018.01) G06N 3/08 (2006.01)
G06T 3/00 (2019.01) G06T 5/50 (2006.01)
H04N 13/156 (2018.01)
(52) CPC특허분류
H04N 13/128 (2021.08)
G06N 3/08 (2013.01)
(21) 출원번호 10-2021-0105961
(22) 출원일자 2021년08월11일
심사청구일자 2021년08월11일
(30) 우선권주장
1020200160113 2020년11월25일 대한민국(KR)

(71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
충남대학교산학협력단
대전광역시 유성구 대학로 99 (궁동, 충남대학교)
(72) 발명자
엄기문
대전광역시 유성구 가정로 218 (가정동)
이희경
대전광역시 유성구 가정로 218 (가정동)
(74) 대리인
특허법인 무한

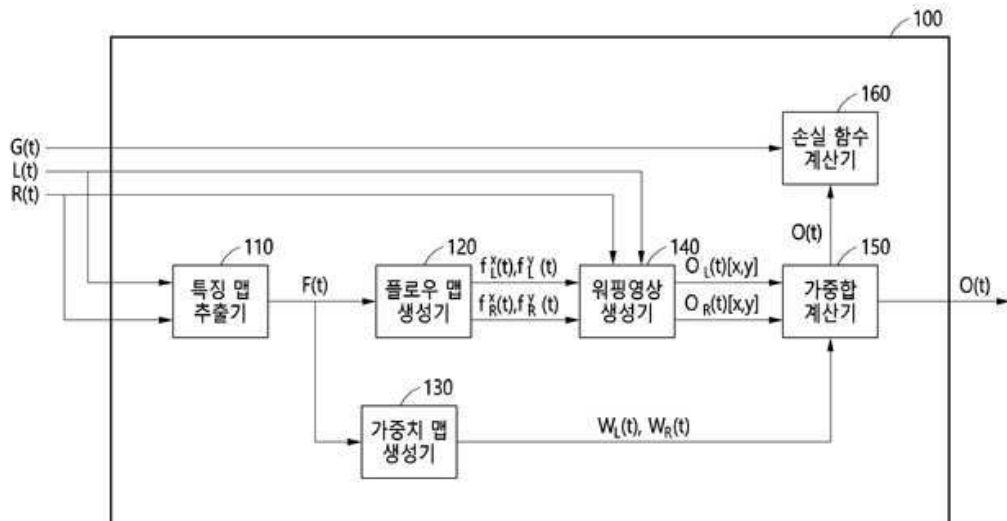
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 딥 러닝 기반 영상 정합 방법 및 장치

(57) 요약

딥 러닝 기반 영상 정합 방법 및 장치가 개시된다. 일 실시예에 따른 딥 러닝 기반 영상 방법은, 제1 영상 및 상기 제1 영상과 정합되는 제2 영상을 포함하는 복수의 영상을 수신하는 동작; 상기 복수의 영상의 특징 맵을 추출하는 동작; 상기 특징 맵에 기초하여 상기 복수의 영상의 픽셀들의 이동 정보인 복수의 플로우 맵을 생성하는 동작; 상기 복수의 영상에 상기 복수의 플로우 맵을 각각 적용하여 상기 복수의 영상의 픽셀들이 이동한 복수의 워핑 영상을 생성하는 동작; 상기 특징 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩하기 위한 복수의 가중치 맵을 생성하는 동작; 및 상기 복수의 가중치 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩함으로써 상기 제1 영상과 상기 제2 영상이 정합된 정합 영상을 생성하는 동작을 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06T 3/0093 (2013.01)

G06T 5/50 (2013.01)

H04N 13/156 (2018.05)

G06T 2207/20084 (2013.01)

(72) 발명자

송대영

대전광역시 유성구 대학로 99, 공학2호관(E2) 472
호 (궁동, 충남대학교)

조동현

대전광역시 유성구 대학로 99, 공학2호관(E2) 407
호 (궁동, 충남대학교)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711102773
과제번호	2018-0-00207
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	ETRI 연구개발지원사업
연구과제명	[전문연구실]이머시브 미디어 전문연구실
기 여 율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2020.01.01 ~ 2020.12.31

명세서

청구범위

청구항 1

제1 영상 및 상기 제1 영상과 정합되는 제2 영상을 포함하는 복수의 영상을 수신하는 동작;

상기 복수의 영상의 특징 맵을 추출하는 동작;

상기 특징 맵에 기초하여 상기 복수의 영상의 픽셀들의 이동 정보인 복수의 플로우 맵을 생성하는 동작;

상기 복수의 영상에 상기 복수의 플로우 맵을 각각 적용하여 상기 복수의 영상의 픽셀들이 이동한 복수의 워핑 영상을 생성하는 동작;

상기 특징 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩하기 위한 복수의 가중치 맵을 생성하는 동작; 및

상기 복수의 가중치 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩함으로써 상기 제1 영상과 상기 제2 영상이 정합된 정합 영상을 생성하는 동작

을 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 2

제1항에 있어서,

상기 복수의 플로우 맵을 생성하는 동작은,

상기 제1 영상의 픽셀들의 이동 정보인 제1 플로우 맵을 생성하는 동작; 및

상기 제2 영상의 픽셀들의 이동 정보인 제2 플로우 맵을 생성하는 동작

을 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 3

제2항에 있어서,

상기 복수의 워핑 영상을 생성하는 동작은,

상기 제1 영상에 상기 제1 플로우 맵을 적용하여 제1 워핑 영상을 생성하는 동작; 및

상기 제2 영상에 상기 제2 플로우 맵을 적용하여 제2 워핑 영상을 생성하는 동작

을 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 4

제3항에 있어서,

상기 복수의 가중치 맵을 생성하는 동작은,

상기 제1 워핑 영상에 대한 제1 가중치 맵을 생성하는 동작; 및

상기 제2 워핑 영상에 대한 제2 가중치 맵을 생성하는 동작

을 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 5

제4항에 있어서,
 상기 정합 영상을 생성하는 동작은,
 상기 제1 워핑 영상에 상기 제1 가중치 맵을 적용하는 동작;
 상기 제2 워핑 영상에 상기 제2 가중치 맵을 적용하는 동작; 및
 상기 제1 가중치가 적용된 제1 워핑 영상과 상기 제2 가중치가 적용된 제2 워핑 영상을 블렌딩하여 상기 정합 영상을 생성하는 동작
 을 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 6

제1항에 있어서,
 제1 손실함수와 제2 손실함수가 가중합된 제3 손실 함수를 계산하는 동작; 및
 상기 제3 손실 함수에 기초하여 학습하는 동작
 을 더 포함하는, 딥 러닝 기반 영상 정합 방법.

청구항 7

제6항에 있어서,
 상기 제1 손실함수는,
 상기 정합 영상의 픽셀 값과 상기 정답 영상의 픽셀 값의 차이를 계산하는 L1 손실 함수이고,
 상기 제2 손실 함수는,
 상기 정합 영상의 손실 정보를 평가하는 SSIM(Structural similarity index measure) 손실 함수인,
 딥 러닝 기반 영상 정합 방법.

청구항 8

하드웨어와 결합되어 제1항 내지 제7항 중 어느 하나의 항의 방법을 실행시키기 위하여 컴퓨터 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

청구항 9

하나 이상의 인스트럭션을 저장하는 메모리; 및
 상기 인스트럭션을 실행시키기 위한 프로세서
 를 포함하고,
 상기 인스트럭션이 실행될 때, 상기 프로세서는,
 제1 영상 및 상기 제1 영상과 정합되는 제2 영상을 포함하는 복수의 영상을 수신하고,
 상기 복수의 영상의 특징 맵을 추출하고,
 상기 특징 맵에 기초하여 상기 복수의 영상의 픽셀들의 이동 정보인 복수의 플로우 맵을 생성하고,

상기 복수의 영상에 상기 복수의 플로우 맵을 각각 적용하여 상기 복수의 영상의 픽셀들이 이동한 복수의 워핑 영상을 생성하고,

상기 특징 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩하기 위한 복수의 가중치 맵을 생성하고,

상기 복수의 가중치 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩함으로써 상기 제1 영상과 상기 제2 영상이 정합된 정합 영상을 생성하는,

딥 러닝 기반 영상 정합 장치.

청구항 10

제9항에 있어서,

상기 프로세서는,

상기 제1 영상의 픽셀들의 이동 정보인 제1 플로우 맵을 생성하고,

상기 제2 영상의 픽셀들의 이동 정보인 제2 플로우 맵을 생성하는,

딥 러닝 기반 영상 정합 장치.

청구항 11

제10항에 있어서,

상기 프로세서는,

상기 제1 영상에 상기 제1 플로우 맵을 적용하여 제1 워핑 영상을 생성하고,

상기 제2 영상에 상기 제2 플로우 맵을 적용하여 제2 워핑 영상을 생성하는,

딥러닝 기반 영상 정합 장치.

청구항 12

제11항에 있어서,

상기 프로세서는,

상기 제1 워핑 영상에 대한 제1 가중치 맵을 생성하고,

상기 제2 워핑 영상에 대한 제2 가중치 맵을 생성하는,

딥러닝 기반 영상 정합 장치.

청구항 13

제12항에 있어서,

상기 프로세서는,

상기 제1 워핑 영상에 상기 제1 가중치 맵을 적용하고,

상기 제2 워핑 영상에 상기 제2 가중치 맵을 적용하고,

상기 제1 가중치가 적용된 제1 워핑 영상과 상기 제2 가중치가 적용된 제2 워핑 영상을 블렌딩하여 상기 정합 영상을 생성하는,

딥 러닝 기반 영상 정합 장치.

청구항 14

제9항에 있어서,
상기 프로세서는,
제1 손실함수와 제2 손실함수가 가중합된 제3 손실 함수를 계산하고,
상기 제3 손실 함수에 기초하여 학습하는,
딥 러닝 기반 영상 정합 장치.

청구항 15

제14항에 있어서,
상기 제1 손실함수는,
상기 정합 영상의 픽셀 값과 상기 정답 영상의 픽셀 값의 차이를 계산하는 L1 손실 함수이고,
상기 제2 손실 함수는,
상기 정합 영상의 손실 정보를 평가하는 SSIM(Structural similarity index measure) 손실 함수인,
딥 러닝 기반 영상 정합 장치.

발명의 설명

기술 분야

[0001] 아래 개시는 딥 러닝 기반 영상 정합 방법 및 장치에 관한 것이다.

배경 기술

[0003] 하나의 장면이나 대상을 다른 시간이나 관점에서 촬영할 경우, 영상은 서로 다른 좌표계에서 획득된다. 영상 정합(image stitching) 기술은 겹치는 부분이 존재하는 두 장 이상의 영상을 자연스럽게 연결하여 하나의 영상을 생성하는 기술이다.

[0004] 영상 정합 기술은 한 가지 방식의 영상들을 맞추는 단일 방식 정합 및 여러 방식의 영상을 서로 맞추는 여러 방식 정합으로 구분될 수 있다. 단일 방식 정합은 한 가지 스캐너 및/또는 측정 장치에서 얻어진 영상만을 정합하는 것이고, 여러 영상 방식은 서로 다른 스캐너 및/또는 측정 장치에서 얻어진 영상들을 정합하는 것이다.

[0005] 영상 정합 기술은 특징점 추출, 중첩 영역에서의 특징점 매칭, 호모그래피(homography) 추정, 워핑(warping) 및 블렌딩(blending) 등의 기술을 이용하여 구현된다.

발명의 내용

해결하려는 과제

[0007] 영상을 복수의 시점에서 촬영하는 경우, 물체와 촬영 시점 간의 거리가 촬영 시점 마다 서로 다를 수 있어 거리 차에 의한 시차 왜곡은 빈번하게 발생할 수 있다. 이에, 거리차에 의한 왜곡 보정을 위한 기술이 필요할 수 있다.

[0008] 다양한 실시예들은 딥러닝 네트워크에 기반하여 시차(예: 시차 왜곡)에 강인한 영상 정합 기술을 제공할 수 있다.

[0009] 다만, 기술적 과제는 상술한 기술적 과제들로 한정되는 것은 아니며, 또 다른 기술적 과제들이 존재할 수 있다.

과제의 해결 수단

[0011] 일 실시예에 따른 영상 정합 방법은, 제1 영상 및 상기 제1 영상과 정합(stitch)되는 제2 영상을 포함하는 복수의 영상을 수신하는 동작과 상기 복수의 영상의 특징 맵을 추출하는 동작과 상기 특징 맵에 기초하여 상기 복수의 영상의 픽셀들의 이동 정보인 복수의 플로우 맵을 생성하는 동작과 상기 복수의 영상에 상기 복수의 플로우 맵을 각각 적용하여 상기 복수의 영상의 픽셀들이 이동한 복수의 워핑 영상을 생성하는 동작과 상기 특징 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩하기 위한 복수의 가중치 맵을 생성하는 동작과 상기 복수의 가중치 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩함으로써 상기 제1 영상과 상기 제2 영상이 정합된 정합 영상을 생성하는 동작을 포함할 수 있다.

[0012] 상기 복수의 플로우 맵을 생성하는 동작은, 상기 제1 영상의 픽셀들의 이동 정보인 제1 플로우 맵을 생성하는 동작과 상기 제2 영상의 픽셀들의 이동 정보인 제2 플로우 맵을 생성하는 동작을 포함할 수 있다.

[0013] 상기 복수의 워핑 영상을 생성하는 동작은, 상기 제1 영상에 상기 제1 플로우 맵을 적용하여 제1 워핑 영상을 생성하는 동작과 상기 제2 영상에 상기 제2 플로우 맵을 적용하여 제2 워핑 영상을 생성하는 동작을 포함할 수 있다.

[0014] 상기 복수의 가중치 맵을 생성하는 동작은, 상기 제1 워핑 영상에 대한 제1 가중치 맵을 생성하는 동작과 상기 제2 워핑 영상에 대한 제2 가중치 맵을 생성하는 동작을 포함할 수 있다.

[0015] 상기 정합 영상을 생성하는 동작은, 상기 제1 워핑 영상에 상기 제1 가중치 맵을 적용하는 동작과 상기 제2 워핑 영상에 상기 제2 가중치 맵을 적용하는 동작과 상기 제1 가중치가 적용된 제1 워핑 영상과 상기 제2 가중치가 적용된 제2 워핑 영상을 블렌딩하여 상기 정합 영상을 생성하는 동작을 포함할 수 있다.

[0016] 영상 정합 방법은, 제1 손실함수와 제2 손실함수가 가중합된 제3 손실 함수를 계산하는 동작과 상기 제3 손실 함수에 기초하여 학습하는 동작을 더 포함할 수 있다.

[0017] 상기 제1 손실함수는, 상기 정합 영상의 픽셀 값과 상기 정답 영상의 픽셀 값의 차이를 계산하는 L1 손실 함수이고, 상기 제2 손실 함수는, 상기 정합 영상의 손실 정보를 평가하는 SSIM(Structural similarity index measure) 손실 함수일 수 있다.

[0018] 일 실시예에 따른 딥러닝 기반 영상 정합 장치는, 하나 이상의 인스트럭션을 저장하는 메모리와 상기 인스트럭션을 실행시키기 위한 프로세서를 포함하고, 상기 인스트럭션이 실행될 때, 상기 프로세서는, 제1 영상 및 상기 제1 영상과 정합되는 제2 영상을 포함하는 복수의 영상을 수신하고, 상기 복수의 영상의 특징 맵을 추출하고, 상기 특징 맵에 기초하여 상기 복수의 영상의 픽셀들의 이동 정보인 복수의 플로우 맵을 생성하고, 상기 복수의 영상에 상기 복수의 플로우 맵을 각각 적용하여 상기 복수의 영상의 픽셀들이 이동한 복수의 워핑 영상을 생성하고, 상기 특징 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩하기 위한 복수의 가중치 맵을 생성하고, 상기 복수의 가중치 맵에 기초하여 상기 복수의 워핑 영상을 블렌딩함으로써 상기 제1 영상과 상기 제2 영상이 정합된 정합 영상을 생성할 수 있다.

[0019] 상기 프로세서는, 상기 제1 영상의 픽셀들의 이동 정보인 제1 플로우 맵을 생성하고, 상기 제2 영상의 픽셀들의 이동 정보인 제2 플로우 맵을 생성할 수 있다.

[0020] 상기 프로세서는, 상기 제1 영상에 상기 제1 플로우 맵을 적용하여 제1 워핑 영상을 생성하고, 상기 제2 영상에 상기 제2 플로우 맵을 적용하여 제2 워핑 영상을 생성할 수 있다.

[0021] 상기 프로세서는, 상기 제1 워핑 영상에 대한 제1 가중치 맵을 생성하고, 상기 제2 워핑 영상에 대한 제2 가중치 맵을 생성할 수 있다.

[0022] 상기 프로세서는, 상기 제1 워핑 영상에 상기 제1 가중치 맵을 적용하고, 상기 제2 워핑 영상에 상기 제2 가중치 맵을 적용하고, 상기 제1 가중치가 적용된 제1 워핑 영상과 상기 제2 가중치가 적용된 제2 워핑 영상을 블렌딩하여 상기 정합 영상을 생성할 수 있다.

- [0023] 상기 프로세서는, 제1 손실함수와 제2 손실함수가 가중합 된 제3 손실 함수를 계산하고, 상기 제3 손실 함수에 기초하여 학습할 수 있다.
- [0024] 상기 제1 손실함수는, 상기 정합 영상의 픽셀 값과 상기 정답 영상의 픽셀 값의 차이를 계산하는 L1 손실 함수이고, 상기 제2 손실 함수는, 상기 정합 영상의 손실 정보를 평가하는 SSIM(Structural similarity index measure) 손실 함수일 수 있다.

도면의 간단한 설명

- [0026] 도 1은 일 실시예에 따른 영상 정합 장치의 개략적인 블록도를 나타낸다.
- 도 2a는 도1에 도시된 영상 정합 장치에 입력되는 제1 영상의 일 예를 나타낸다.
- 도 2b는 도1에 도시된 영상 정합 장치에 입력되는 제2 영상의 일 예를 나타낸다.
- 도 3a는 도1에 도시된 영상 정합 장치에 의해 생성된 제1 워핑 영상의 일 예를 나타낸다.
- 도 3b는 도1에 도시된 영상 정합 장치에 의해 생성된 제2 워핑 영상의 일 예를 나타낸다.
- 도 4는 도 1에 도시된 영상 정합 장치에 의해 생성된 정합 영상의 일 예를 나타낸다.
- 도 5는 일 실시예에 따른 영상 정합 장치의 다른 예를 나타낸다.

발명을 실시하기 위한 구체적인 내용

- [0027] 실시예들에 대한 특정한 구조적 또는 기능적 설명들은 단지 예시를 위한 목적으로 개시된 것으로서, 다양한 형태로 변경되어 구현될 수 있다. 따라서, 실제 구현되는 형태는 개시된 특정 실시예로만 한정되는 것이 아니며, 본 명세서의 범위는 실시예들로 설명한 기술적 사상에 포함되는 변경, 균등물, 또는 대체물을 포함한다.
- [0028] 제1 또는 제2 등의 용어를 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 해석되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.
- [0029] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다.
- [0030] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 설명된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함으로 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0031] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 해당 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0032] 이하, 실시예들을 첨부된 도면들을 참조하여 상세하게 설명한다. 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고, 이에 대한 중복되는 설명은 생략하기로 한다.
- [0034] 도 1은 일 실시예에 따른 영상 정합 장치의 개략적인 블록도를 나타낸다.
- [0035] 영상 정합 장치(100)는 제1 영상 $L(t)$ 및 제2 영상 $R(t)$ 을 수신하여, 제1 영상 $L(t)$ 과 제2 영상 $R(t)$ 을 정합한 정합 영상 $O(t)$ 을 생성할 수 있다. 또한, 영상 정합 장치(100)는 제1 영상 $L(t)$ 과 제2 영상 $R(t)$ 이 정확하게 정합된 정답 영상 $G(t)$ 을 수신하여 학습할 수 있다. 이때, 영상 정합 장치(100)는 복수 개의 손실 함수가 가중합

(weighted sum)된 손실 함수를 이용하여 학습할 수 있다.

- [0036] 영상 정합 장치(100)는 학습된 뉴럴 네트워크에 기초하여 영상을 정합할 수 있다.
- [0037] 뉴럴 네트워크(또는 인공 신경망)는 기계학습과 인지과학에서 생물학의 신경을 모방한 통계학적 학습 알고리즘을 포함할 수 있다. 뉴럴 네트워크는 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜, 문제 해결 능력을 가지는 모델 전반을 의미할 수 있다.
- [0038] 뉴럴 네트워크의 뉴런은 가중치 또는 바이어스의 조합을 포함할 수 있다. 뉴럴 네트워크는 하나 이상의 뉴런 또는 노드로 구성된 하나 이상의 레이어(layer)를 포함할 수 있다. 뉴럴 네트워크는 뉴런의 가중치를 학습을 통해 변화시킴으로써 임의의 입력으로부터 예측하고자 하는 결과를 추론할 수 있다.
- [0039] 뉴럴 네트워크는 심층 뉴럴 네트워크 (Deep Neural Network)를 포함할 수 있다. 뉴럴 네트워크는 CNN(Convolutional Neural Network), RNN(Recurrent Neural Network), 퍼셉트론(perceptron), 다층 퍼셉트론(multilayer perceptron), FF(Feed Forward), RBF(Radial Basis Network), DFF(Deep Feed Forward), LSTM(Long Short Term Memory), GRU(Gated Recurrent Unit), AE(Auto Encoder), VAE(Variational Auto Encoder), DAE(Denoising Auto Encoder), SAE(Sparse Auto Encoder), MC(Markov Chain), HN(Hopfield Network), BM(Boltzmann Machine), RBM(Restricted Boltzmann Machine), DBN(Deep Belief Network), DCN(Deep Convolutional Network), DN(Deconvolutional Network), DCIGN(Deep Convolutional Inverse Graphics Network), GAN(Generative Adversarial Network), LSM(Liquid State Machine), ELM(Extreme Learning Machine), ESN(Echo State Network), DRN(Deep Residual Network), DNC(Differentiable Neural Computer), NTM(Neural Turing Machine), CN(Capsule Network), KN(Kohonen Network) 및 AN(Attention Network)를 포함할 수 있다.
- [0040] 영상 정합 장치(100)는 마더보드(motherboard)와 같은 인쇄 회로 기판(printed circuit board(PCB)), 집적 회로(integrated circuit(IC)), 또는 SoC(system on chip)로 구현될 수 있다. 예를 들어, 영상 정합 장치(100)는 애플리케이션 프로세서(application processor)로 구현될 수 있다.
- [0041] 또한, 영상 정합 장치(100)는 PC(personal computer), 데이터 서버 등으로 구현될 수 있다.
- [0042] 영상 정합 장치(100)는 특징 맵 추출기(110), 플로우 맵 생성기(120), 가중치 맵 생성기(130), 워핑 영상 생성기(140), 가중합 계산기(150), 및 손실 함수 계산기(160)를 포함할 수 있다.
- [0043] 특징 맵 추출기(110)는 제1 영상 $L(t)$ 및 제2 영상 $R(t)$ 을 수신하여, 특징 맵 $F(t)$ 을 추출할 수 있다. 특징 맵 추출기(110)는 제1 영상 $L(t)$ 에 대응하는 제1 특징 맵 $F_L(t)$ 및 제2 영상 $R(t)$ 에 대응하는 제2 특징 맵 $F_R(t)$ 을 추출할 수 있다. 제1 영상 $L(t)$ 은 특정 시점의 왼쪽 방향을 촬영한 영상일 수 있고, 제2 영상 $R(t)$ 은 특정 시점의 오른쪽 방향을 촬영한 영상일 수 있다. 제1 영상 $L(t)$ 및 제2 영상 $R(t)$ 은 가상 시뮬레이터에 의해 획득될 수도 있다. 제1 영상 $L(t)$ 및 제2 영상 $R(t)$ 은 중첩되는 부분이 있을 수 있고, 제1 영상 $L(t)$ 과 대응되는 제1 특징 맵 $F_L(t)$ 과 제2 영상 $R(t)$ 과 대응되는 제2 특징 맵 $F_R(t)$ 은 결합될 수 있다. 특징 맵 추출기(110)는 제1 특징 맵 $F_L(t)$ 과 제2 특징 맵 $F_R(t)$ 을 결합하여 특징 맵 $F(t) = \{F_L(t), F_R(t)\}$ 을 추출할 수 있다.
- [0044] 플로우 맵 생성기(120)는 특징 맵 $F(t)$ 을 수신하여 플로우 맵(예: 제1 플로우 맵, 제2 플로우 맵)을 출력할 수 있다. 플로우 맵 생성기(120)는 특징 맵 $F(t)$ 으로부터 제1 영상 $L(t)$ 의 픽셀들의 이동 정보(movement information)인 제1 플로우 맵 $f_L^x(t), f_L^y(t)$, 제2 영상 $R(t)$ 의 픽셀들의 이동 정보인 제2 플로우 맵 $f_R^x(t), f_R^y(t)$ 을 생성할 수 있다. 영상의 픽셀의 좌표는 (x,y)로 표현될 수 있고, 플로우 맵 생성기(120)는 제1 영상 $L(t)$ 의 픽셀의 x좌표에 대한 플로우 맵 $f_L^x(t)$, 제1 영상 $L(t)$ 의 픽셀의 y좌표에 대한 플로우 맵 $f_L^y(t)$, 제2 영상 $R(t)$ 의 픽셀의 x좌표에 대한 플로우 맵 $f_R^x(t)$, 및 제2 영상 $R(t)$ 의 픽셀의 y좌표에 대한 플로우 맵 $f_R^y(t)$ 을 생성할 수 있다. 제1 플로우 맵 $f_L^x(t), f_L^y(t)$ 및 제2 플로우 맵 $f_R^x(t), f_R^y(t)$ 은 정합 영상 $O(t)$ 과 크기가 동일할 수 있다.
- [0045] 워핑 영상 생성기(140)는 영상(예: 제1 영상, 제2 영상) 및 플로우 맵(예: 제1 플로우 맵, 제2 플로우 맵)을 수신하여, 워핑 영상(예: 제1 워핑 영상, 제2 워핑 영상)을 출력할 수 있다. 워핑 영상 생성기(140)는 제1 영상

$L(t)$ 에 제1 플로우 맵 $f_L^x(t), f_L^y(t)$ 을 적용(예: 워핑)하여 제1 영상 $L(t)$ 의 픽셀들이 제1 플로우 맵 $f_L^x(t), f_L^y(t)$ 에 따라 이동한 제1 워핑 영상 $O_L(t)[x, y]$ 을 생성할 수 있고, 제2 영상 $R(t)$ 에 제2 플로우 맵 $f_R^x(t), f_R^y(t)$ 을 적용(예: 워핑)하여 제2 영상 $R(t)$ 의 픽셀들이 제2 플로우 맵 $f_R^x(t), f_R^y(t)$ 에 따라 이동한 제2 워핑 영상 $O_R(t)[x, y]$ 을 생성할 수 있다. 제1 워핑 영상 $O_L(t)[x, y]$ 및 제2 워핑 영상 $O_R(t)[x, y]$ 은 수학적식 1을 통해 계산될 수 있다.

[0047] [수학적식 1]

[0048] $O_L(t)[x, y] = L(t)[x + f_L^x(t), y + f_L^y(t)]$

[0049] $O_R(t)[x, y] = R(t)[x + f_R^x(t), y + f_R^y(t)]$

[0051] 여기서, $L(t)$ 는 특정 시점의 왼쪽 방향을 촬영한 제1 영상이고, $R(t)$ 는 특정 시점의 오른쪽 방향을 촬영한 제2 영상이고, x, y 는 영상의 픽셀의 좌표이고, t 는 시간이고, $f_L^x(t), f_L^y(t)$ 는 제1 영상의 픽셀들의 이동 정보인 제1 플로우 맵이고, $f_R^x(t), f_R^y(t)$ 는 제2 영상의 픽셀들의 이동 정보인 제2 플로우 맵일 수 있다. 워핑 영상 생성기(140)는 영상의 픽셀마다 워핑을 수행하여 시차 왜곡에 강인한 정합 영상 생성에 기여할 수 있다.

[0052] 가중치 맵 생성기(130)는 특징 맵 $F(t)$ 을 수신하여 가중치 맵(예: 제1 가중치 맵, 제2 가중치 맵)을 출력할 수 있다. 가중치 맵 생성기(130)는 제1 워핑 영상 $O_L(t)[x, y]$ 과 제2 워핑 영상 $O_R(t)[x, y]$ 를 블렌딩하기 위한 가중치 맵을 생성할 수 있다. 가중치 맵 생성기는 제1 워핑 영상 $O_L(t)[x, y]$ 에 대한 제1 가중치 맵 $w_L(t)$ 을 생성할 수 있고, 제2 워핑 영상 $O_R(t)[x, y]$ 에 대한 제2 가중치 맵 $w_R(t)$ 을 생성할 수 있다. 제1 가중치 맵 $w_L(t)$ 및 제2 가중치 맵 $w_R(t)$ 는 정합 영상 $O(t)$ 과 크기가 동일할 수 있다.

[0053] 가중합 계산기(150)는 워핑 영상(예: 제1 워핑 영상, 제2 워핑 영상) 및 가중치 맵(예: 제1 가중치 맵, 제2 가중치 맵)을 수신하여 제1 영상 $L(t)$ 과 제2 영상 $R(t)$ 이 정합된 정합 영상 $O(t)$ 을 출력할 수 있다. 가중합 계산기(150)는 제1 워핑 영상 $O_L(t)[x, y]$ 에 제1 가중치 맵 $w_L(t)$ 을 적용할 수 있고, 제2 워핑 영상 $O_R(t)[x, y]$ 에 제2 가중치 맵 $w_R(t)$ 을 적용할 수 있다. 가중합 계산기(150)는 제1 가중치 맵이 적용된 영상 $w_L(t) * O_L(t)$ 과 제2 가중치 맵 $w_R(t)$ 이 적용된 영상 $w_R(t) * O_R(t)$ 을 블렌딩하여 정합 영상 $O(t)$ 을 생성할 수 있다. 정합 영상 $O(t)$ 은 수학적식 2를 통해 계산될 수 있다.

[0055] [수학적식 2]

[0056] $O(t) = w_L(t) * O_L(t) + w_R(t) * O_R(t)$

[0058] 여기서, t 는 시간이고, $w_L(t)$ 는 제1 가중치이고, $w_R(t)$ 는 제2 가중치일 수 있다.

[0059] 정합 영상 $O(t)$ 은 특정 시점의 왼쪽 방향을 촬영한 제1 영상 $L(t)$ 및 특정 시점의 오른쪽 방향을 촬영한 제2 영상 $R(t)$ 을 정합한 영상일 수 있다.

[0060] 손실 함수 계산기(160)는 복수 개의 손실 함수(예: 제1 손실 함수, 제2 손실 함수)가 가중합된 손실 함수(예: 제3 손실 함수)를 계산할 수 있다. 손실 함수 계산기는 수신한 정합 영상 $G(t)$ 및 수신한 정합 영상 $G(t)$ 에 기초하여 제1 손실 함수(예: L1 함수)에 대한 제1 가중치 및 제2 손실 함수(예: SSIM(Structural similarity index measure) 함수)에 대한 제2 가중치를 설정할 수 있다. 제1 손실 함수는 수학적식 3을 통해 계산될 수 있다.

[0062] [수학식 3]

[0063]
$$L1(a, b) = \sum |a - b|$$

[0065] 여기서, a는 정합 영상 $O(t)$ 의 픽셀 값이고, b는 정답 영상 $G(t)$ 의 픽셀 값일 수 있다. 제1 손실 함수는 정합 영상 $O(t)$ 의 픽셀 값과 정답 영상 $G(t)$ 의 픽셀 값의 차이를 계산하는 L1함수일 수 있다. 제1 손실 함수를 이용하여 학습하면, 정합 영상 $O(t)$ 의 픽셀 값이 정답 영상 $G(t)$ 의 픽셀 값과 같아질 수 있다. 제2 손실 함수는 수학식 4를 통해 계산될 수 있다.

[0067] [수학식 4]

[0068]
$$SSIM(a, b) = \frac{(2\mu_a\mu_b+c_1)(\sigma_{ab}+c_2)}{(\mu_a^2+\mu_b^2+c_1)(\sigma_a^2+\sigma_b^2+c_2)}$$

[0070] 여기서, a는 정합 영상 $O(t)$ 의 픽셀 값이고, b는 정답 영상 $G(t)$ 의 픽셀 값이고, μ_a, μ_b 는 각각 정합 영상 $O(t)$ 의 픽셀 값의 평균, 정답 영상 $G(t)$ 의 픽셀 값의 평균이고, σ_a^2, σ_b^2 는 각각 정합 영상 $O(t)$ 의 픽셀 값의 분산, 정답 영상 $G(t)$ 의 픽셀 값의 분산이고, σ_{ab} 는 정합 영상 $O(t)$ 과 정답 영상 $G(t)$ 의 픽셀 값에 대한 공분산이고, c_1, c_2 는 계산의 안정화를 위한 상수일 수 있다. 제2 손실 함수는 정합 영상 $O(t)$ 의 손실 정보를 평가하는 SSIM(Structural similarity index measure) 함수일 수 있다. 제2 손실 함수를 이용하여 학습하면, 정합 영상 $O(t)$ 이 정답 영상 $G(t)$ 의 구조적인 자연스러움을 모사할 수 있다.

[0071] 손실 함수 계산기(160)는 제1 손실 함수에 대한 제1 가중치와 제2 손실 함수에 대한 제2 가중치를 설정하여 서로의 균형을 조절할 수 있다.

[0072] 영상 정합 장치(100)는 영상(예: 제1 영상, 제2 영상)의 픽셀들을 각각 옮기는 플로우 기반 딥러닝 네트워크를 통해 거리 및 깊이 차이에 의한 시차 왜곡에 강인한 영상 정합 방법을 제공할 수 있다. 또한 영상 정합 장치(100)는 복수의 손실 함수의 가중합을 통해 획득한 손실 함수를 통해 균형적인 학습을 수행할 수 있다.

[0074] 도 2a는 도1에 도시된 영상 정합 장치에 입력되는 제1 영상의 일 예를 나타내고, 도 2b는 도1에 도시된 영상 정합 장치에 입력되는 제2 영상의 일 예를 나타낸다.

[0075] 도 2a를 참조하면 특정 시점의 왼쪽 방향을 촬영한 영상인 제1 영상 $L(t)$ 의 일 예를 확인할 수 있고, 도 2b를 참조하면 특정 시점의 오른쪽 방향을 촬영한 영상인 제2 영상 $R(t)$ 의 일 예를 확인할 수 있다.

[0076] 제1 영상 $L(t)$ 의 맨 우측과 제2 영상 $R(t)$ 의 맨 좌측에 존재하는 가로등은 동일한 가로등이다. 즉, 제1 영상 $L(t)$ 과 제2 영상 $R(t)$ 은 중첩되는 부분이 존재한다. 영상 정합 장치(100)는 도 2a 및 도 2b에 각각 도시된 제1 영상 $L(t)$ 과 제2 영상 $R(t)$ 을 정합하여 정합 영상을 생성할 수 있다.

[0078] 도 3a는 도1에 도시된 영상 정합 장치에 의해 생성된 제1 위평 영상의 일 예를 나타내고, 도 3b는 도1에 도시된 영상 정합 장치에 의해 생성된 제2 위평 영상의 일 예를 나타낸다.

[0079] 제1 위평 영상 $O_L(t)[x, y]$ 은 제1 영상 $L(t)$ 의 픽셀들이 제1 플로우 맵 $f_L^x(t), f_L^y(t)$ 에 따라 이동한 영상일 수 있고, 제2 위평 영상 $O_R(t)[x, y]$ 은 제2 영상 $R(t)$ 의 픽셀들이 제1 플로우 맵 $f_R^x(t), f_R^y(t)$ 에 따라 이동한 영상일 수 있다.

- [0080] 도 2a와 도 3a를 비교하면, 제1 워핑 영상 $O_L(t)[x, y]$ 은 제1 영상 $L(t)$ 의 픽셀들이 이동하여 제1 워핑 영상 $O_L(t)[x, y]$ 의 좌측 및 제1 워핑 영상 $O_L(t)[x, y]$ 의 우측의 가로등이 흰 것을 확인할 수 있다. 또한 도 2b와 도 3b를 비교하면, 제2 워핑 영상 $O_R(t)[x, y]$ 은 제2 영상 $R(t)$ 의 픽셀들이 이동하여 제2 워핑 영상 $O_R(t)[x, y]$ 의 좌측이 흰 것을 확인할 수 있다.
- [0082] 도 4는 도 1에 도시된 영상 정합 장치에 의해 생성된 정합 영상의 일 예를 나타낸다.
- [0083] 영상 정합 장치(100)는 특정 시점의 왼쪽 방향을 촬영한 제1 영상 $L(t)$ 및 특정 시점의 오른쪽 방향을 촬영한 제2 영상 $R(t)$ 을 정합하여 정합 영상 $O(t)$ 을 생성할 수 있다.
- [0084] 영상 정합 장치(100)는 영상(예: 제1 영상, 제2 영상)의 픽셀들의 이동 정보인 플로우 맵과 영상을 적용(예: 워핑)하여 워핑 영상(예: 제1 워핑 영상, 제2 워핑 영상)을 생성할 수 있고, 워핑 영상을 가중합하여 정합 영상을 생성할 수 있다. 영상 정합 장치(100)는 플로우 맵(예: 제1 플로우 맵, 제2 플로우 맵)을 학습하여 높은 품질의 정합 영상을 생성할 수 있다.
- [0085] 영상 정합 장치(100)는 영상(예: 제1 영상, 제2 영상)의 픽셀들을 각각 옮기는 플로우 기반 딥러닝 네트워크를 통해 거리 및 깊이 차이에 의한 시차 왜곡에 강인한 영상 정합 방법을 제공할 수 있다.
- [0087] 도 5는 일 실시예에 따른 영상 정합 장치의 다른 예를 나타낸다.
- [0088] 영상 정합 장치(500)는 도 1의 영상 정합 장치(100)와 실질적으로 동일할 수 있다. 영상 정합 장치(500)는 메모리(510) 및 프로세서(530)를 포함할 수 있다.
- [0089] 메모리(510)는 프로세서(530)에 의해 실행가능한 인스트럭션들(예: 프로그램)을 저장할 수 있다. 예를 들어, 인스트럭션들은 프로세서(530)의 동작 및/또는 프로세서(530)의 각 구성의 동작을 실행하기 위한 인스트럭션들을 포함할 수 있다.
- [0090] 프로세서(530)는 메모리(510)에 저장된 데이터를 처리할 수 있다. 프로세서(530)는 메모리(510)에 저장된 컴퓨터로 읽을 수 있는 코드(예를 들어, 소프트웨어) 및 프로세서(530)에 의해 유발된 인스트럭션(instruction)들을 실행할 수 있다.
- [0091] 프로세서(530)는 목적하는 동작들(desired operations)을 실행시키기 위한 물리적인 구조를 갖는 회로를 가지는 하드웨어로 구현된 데이터 처리 장치일 수 있다. 예를 들어, 목적하는 동작들은 프로그램에 포함된 코드(code) 또는 인스트럭션들(instructions)을 포함할 수 있다.
- [0092] 예를 들어, 하드웨어로 구현된 데이터 처리 장치는 마이크로프로세서(microprocessor), 중앙 처리 장치(central processing unit), 프로세서 코어(processor core), 멀티-코어 프로세서(multi-core processor), 멀티프로세서(multiprocessor), ASIC(Application-Specific Integrated Circuit), FPGA(Field Programmable Gate Array)를 포함할 수 있다.
- [0093] 영상 정합 장치(500)에 의해 수행되는 동작은 도 1 내지 도 4를 참조하여 설명한 영상 정합 장치(100)의 동작과 실질적으로 동일할 수 있다. 이에, 상세한 설명은 생략하도록 한다.
- [0095] 이상에서 설명된 실시예들은 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치, 방법 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만,

해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0096] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0097] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 저장할 수 있으며 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.

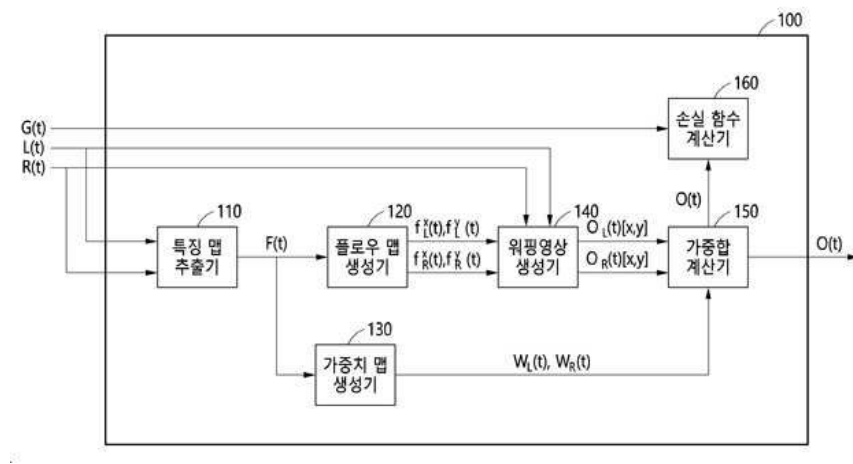
[0099] 위에서 설명한 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 또는 복수의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0100] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 이를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0101] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

도면1



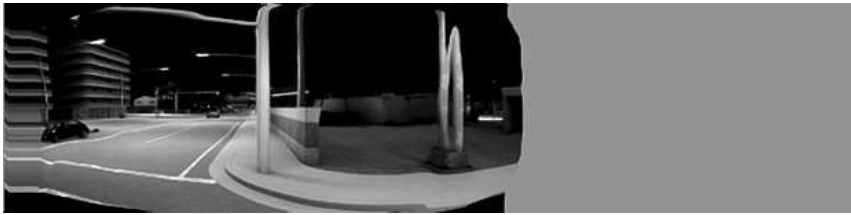
도면2a



도면2b



도면3a



도면3b



도면4



도면5

