(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년07월12일
(11) 등록번호 10-2555165
(24) 등록일자 2023년07월10일

(51) 국제특허분류(Int. Cl.)
G06T 5/50 (2006.01) *G06N 3/08* (2023.01)
G06T 7/557 (2017.01) *H04N 13/156* (2018.01)
(52) CPC특허분류
G06T 5/50 (2023.01)
G06N 3/08 (2023.01)
(21) 출원번호 10-2022-0126531
(22) 출원일자 2022년10월04일
심사청구일자 2022년10월04일
(56) 선행기술조사문헌
KR1020160034513 A
(뒷면에 계속)

(73) 특허권자
인하대학교 산학협력단
인천광역시 미추홀구 인하로 100(용현동, 인하대학교)
(72) 발명자
박인규
서울특별시 강남구 선릉로 221, 103동 1403호(도곡동, 도곡렉슬아파트)
백형선
경기도 시흥시 옥구천동로 451, 110동 801호(정왕동, 부성파스텔아파트)
(74) 대리인
양성보

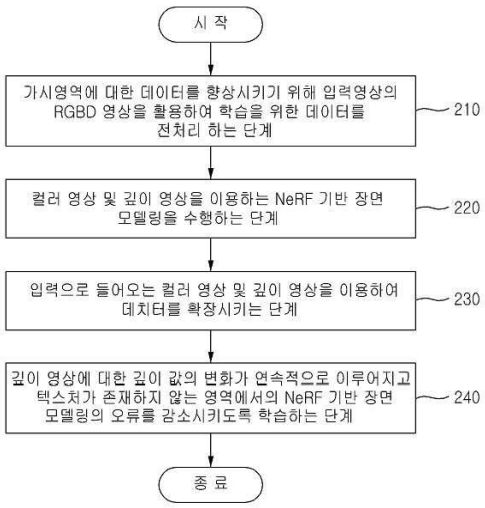
전체 청구항 수 : 총 8 항 심사관 : 김광식

(54) 발명의 명칭 **단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법 및 시스템**

(57) 요 약

단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법 및 시스템이 제시된다. 본 발명에서 제안하는 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법은 전처리부를 통해 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리하는 단계, 모델링부를 통해 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계, 학습부를 통해 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계 및 학습부를 통해 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 단계를 포함한다.

대 표 도 - 도2



(52) CPC특허분류
G06T 7/557 (2017.01)
H04N 13/156 (2018.05)
G06T 2207/10016 (2013.01)

(56) 선행기술조사문헌
KR1020210149425 A
KR1020210001005 A
KR1020170107269 A
JP2022020675 A

이 발명을 지원한 국가연구개발사업
과제고유번호 1711174118
과제번호 RS-2022-00155915
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 인공지능융합혁신인재양성(R&D)
연구과제명 인공지능융합혁신인재양성(인하대학교)
기 여 율 25/100
과제수행기관명 인하대학교 산학협력단
연구기간 2022.07.01 ~ 2022.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711160508
과제번호 2022-0-00981-001
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 방송통신산업기술개발사업
연구과제명 [Ezbaro] 전배경 정합 3D 객체 스트리밍 기술개발(1차년도)
기 여 율 50/100
과제수행기관명 인하대학교 산학협력단
연구기간 2022.04.01 ~ 2022.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711152759
과제번호 2020-0-01389-003
부처명 과학기술정보통신부
과제관리(전문)기관명 정보통신기획평가원
연구사업명 인공지능융합연구센터지원사업(국고)
연구과제명 [Ezbaro][정부] 인공지능융합연구센터지원(3차년도)
기 여 율 25/100
과제수행기관명 인하대학교 산학협력단
연구기간 2022.01.01 ~ 2022.12.31

공지예외적용 : 있음

명 세 서

청구범위

청구항 1

전처리부를 통해 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 상기 입력영상의 RGBD 영상에 관한 데이터를 전처리하는 단계;

모델링부를 통해 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계;

학습부를 통해 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계; 및

학습부를 통해 상기 학습부를 통해 데이터 확장된 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 단계

를 포함하는 라이트필드 합성 방법.

청구항 2

제1항에 있어서,

상기 전처리부를 통해 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 상기 입력영상의 RGBD 영상에 관한 데이터를 전처리하는 단계는,

DIBR(Depth Image Based Rendering) 방식으로 상기 가시영역에 대한 가상시점 영상을 획득하고 해당 광선들이 가시 영역인지 여부를 결정할 마스킹를 생성하여 이후에 학습될 NeRF에 입력으로 들어가며, 마스킹 되는 부분의 광선은 학습에서 제외되는

라이트필드 합성 방법.

청구항 3

제1항에 있어서,

상기 모델링부를 통해 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계는,

가시 영역에 대한 샘플링 레이트를 증가 시키고 폐색 영역의 경우에는 비디오 방향에 따른 입력으로부터 데이터를 추정하여, 렌더링 기반의 입력 영상으로부터 라이트필드의 SAIs(Sub-aperture images) 간의 베이스라인이 넓은 데이터들을 대규모로 취득하여 라이트 필드 데이터셋을 생성하는

라이트필드 합성 방법.

청구항 4

제1항에 있어서,

상기 학습부를 통해 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계는,

깊이 기반 렌더링 방식을 사용하여 입력되는 시점의 주변 시점들을 생성하고, 넓은 베이스라인을 갖도록 DIBR로 워핑시키는 범위를 조절하며, 전체 SAIs를 생성하지 않고 일정 부분의 SAIs를 선택하여 데이터를 확장시키는

라이트필드 합성 방법.

청구항 5

가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 상기 입력영상의

RGBD 영상에 관한 데이터를 전처리하는 전처리부;

상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 모델링부; 및

상기 입력으로 들어오는 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키고, 학습부를 통해 데이터 확장된 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 학습부

를 포함하는 라이트필드 합성 시스템.

청구항 6

제5항에 있어서,

상기 전처리부는,

DIBR(Depth Image Based Rendering) 방식으로 상기 가시영역에 대한 가상시점 영상을 획득하고 해당 광선들이 가시 영역인지 여부를 결정할 마스킹를 생성하여 이후에 학습될 NeRF에 입력으로 들어가며, 마스킹 되는 부분의 광선은 학습에서 제외하는

라이트필드 합성 시스템.

청구항 7

제5항에 있어서,

상기 모델링부는,

가시 영역에 대한 샘플링 레이트를 증가 시키고 폐색 영역의 경우에는 비디오 방향에 따른 입력으로부터 데이터를 추정하여, 렌더링 기반의 입력 영상으로부터 라이트필드의 SAIs(Sub-aperture images) 간의 베이스라인이 넓은 데이터들을 대규모로 취득하여 라이트 필드 데이터셋을 생성하는

라이트필드 합성 시스템.

청구항 8

제5항에 있어서,

상기 학습부는,

깊이 기반 렌더링 방식을 사용하여 입력되는 시점의 주변 시점들을 생성하고, 넓은 베이스라인을 갖도록 DIBR로 워핑시키는 범위를 조절하며, 전체 SAIs를 생성하지 않고 일정 부분의 SAIs를 선택하여 데이터를 확장시키는

라이트필드 합성 시스템.

발명의 설명

기 술 분 야

[0001] 본 발명은 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법 및 시스템에 관한 것이다.

배 경 기 술

[0002] 하나 이상의 입력 영상으로부터 가상시점 영상을 합성하는 연구는 VR/AR, 3D 디스플레이, 홀로그램 등의 콘텐츠와 함께 실감 미디어 환경에서의 다양한 적용방향이 존재한다는 점에서 많은 관심과 연구가 진행되어왔다.

[0003] 일반적으로 가상시점 생성은 입력 영상과 카메라 파라미터로부터 영상들의 픽셀들 간의 관계를 활용하거나, 부족한 기하학적 정보 추정을 위한 깊이 정보, 포인트 클라우드 등의 추가적인 정보를 활용하여 장면을 표현하는 것이 일반적이다. 따라서 기존의 실감 미디어 구성을 위한 시스템은 다수의 카메라를 이용하여 특정 목적에 맞도록 배치함으로써 데이터를 수집, 가공하는 방식을 사용한다. 이로인해 다중 카메라 구성에 따른 증가하는 비

용 문제와 더불어 다시점 영상들 간의 동기화 문제, 다시점 영상 간의 관계 추정에 필요한 연산량 증가 등의 여러 문제를 갖고 있다는 점에서 여러 제약이 존재한다.

[0004] 이러한 문제점을 해결하기 위해서 다시점 영상을 동시에 획득하는 하드웨어들이 제안되었으며 대표적으로는 양안 카메라, 라이트필드 카메라 등이 존재한다. 또한 그래픽 처리 장치(Graphic Processing Unit; GPU)의 기술 발전과 대규모의 데이터로부터의 딥러닝 아키텍처의 급격한 성장으로부터 카메라의 개수를 최소화하면서 기존의 목적에 부합하는 다양한 시점의 영상들을 생성할 수 있는 연구들이 수행되었다.

[0005] 딥러닝을 활용한 다수의 가상시점 연구들은 다중 시점 기반의 가상시점 생성부터 더욱 나아가 단안영상까지도 딥러닝 프레임워크를 사용하여 가상시점 생성을 가능하게 하고 있다. 세부적으로는 기존의 장면 표현 방식에 학습 기반의 네트워크를 결합하여 학습함으로써 장면을 모델링 하거나, 새로운 장면표현 방식을 제안함으로써 이전의 장면 표현 방식들에 비해 속도 또는 영상 품질면에서 높은 성능을 보여주는 연구들도 수행되었다. 이는 새로운 장면 표현 방식이 입력 영상 간의 관계를 추정하고 최적화하는 전통적인 컴퓨터 비전 기법에 비해 일반화와 에러에 강인할 수 있다는 점을 알 수 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) [1] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in Proc. of European Conference on Computer Vision, pp. 405- 421, Springer, 2020.

(특허문헌 0002) [2] K. Deng, A. Liu, J.-Y. Zhu, and D. Ramanan, "Depth-supervised NeRF:

(특허문헌 0003) Fewer views and faster training for free," in Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12882-12891, 2022.

발명의 내용

해결하려는 과제

[0007] 본 발명이 이루고자 하는 기술적 과제는 정방향으로 움직이는 비디오의 프레임들을 다시점 영상과 같이 사용함으로써 가상시점 영상 합성의 기법들을 활용하여 라이트필드를 합성하는 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법 및 시스템을 제공하는데 있다. 라이트필드는 구조화된 다시점 영상을 갖는다는 장점으로 실감 미디어 환경을 구축할 수 있다는 점에서 가상시점 합성 연구의 필요성과 합성한 결과의 평가에 효과적이다. 본 발명에서는 두 가지 장면 표현 방식을 통해 라이트필드를 생성하는 방안을 제안한다.

과제의 해결 수단

[0008] 일 측면에 있어서, 본 발명에서 제안하는 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법은 전처리부를 통해 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리하는 단계, 모델링부를 통해 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계, 학습부를 통해 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계 및 학습부를 통해 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 단계를 포함한다.

[0009] 상기 전처리부를 통해 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리하는 단계는 DIBR(Depth Image Based Rendering) 방식으로 상기 가시영역에 대한 가상시점 영상을 획득하고 해당 광선들이 가시 영역인지 여부를 결정할 마스킹를 생성하여 이후에 학습될 NeRF에 입력으로 들어가며, 마스킹 되는 부분의 광선은 학습에서 제외된다.

[0010] 상기 모델링부를 통해 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계는 가시 영역에 대한 샘플링 레이트를 증가 시키고 폐색 영역의 경우에는 비디오 방향에 따른 입력으로부터 데이터를 추정하여, 렌더링 기반의 입력 영상으로부터 라이트필드의 SAIs(Sub-aperture images) 간의

베이스라인이 넓은 데이터들을 대규모로 취득하여 라이트 필드 데이터셋을 생성한다.

[0011] 상기 학습부를 통해 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계는 깊이 기반 렌더링 방식을 사용하여 입력되는 시점의 주변 시점들을 생성하고, 넓은 베이스라인을 갖도록 DIBR로 워핑시키는 범위를 조절하며, 전체 SAIs를 생성하지 않고 일정 부분의 SAIs를 선택하여 데이터를 확장시킨다.

[0012] 또 다른 일 측면에 있어서, 본 발명에서 제안하는 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 시스템은 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리하는 전처리부, 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 모델링부, 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키고, 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 학습부를 포함한다.

발명의 효과

[0013] 본 발명의 실시예들에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법 및 시스템을 통해 정방향으로 움직이는 비디오의 프레임들을 다시점 영상과 같이 사용함으로써 가상시점 영상 합성의 기법들을 활용하여 라이트필드를 합성할 수 있다. 라이트필드는 구조화된 다시점 영상을 갖는다는 장점으로 실감 미디어 환경을 구축할 수 있다는 점에서 가상시점 합성 연구의 필요성과 합성한 결과의 평가에 효과적이다. 본 발명에서는 두 가지 장면 표현 방식을 통해 라이트필드를 생성하는 방안을 제안한다.

도면의 간단한 설명

[0014] 도 1은 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성을 위한 전체 프레임워크를 나타내는 도면이다.

도 2는 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법을 설명하기 위한 흐름도이다.

도 3은 본 발명의 일 실시예에 따른 실제환경 실험 시 입력으로 받는 컬러 영상 및 깊이 영상을 나타내는 도면이다.

도 4는 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 시스템의 구성을 나타내는 도면이다.

도 5는 본 발명의 일 실시예에 따른 합성한 결과의 SAI의 예시를 나타내는 도면이다.

도 6은 본 발명의 일 실시예에 따른 본 발명의 일 실시예에 따른 네트워크를 사용하여 생성한 라이트필드를 나타내는 도면이다.

도 7은 본 발명의 일 실시예에 따른 KITTI 데이터 넓은 베이스라인 정성적 결과를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0015] 본 발명에서는 넓은 베이스라인을 갖는 라이트필드를 합성함으로써 쉽게 취득가능한 라이트 필드 데이터 생성 방법을 제안하고 이를 통해 취득한 대규모의 라이트 필드 비디오를 통하여 단안 비디오로부터 라이트필드를 생성하는 네트워크를 제안한다. 제안하는 네트워크는 기존의 가상 시점 합성에서 높은 활용성을 보이는 MPIs(Multiplane-images)와 깊이 추정을 결합하는 방식으로 두 가지의 장면 표현 방식이 가지고 있는 특성을 결합하는 프레임워크이다.

[0016] 본 발명의 실시예에 따르면, 사실적 렌더링 기반의 비디오 게임인 GTAV를 활용하여 8×8 SAIs로 구성된 라이트필드 영상 비디오를 생성한다.

[0017] 본 발명의 실시예에 따르면, 사실적 렌더링 기반의 비디오 게임으로부터 라이트필드의 SAIs(Sub-aperture images) 간의 베이스라인이 넓은 데이터들을 대규모로 취득하여 라이트 필드 데이터셋을 생성한다.

[0018] 또한, 다중 평면 기반과 DIBR(Depth image-based rendering) 기법을 결합하여 렌더링 가용범위를 향상시키면서도 높은 품질의 라이트필드를 합성하고, NeRF(Neural radiance field) 기반의 장면 모델링 방식을 통해 기존의 라이트 필드 합성 기법과 다르게 넓은 베이스라인을 갖는 라이트필드를 합성한다. 이하, 본 발명의 실시 예를

첨부된 도면을 참조하여 상세하게 설명한다.

[0020] NeRF는 하나의 장면을 다양한 시점으로부터 입력된 정보를 토대로 장면 그 자체를 하나의 네트워크로 표현하는 것을 의미한다. NeRF 모델링 과정에서 네트워크는 단일한 5D 좌표를 입력으로 이용하고, 공간상 한 지점에서의 볼륨(volume) 밀도와 빛에 따라 달라지는 빛을 출력하는 완전 연결 네트워크(fully connected network)를 이용해 장면을 모델링한다. 훈련된 네트워크를 활용하여 가상시점을 합성하는 경우에는 카메라 빛을 따라서 5D 좌표의 출력값을 이용해 시점을 합성하고, 전통적인 볼륨 렌더링 기법을 이용해 샘플링된 포인트들의 색과 밀도를 영상으로 합성한다. 처음으로 이러한 장면 표현 방식을 제안한 NeRF의 경우 정적인 장면을 다중 레이어의 퍼셉트론을 활용하여 공간상의 점 (x, y, z)와 방향 (λ,θ)으로 방출하는 빛, 각 점 (x, y, z)마다 빛이 얼마나 통과할 것인지를 결정하는 밀도를 출력으로 하는 연속적인 5D 함수로 표현하였다.

[0021] $r(t) = o + td$ (1)

[0022] 입력영상에서의 하나의 픽셀은 카메라 투영중심 o와 방향 d를 따라 진행된다. 따라서 특정 시점의 영상을 렌더링 하기 위해서는 장면의 카메라 빛들을 따라가며 공간상의 3D 포인트들을 샘플링하고 이 포인트들과 대응하는 2D 시점 방향을 신경망의 입력으로 사용해 각 포인트들의 색과 밀도를 얻어낸 후, 전통적인 볼륨 렌더링(volume rendering) 방식을 이용해 색과 밀도를 누적하는 방식으로 2D 영상을 만들어낸다.

[0023] 이러한 과정은 모두 미분 가능하기 때문에, 관찰한 이미지와 이에 대응되는 렌더링된 결과물 간의 차이를 줄여나가며 모델을 최적화하는 과정에서 그래디언트 경사하강법을 사용할 수 있다. NeRF의 경우 이러한 광선기반의 표현방식만을 사용하여 훈련을 진행하면 고주파수의 영상정보는 손실된 상태로 복원되거나 세부표현들이 제거된다는 특성에 주목하여 아래와 같은 기법을 같이 제안하였다. 첫 번째로는 공간상의 위치와 방향의 차원이 (x, y, z), (λ,θ)로 5개의 차원으로만 이루어져 있기에 MLP(Multi-Layer Perceptron)에 충분한 정보를 제공하지 못하고 손실된다는 점을 보완하고자 위치 인코딩(positional encoding)으로 입력차원을 증가시켜 고차원 데이터로 매핑하여 입력으로 사용하였다. 이러한 위치 인코딩은 고주파 영역의 처리 성능을 향상된 방향으로 학습되도록 유도하며 주기함수를 통해 다음과 같이 표현할 수 있다.

[0024] $\gamma(x) = [\sin(x), \cos(x), \ldots, \sin(2^{L-1}x), \cos(2^{L-1}x)]$ (2)

[0025] 위 식은 단순히 3D 위치 공간상의 한 점 x의 각 차원의 사인함수와 코사인함수로 1에서 2^{L-1}까지 2의 거듭제곱으로 표현한다. 여기서 L은 하이퍼파라미터에 해당 되며 NeRF에서는 위치정보와 각정보에 특정값을 이용하여 주기함수를 통해 임베딩시킨다. 하지만 일반적인 도로장면에서는 단순히 모델링이 이루어지기 어렵다. NeRF의 경우에는 입력으로 들어오는 영상들이 모두 유사한 거리에서 획득한 다량의 데이터라는 점에서 멀티뷰를 통한 장면의 기하학적인 내용을 네트워크를 통해 쉽게 학습되도록 제한하였다. 이에 따라 입력으로 사용되는 영상의 형태가 제한적이며 하나의 객체만이 존재하는 장면은 높은 성능을 보여주며 시점합성에 뛰어난 성능을 보여준다. 그러나 일반적인 상황에서 이러한 영상들을 획득하는 것은 어려운 일에 속하므로 입력영상의 개수를 줄이거나 일반적인 장면으로 학습하는 연구들이 진행되고 있다.

[0026] 본 발명에서는 기존의 사용되는 데이터에서 벗어나 일반적인 도로상황을 가정하고 폐색영역이 빈번하게 나타나는 장면을 NeRF기반으로 모델링하는 기법을 제안한다. 이러한 복합적인 장면에 대하여 기존의 NeRF는 고해상도의 결과물을 낼 만큼 충분히 수렴하지 못하고, 카메라 광선당 요구 샘플 개수 측면에서 부족하기 때문에 정확한 기하학적 추정에 실패한다. 이러한 문제를 해결하기 위해 컬러 정보 이외의 깊이 정보를 추가적으로 제공함으로써 기존의 모델에서 필요한 입력영상의 개수를 줄이도록 하였다. 이는 [2]에서 제안된 방법과 유사하게 진행된다. 하지만 [2]에서는 자세정보 추정과정에서 영상들 간의 관계를 파악하고 영상 간의 일치하는 3차원 포인트들을 학습에 사용하지만 본 발명에서는 깊이 정보 영상 전체를 사용하여 네트워크를 훈련한다. 이는 MLP가 카메라의 움직임에 따라 변화하는 정보를 학습할 수 있게 하고 이에 따라 3차원 기하학적 관계 추정에 큰 방향성을 제시하므로 학습의 수렴에 도움이 된다. 다음으로는 RGBD 영상을 활용하는 방식으로 단일 MLP만으로 학습에 어려움을 겪는 파트를 따로 모델링함으로써 샘플링하는 과정에서 요구되는 적정 샘플의 개수에 맞춰 광선들을 확장한다.

[0028] 본 발명의 실시예에 따르면, 제안하는 라이트필드 합성 네트워크 학습을 위한 데이터를 취득하기 위해 사실적 렌더링 기반의 비디오 게임을 사용하였다. 본 발명에서는 깊이 영상을 추가적으로 획득했다는 점에서 종래기술과 차이가 있다.

[0029] 종래기술에 따르면, CNN을 학습시키기 위해 대량의 데이터를 취득해야 했지만 본 발명에서는 여러 장면을 구성

할 정도의 데이터만이 필요하므로 이에 맞도록 수정하여 데이터를 생성한다.

[0030] 본 발명의 실시예에 따르면, 실세계를 모델링한 가상환경 기반의 게임이라는 점에서 학습에 사용할 데이터를 취득한 것이며 이는 실제환경 데이터에서의 동작을 통해 확인할 수 있다. 가상 데이터의 경우 세부적인 데이터 셋 취득 방법은 다음과 같다.

[0031] 본 발명의 실시예에 따르면, 가상 환경상의 데이터를 다루기 위해서 Script Hook V library를 사용하였다. 해당 라이브러리를 사용하여 가상환경 내에 다중 카메라 배열을 구성하고, 각각의 카메라를 렌더링하여 컬러영상과 깊이 영상을 추출함으로써 하나의 라이트 필드 영상과 대응되는 깊이 영상들을 취득하였다. 실제 환경의 데이터는 KITTI 데이터를 사용하였으며, 라이다 센서를 통해 취득한 데이터를 활용하여 학습에 사용할 깊이 정보 영상으로 가공하였다.

[0032] 본 발명에서는 기존의 NeRF 기반의 방식이 주요하게 다루던 단일 객체를 정면 또는 360 바라보는 데이터에서 벗어나 좀더 일반적인 데이터에서 작동하는 노벨 뷰 합성(Novel view synthesis)을 통해 베이스라인이 기존의 기법들보다 확장된 라이트필드를 생성하는 것을 목표로 한다.

[0034] 도 1 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성을 위한 전체 프레임워크를 나타내는 도면이다.

[0035] 도 1은 발명에서 제안하고자 하는 알고리즘의 전체적인 구성을 보여준다.

[0036] 제안하는 뉴럴 레이디언스 기반의 라이트필드 합성 방법은 하나의 장면을 다양한 시점으로부터 입력된 정보(110)를 토대로 장면 그 자체를 하나의 네트워크로 표현한다.

[0037] 첫 번째로 컬러 영상을 감독(supervision)으로 간주하여 모델링(121)하는 NeRF의 방식을 전체적으로 따른다. 그러나 다양한 각도의 조밀한 입력을 가정하고 만들어진 기법인 만큼 비디오를 입력으로 함에 따라 제한된 입력으로부터 퀄리티 높은 결과를 생성하기 위해 깊이 정보를 추가적으로 감독(supervision)으로 사용한다(122). 이것은 RGB영상만으로 입력을 받을 경우 멀티뷰의 특성을 이용하여 장면의 기하학적인 해석이 가능하지만 비디오를 입력으로 할 경우에는 이를 보완하고자 하는 목적으로 사용된다. 두 번째로는 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 거의 존재하지 않는 도로부분에서 뉴럴 레이디언스 필드(Neural radiance field)의 생성에 문제점을 확인하여 이를 해결하고자 학습에 악영향을 미치는 도로부분을 해결하기 위한 방안을 제안한다(130). 위와 같이 본 발명에서 제안하는 기법은 좀더 일반적인 환경의 데이터로의 확장을 목표로 한다.

[0039] 도 2는 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법을 설명하기 위한 흐름도이다.

[0040] 제안하는 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 방법은 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리하는 단계(210), 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 단계(220), 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키는 단계(230) 및 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 단계(240)를 포함한다.

[0041] 단계(210)에서, 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리 한다.

[0042] 본 발명에서 목표로 하고 있는 도로 주행의 데이터의 경우에는 기존의 NeRF 기법을 활용하여 해결 하는데 여러 한계점이 존재한다. 기존의 입력과는 다르게 카메라가 움직이고 장면이 고정되지 않았다는 점에서 전체적인 학습에 어려움을 갖고 있다. 특히 텍스처를 구별하기 힘들고 큰 특징이 존재하지 않는 도로부분에서 네트워크는 깊이 추정 또는 3차원 구조 추정에 어려움을 갖는다. 또한 도로부분은 영상의 전체적인 부분에서 나타나면서도 다양한 깊이로 연속적으로 픽셀들이 이어진다는 점은 더욱 문제를 어렵게 한다.

[0043] 종래기술에 따른 NeRF는 동일한 거리에 있는 데이터를 중심으로 네트워크를 설계하였기에 데이터의 경우 많은 입력을 통해 기하학적인 추정을 할 수 있지만 직선방향의 소수의 프레임으로부터는 획득하기 어려운 것으로 보인다. 따라서 대부분 가시영역인 도로부분에 대한 데이터를 향상시키기 위해 입력으로 들어오는 RGBD 영상을 활용한다. 기존의 DIBR방식으로 시점을 합성할 때의 가장 큰 문제점은 폐색영역에 대한 정보를 얻기 힘들다는 것이다. 하지만 도로의 경우 폐색영역이 존재하는 경우가 거의 없고 대부분 시야에 보이므로 DIBR을 활용하여 가상시점의 데이터를 생성한다. 결과적으로는 DIBR 방식으로 가시영역에 대한 가상시점 영상을 획득하고 해당 광

선들이 가시 영역인지 결정할 마스킹를 생성하게 된다. 이는 이후에 학습할 뉴럴 레디언스 필드에 입력으로 들어가며, 마스킹 되는 부분의 광선은 학습에서 제외한다.

[0044] 단계(220)에서, 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행한다.

[0045] 본 발명의 실시예에 따른 전체적인 프레임워크는 NeRF를 제안하였던 기존의 NeRF의 파이프라인과 유사하게 진행이 된다. 그러나 기존의 방식에서는 고려하지 않았던 깊이 정보를 추가적으로 사용하였고 전처리 과정을 통해 적은 수의 입력으로부터 가시 영역에 대한 샘플들에 대해 여러 데이터를 확보할 수 있다.

[0047] 도 3은 본 발명의 일 실시예에 따른 실제환경 실험 시 입력으로 받는 컬러 영상 및 깊이 영상을 나타내는 도면이다.

[0048] 도 3(a)는 컬러 영상을 나타내는 도면이고, 도 3(b)는 깊이 영상을 나타내는 도면이다.

[0049] 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링은 결과적으로 가시 영역에 대한 샘플링 레이트를 증가 시키고 폐색 영역의 경우에는 비디오 방향에 따른 입력으로부터 추정한다. 8 × 8 SAI(Sub-aperture images) 모두를 생성하여 학습시키기에는 연산량 증가가 크므로 일정 SAI를 선택하여 진행하였다.

[0051] 단계(230)에서, 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시킨다.

[0052] 본 발명에서는 입력 단안 비디오로부터 넓은 베이스라인을 갖는 라이트 필드를 합성하는 NeRF 기반의 네트워크를 제안한다. 또한 제안된 네트워크에 사용할 데이터를 생성하여 합성한 라이트필드의 성능을 검증한다. 제안하는 방 법은 두 개의 스테이지로 구별된다. 첫 번째는 입력으로 들어오는 컬러 영상과 깊이 영상을 이용하여 데이터를 확장시킨다. 깊이 기반 렌더링 방식을 사용하여 입력으로 들어온 시점 주변 시점들을 생성한다. 이는 직선방향 주행으로 발생하는 희소한 입력으로부터 MLP가 기하학적인 추론을 원활하게 하도록 돕는다. 또한 기존의 기법에 비해 넓은 베이스라인을 갖기 위해 DIBR로 워핑시키는 범위를 조 절하였으며, 전체 SAI를 생성하지는 않고 일정 부분의 SAI 를 선택하여 데이터를 증강하였다.

[0053] 단계(240)에서, 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습한다.

[0054] 본 발명의 실시예에 따른 학습의 전반적인 과정은 가시영역을 포함하는 가상시점을 생성하고 입력 RGBD 를 통해 NeRF를 모델링하게 된다. 이외에는 [1]에서 제안한 NeRF의 원활한 학습을 위한 위치 임베딩(positional embedding), 계층적 볼륨 샘플링(Hierarchical volume sampling)과 같은 기법들을 사용하였다. End-to-End로 네트워크가 훈련되며 훈련 과정에서 사용한 손실함수는 단순히 L2함수만을 사용하였다. 본 발명에서 제안한 네트워크는 ADAM 옵티마이저(optimizer)를 사용하였으며, 학습률은 0.0005으로 시작하여 학습률을 감소시켜가며 학습을 진행한다. 실험은 NVIDIA RTX A6000 48GB GPU를 통해 약 1일 정도의 기간 동안 학습하였다.

[0056] 도 4는 본 발명의 일 실시예에 따른 단안 비디오에서 뉴럴 레이디언스 기반의 라이트필드 합성 시스템의 구성을 나타내는 도면이다.

[0057] 본 실시예에 따른 라이트필드 합성 시스템(400)은 프로세서(410), 버스(420), 네트워크 인터페이스(430), 메모리(440) 및 데이터베이스(450)를 포함할 수 있다. 메모리(440)는 운영체제(441) 및 라이트필드 합성 루틴(442)을 포함할 수 있다. 프로세서(410)는 전처리부(411), 모델링부(412) 및 학습부(413)를 포함할 수 있다. 다른 실시예들에서 라이트필드 합성 시스템(400)은 도 4의 구성요소들보다 더 많은 구성요소들을 포함할 수도 있다. 그러나, 대부분의 종래기술적 구성요소들을 명확하게 도시할 필요성은 없다. 예를 들어, 라이트필드 합성 시스템(400)은 디스플레이나 트랜시버(transceiver)와 같은 다른 구성요소들을 포함할 수도 있다.

[0058] 메모리(440)는 컴퓨터에서 판독 가능한 기록 매체로서, RAM(random access memory), ROM(read only memory) 및 디스크 드라이브와 같은 비소멸성 대용량 기록장치(permanent mass storage device)를 포함할 수 있다. 또한, 메모리(440)에는 운영체제(441)와 라이트필드 합성 루틴(442)을 위한 프로그램 코드가 저장될 수 있다. 이러한 소프트웨어 구성요소들은 드라이브 메커니즘(drive mechanism, 미도시)을 이용하여 메모리(440)와는 별도의 컴퓨터에서 판독 가능한 기록 매체로부터 로딩될 수 있다. 이러한 별도의 컴퓨터에서 판독 가능한 기록 매체는 플로피 드라이브, 디스크, 테이프, DVD/CD-ROM 드라이브, 메모리 카드 등의 컴퓨터에서 판독 가능한 기록 매체(미도시)를 포함할 수 있다. 다른 실시예에서 소프트웨어 구성요소들은 컴퓨터에서 판독 가능한 기록 매체가 아닌 네트워크 인터페이스(430)를 통해 메모리(440)에 로딩될 수도 있다.

[0059] 버스(420)는 라이트필드 합성 시스템(400)의 구성요소들간의 통신 및 데이터 전송을 가능하게 할 수 있다. 버스(420)는 고속 시리얼 버스(high-speed serial bus), 병렬 버스(parallel bus), SAN(Storage Area Network) 및/또는 다른 적절한 통신 기술을 이용하여 구성될 수 있다.

[0060] 네트워크 인터페이스(430)는 라이트필드 합성 시스템(400)을 컴퓨터 네트워크에 연결하기 위한 컴퓨터 하드웨어 구성요소일 수 있다. 네트워크 인터페이스(430)는 라이트필드 합성 시스템(400)을 무선 또는 유선 커넥션을 통해 컴퓨터 네트워크에 연결시킬 수 있다.

[0061] 데이터베이스(450)는 라이트필드 합성을 위해 필요한 모든 정보를 저장 및 유지하는 역할을 할 수 있다. 도 4에서는 라이트필드 합성 시스템(400)의 내부에 데이터베이스(450)를 구축하여 포함하는 것으로 도시하고 있으나, 이에 한정되는 것은 아니며 시스템 구현 방식이나 환경 등에 따라 생략될 수 있고 혹은 전체 또는 일부의 데이터베이스가 별개의 다른 시스템 상에 구축된 외부 데이터베이스로서 존재하는 것 또한 가능하다.

[0062] 프로세서(410)는 기본적인 산술, 로직 및 라이트필드 합성 시스템(400)의 입출력 연산을 수행함으로써, 컴퓨터 프로그램의 명령을 처리하도록 구성될 수 있다. 명령은 메모리(440) 또는 네트워크 인터페이스(430)에 의해, 그리고 버스(420)를 통해 프로세서(410)로 제공될 수 있다. 프로세서(410)는 전처리부(411), 모델링부(412) 및 학습부(413)를 위한 프로그램 코드를 실행하도록 구성될 수 있다. 이러한 프로그램 코드는 메모리(440)와 같은 기록 장치에 저장될 수 있다.

[0063] 전처리부(411), 모델링부(412) 및 학습부(413)는 도 2의 단계들(210~240)을 수행하기 위해 구성될 수 있다.

[0064] 라이트필드 합성 시스템(400)은 전처리부(411), 모델링부(412) 및 학습부(413)를 포함할 수 있다.

[0065] 전처리부(411)는 가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리한다.

[0066] 전처리부(411)는 DIBR(Depth Image Based Rendering) 방식으로 상기 가시영역에 대한 가상시점 영상을 획득하고 해당 광선들이 가시 영역인지 여부를 결정할 마스킹를 생성하여 이후에 학습될 NeRF에 입력으로 들어가며, 마스킹 되는 부분의 광선은 학습에서 제외한다.

[0067] 모델링부(412)는 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행한다.

[0068] 모델링부(412)는 가시 영역에 대한 샘플링 레이트를 증가 시키고 폐색 영역의 경우에는 비디오 방향에 따른 입력으로부터 데이터를 추정하여, 렌더링 기반의 입력 영상으로부터 라이트필드의 SAIs(Sub-aperture images) 간의 베이스라인이 넓은 데이터들을 대규모로 취득하여 라이트 필드 데이터셋을 생성한다.

[0069] 학습부(413)는 상기 입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키고, 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습한다.

[0070] 학습부(413)는 깊이 기반 렌더링 방식을 사용하여 입력되는 시점의 주변 시점들을 생성하고, 넓은 베이스라인을 갖도록 DIBR로 워핑시키는 범위를 조절하며, 전체 SAIs를 생성하지 않고 일정 부분의 SAIs를 선택하여 데이터를 확장시킨다.

[0072] 도 5는 본 발명의 일 실시예에 따른 네트워크를 사용하여 생성한 라이트필드를 나타내는 도면이다.

[0073] 도 5(a)는 G.T., 도 5(b)는 Srinivasan, 도 5(c)는 NeRF, 도 5d)는 본 발명의 실시예에 따른 결과를 나타낸다.

[0074] 시점 합성 기술에서는 일반적으로 지도 학습 기반의 방법을 사용 하므로 정량적 평가를 위해서는 GT 영상을 필요로 한다. 본 발명에서의 도로 주행 라이트 필드 데이터셋을 통해 직접 생성한 라이트필드 비디오 데이터를 이용하여 정량적 평가를 수행한다. 제안하는 방법은 학습에 사용되지 않은 시점에서부터 광 선들을 쿼리하여 SAI를 생성한다. 다음으로는 실제환경에서의 검증을 위해 도로 주행 데이터셋인 KITTI를 이용하여 결과를 생성한다. KITTI 데이터셋은 양안 카메라 기반으로 도로를 주행하며 데이터를 생성하였기에 정량적 평가를 불가능하며, 정성적 평가로만 비교 분석하였다. 실험에 사용된 데이터는 학습과정에서 사용한 480X640 공간해상도를 통해 수행하였으며, KITTI 데이터 또한 사이즈를 변경하여 실험에 사용하였다.

[0075] 도 5는 제안하는 네트워크를 사용하여 생성한 라이트필드를 보여준다. 뉴럴 레디언스를 기반으로 하는 방식이 장면에 따른 최적화를 목표로 한다는 점에서 실제환경과 가상환경상에서의 성능차이는 거의 존재하지 않는 것을

확인 할 수 있었다.

[0076] 본 발명에서 제안하는 프레임워크에 대해 정성적 평가를 수행 하였다. 도 6은 각각의 방법으로 생성한 SAI를 보여주고 있다. 도 5(a)의 경우는 단안으로부터 깊이정보를 추정하는 것에 기반을 둔 Srinivasan기법에 해당하는 SAI의 예시이고, 도 5(b)의 경우에는 전체적인 기하학적 구조를 정확히 추정하지 못한 것을 볼 수 있다. 또한 도 5(c)의 경우에는 기하적으로 복잡하거나 다양한 물체들이 존재하는 경우 깊이 값을 정확히 추정하지 못해 플로팅되는 아티팩트를 많이 생성하는 한계를 볼 수 있다. 정성적 평가의 경우에는 가상환경 및 실제환경 모두에서 진행하였다. NeRF와 제안하는 프레임워크로 생성한 결과를 비교하였을 때, 제안하는 방식의 경우가 상대적으로 G.T의 기하적인 형태를 전체적으로 보존하고, 도로부분에서도 정확히 표현하는 것을 확인할 수 있다.

[0078] 도 6은 본 발명의 일 실시예에 따른 합성한 결과의 SAI의 예시를 나타내는 도면이다.

[0079] 도 7은 본 발명의 일 실시예에 따른 KITTI 데이터 넓은 베이스라인 정성적 결과를 나타내는 도면이다.

[0080] 도 6은 라이트필드의 일정 부분의 SAI를 보여주고있다. 도 6(a)는 가상환경 데이터, 도 6(b)는 KITTI 데이터를 나타낸다. 비교에서 확인한 것과 같이 베이스라인이 커짐에도 전체적으로 기하적인 구조를 유지하며 높은 품질을 유지한다.

[0082] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 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, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 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 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

[0083] 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 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

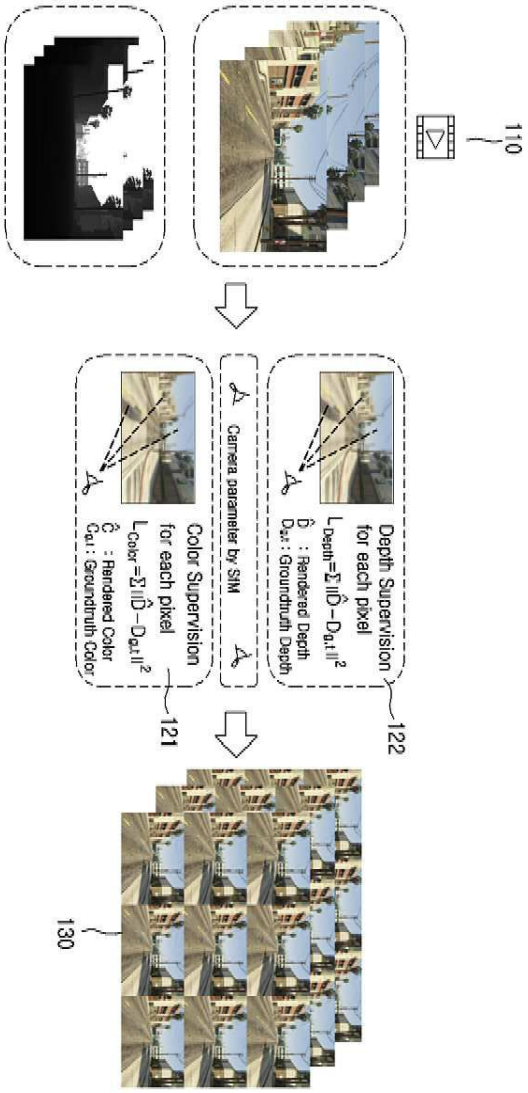
[0084] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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.

[0085] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 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.

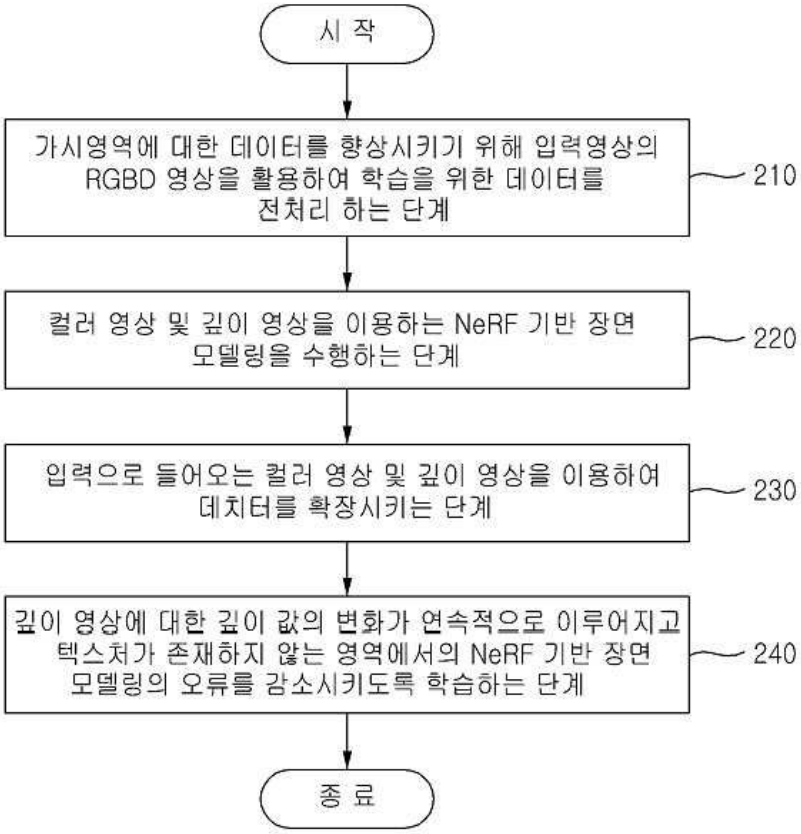
[0086] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

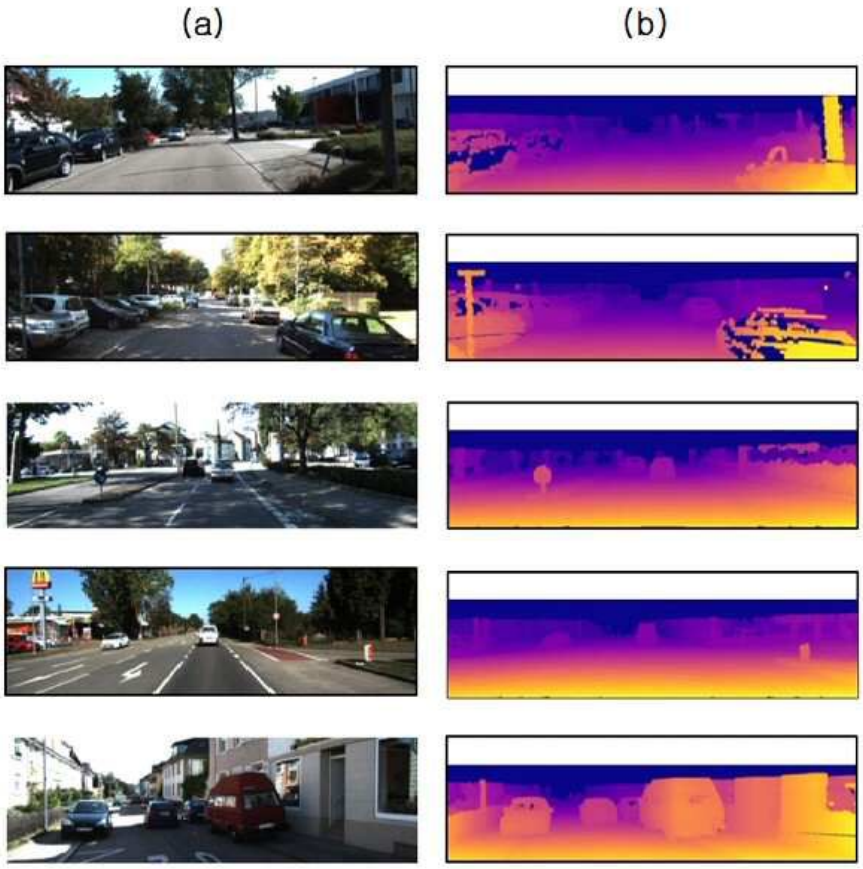
도면1



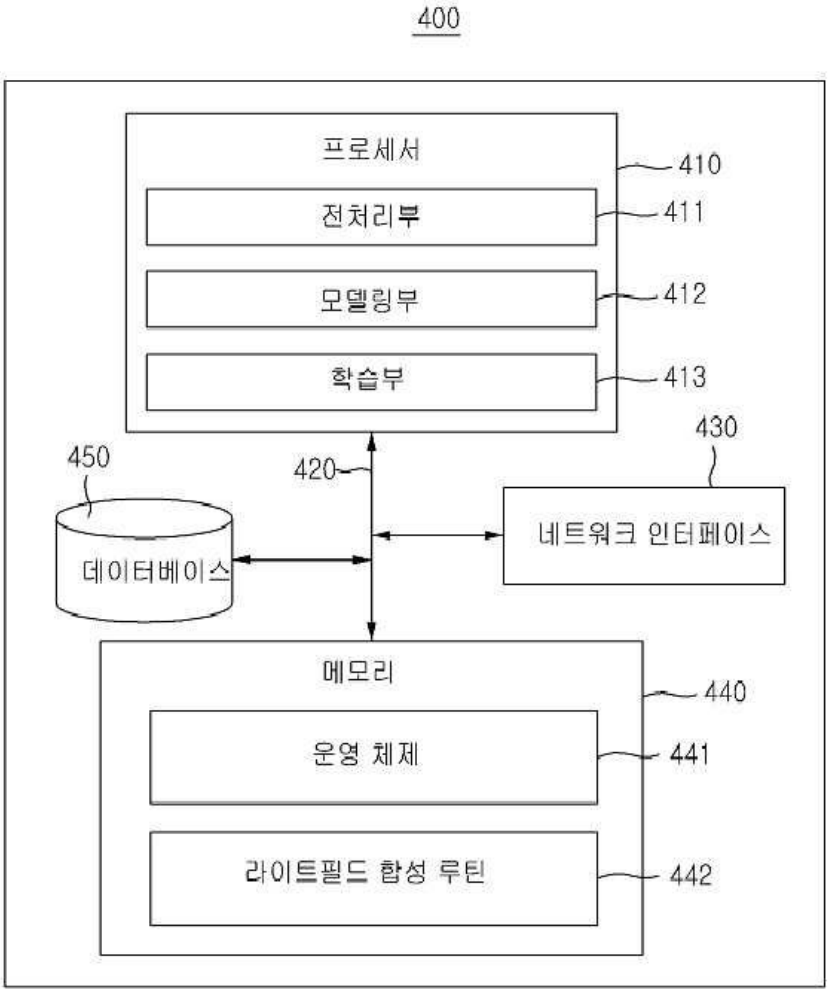
도면2


시 작
가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 데이터를 전처리 하는 단계
210
컬러 영상 및 깊이 영상을 이용하는 NeRF 기반 장면 모델링을 수행하는 단계
220
입력으로 들어오는 컬러 영상 및 깊이 영상을 이용하여 데치터를 확장시키는 단계
230
깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면 모델링의 오류를 감소시키도록 학습하는 단계
240
종 료

도면3



도면4


400
프로세서
410
전처리부
411
모델링부
412
학습부
413
430
450
420
네트워크 인터페이스
데이터베이스
메모리
440
운영 체제
441
라이트필드 합성 루틴
442

도면5



(a)



(b)



(c)



(d)

도면6



(a)



(b)

도면7



(a)



(b)

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 5

【변경전】

가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 상기 입력영상의 RGBD 영상에 관한 데이터를 전처리하는 전처리부;

상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field) 기반 장면모델링을 수행하는 모델링부; 및

상기 입력으로 들어오는 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키고, 상기 학습부를 통해 데이터 확장된 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 학습부

를 포함하는 라이트필드 합성 시스템.

【변경후】

가시영역에 대한 데이터를 향상시키기 위해 입력영상의 RGBD 영상을 활용하여 학습을 위한 상기 입력영상의 RGBD 영상에 관한 데이터를 전처리하는 전처리부;

상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하는 NeRF(Nueral radiance field)

기반 장면모델링을 수행하는 모델링부; 및

상기 입력으로 들어오는 상기 전처리된 입력영상의 RGBD 영상에 대한 컬러 영상 및 깊이 영상을 이용하여 데이터를 확장시키고, 학습부를 통해 데이터 확장된 깊이 영상에 대한 깊이 값의 변화가 연속적으로 이루어지고 텍스처가 존재하지 않는 영역에서의 NeRF 기반 장면모델링의 오류를 감소시키도록 학습하는 학습부

를 포함하는 라이트필드 합성 시스템.