(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0179013
(43) 공개일자 2024년12월31일

(51) 국제특허분류(Int. Cl.)
G06T 15/08 (2011.01) *G06T 15/06* (2011.01)
G06T 15/20 (2011.01)
(52) CPC특허분류
G06T 15/08 (2013.01)
G06T 15/06 (2013.01)
(21) 출원번호 10-2023-0081494
(22) 출원일자 2023년06월23일
심사청구일자 2023년06월23일

(71) 출원인
계명대학교 산학협력단
대구광역시 달서구 달구벌대로 1095, 계명대학교 산학협력관 201호(신당동)
(72) 발명자
고병철
대구광역시 수성구 범어천로 107
김선빈
대구광역시 달서구 비슬로 2724
(74) 대리인
김건우

전체 청구항 수 : 총 16 항

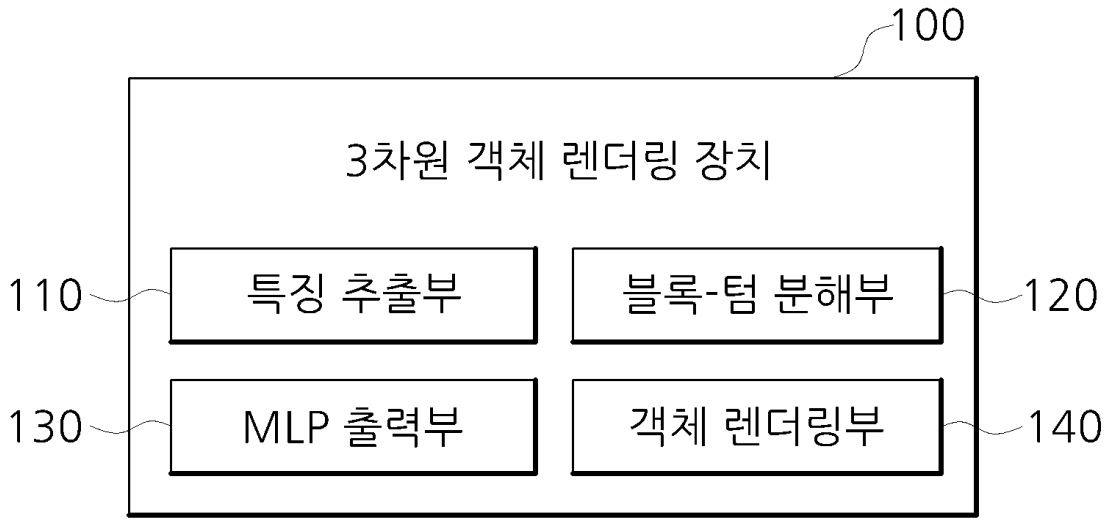
(54) 발명의 명칭 **블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법**

(57) 요 약

본 발명은 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치에 관한 것으로서, 보다 구체적으로는 3차원 객체 렌더링 장치로서, 객체의 3차원 위치와 레이(ray) 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 특징 추출부; 상기 추출한 Q, K 및

(뒷면에 계속)

대 표 도 - 도1



V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 블록-텀 분해부; 상기 결합한 결괏값을 MLP(Multi-Layer Perceptron)에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 MLP 출력부; 및 상기 특징 추출부, 블록-텀 분해부 및 MLP 출력부의 실행을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링하는 객체 렌더링부를 포함하는 것을 그 구성상의 특징으로 한다.

또한, 본 발명은 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법에 관한 것으로서, 보다 구체적으로는 컴퓨터에서 각 단계가 수행되는 3차원 객체 렌더링 방법으로서, (1) 객체의 3차원 위치와 레이(ray) 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 단계; (2) 상기 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 단계; (3) 상기 결합한 결괏값을 MLP(Multi-Layer Perceptron)에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 단계; 및 (4) 상기 단계 (1) 내지 단계 (3)을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

본 발명에서 제안하고 있는 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법에 따르면, 복셀(Voxel) 기반의 NeRF를 이용하되, 고차원의 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하되 벡터 없이 행렬만을 사용하는 블록-텀 분해를 적용함으로써, 기존의 NeRF보다 학습 속도와 렌더링 속도가 빠르고 렌더링 성능이 우수하며, 다른 텐서 분해 기법들보다 상대적으로 적은 수의 파라미터로 목표 텐서에 근사할 수 있어서 계산량을 줄이고 성능을 개선할 수 있다.

(52) CPC특허분류
G06T 15/20 (2013.01)
G06T 2207/20084 (2013.01)

명 세 서

청구범위

청구항 1

3차원 객체 렌더링 장치(100)로서,

객체의 3차원 위치와 레이(ray) 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 특징 추출부(110);

상기 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 블록-텀 분해부(120);

상기 결합한 결괏값을 MLP(Multi-Layer Perceptron)에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 MLP 출력부(130); 및

상기 특징 추출부(110), 블록-텀 분해부(120) 및 MLP 출력부(130)의 실행을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링하는 객체 렌더링부(140)를 포함하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 2

제1항에 있어서, 상기 특징 추출부(110)는,

양선형 보간법(bilinear interpolation)을 이용해 Q, K 및 V에 해당하는 특징을 추출하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 3

제1항에 있어서, 상기 블록은,

각각 CP 분해 또는 터커(Tucker) 분해로 표현되는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 4

제3항에 있어서, 상기 블록-텀 분해부(120)는,

상기 Q, K 및 V 행렬을 이용해 터커(Tucker) 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 터커 분해부(121);

상기 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 멀티헤드 어텐션부(122); 및

상기 P개의 결괏값을 결합하는 연결부(123)를 포함하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 5

제4항에 있어서, 상기 터커 분해부(121)는,

코어 텐서와 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행하는 것을 특징으로 하는, 블록-텀 분해 기반의 신

경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 6

제5항에 있어서, 상기 코어 텐서는,

가우시안 표준정규분포를 따르는 랜덤값으로 설정된 것을 특징을 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 7

제1항에 있어서, 상기 객체 렌더링부(140)는,

모든 레이에 대해 생성된 출력값을 이용해 각각의 레이에 해당하는 픽셀값을 계산하여, 볼륨 렌더링을 수행하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 8

제7항에 있어서, 상기 객체 렌더링부(140)는,

디퍼런셔블(Differentiable) 볼륨 렌더링을 사용해 렌더링하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100).

청구항 9

컴퓨터에서 각 단계가 수행되는 3차원 객체 렌더링 방법으로서,

(1) 객체의 3차원 위치와 레이(ray) 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 단계;

(2) 상기 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 단계;

(3) 상기 결합한 결괏값을 MLP(Multi-Layer Perceptron)에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 단계; 및

(4) 상기 단계 (1) 내지 단계 (3)을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링하는 단계를 포함하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 10

제9항에 있어서, 상기 단계 (1)에서는,

양선형 보간법(bilinear interpolation)을 이용해 Q, K 및 V에 해당하는 특징을 추출하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 11

제9항에 있어서, 상기 단계 (2)에서 상기 블록은,

각각 CP 분해 또는 터커(Tucker) 분해로 표현되는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3

차원 객체 렌더링 방법.

청구항 12

제11항에 있어서, 상기 단계 (2)는,

(2-1) 상기 Q, K 및 V 행렬을 이용해 터커(Tucker) 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 단계;

(2-2) 상기 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 단계; 및

(2-3) 상기 P개의 결괏값을 결합하는 단계를 포함하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 13

제12항에 있어서, 상기 단계 (2-1)에서는,

코어 텐서와 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 14

제13항에 있어서, 상기 코어 텐서는,

가우시안 표준정규분포를 따르는 랜덤값으로 설정된 것을 특징을 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 15

제9항에 있어서, 상기 단계 (4)에서는,

모든 레이에 대해 생성된 출력값을 이용해 각각의 레이에 해당하는 픽셀값을 계산하여, 볼륨 렌더링을 수행하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

청구항 16

제15항에 있어서, 상기 단계 (4)에서는,

디퍼런셔블(Differentiable) 볼륨 렌더링을 사용해 렌더링하는 것을 특징으로 하는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법.

발명의 설명

기 술 분 야

[0001] 본 발명은 3차원 객체 렌더링 장치 및 방법에 관한 것으로서, 보다 구체적으로는 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법에 관한 것이다.

배 경 기 술

[0002] 이 부분에 기술된 내용은 단순히 본 발명의 일실시예에 대한 배경 정보를 제공할 뿐 종래기술을 구성하는 것은 아니다.

[0004] 최근 컴퓨터 비전 분야에서 AR과 VR 등 다양한 기술이 발전함에 따라 3D 객체를 고품질의 이미지로 렌더링하는 인공지능 모델들이 크게 주목을 받고 있다. 특히 Multi-Layer Perceptron(MLP)을 활용하여 실제 객체와 유사한 이미지를 렌더링하는 모델인 신경 방사 필드(Neural Radiance Fields: NeRF)(비특허문헌 0001)는 높은 암시적 신경 표현 능력으로 작은 모델 크기에도 불구하고 우수한 렌더링 성능을 보인다.

[0006] NeRF는 객체를 찍은 여러 장의 이미지를 입력받아 새로운 시점에서의 객체 이미지를 만들어 내는(View synthesis) 모델이다. NeRF는 물체의 위치 정보(spatial location)인 (x,y,z)와 물체가 바라보는 방향(viewing direction, 카메라의 한 점과 각도)인 (θ,φ)를 포함한 5차원 데이터를 입력으로 받아, RGB 값과 물체의 밀도(density)를 예측하는 완전 연결 계층(fully-connected network, FC)을 학습한다. 여기서 물체의 밀도는 투명도의 역수로, 밀도가 커지면 물체가 불투명해지고(뒤에 있는 것들이 잘 보이지 않음), 밀도가 작아지면 물체가 투명해진다.

[0008] 하지만 NeRF는 수많은 뷰포인트(Viewpoints)에 대한 최적화가 필요하기 때문에, 계산 비용이 높다는 단점이 있다. 예를 들어, 학습 속도가 하루 정도로 느리고 렌더링 속도가 30초 정도로 느리다. 또한, 렌더링 성능이 좋지 않은 한계가 있다. 따라서 이를 해결할 수 있는 기술의 개발이 필요하다.

[0010] 전술한 배경 기술은 발명자가 본 발명의 도출을 위해 보유하고 있었거나, 본 발명의 도출 과정에서 습득한 기술 정보로서, 반드시 본 발명의 출원 전에 일반 공중에게 공개된 공지 기술이라 할 수는 없다.

선행기술문헌

비특허문헌

[0011] (비특허문헌 0001) Ben Mildenhall, Pratul P.Srinivasan, Matthew Tancik, Jonathan T.Barron, Ravi Ramamoorthi, Ren Ng, NeRF: Representing scenes as neural radiance fields for view synthesis, European Conference on Computer Vision,pp. 405-421, 2020.

(비특허문헌 0002) Lieven De Lathauwer, Decompositions of a Higher-Order Tensor in Block Terms-Part II: Definitions and uniqueness, SIAM, Journal on Matrix Analysis and Applications, 30(3), pp. 1033-1066, 2008.

(비특허문헌 0003) Anpei Chen, Zexiang Xu, Andreas Geiger, Jingyi Yu, Hao Su, TensoRF: Tensorial radiance fields, European Conference on Computer Vision, pp. 333-350, 2022.

(비특허문헌 0004) Tamara G. Kolda, Brett W. Bader, Tensor decompositions and applications, SIAM review, 51(3), pp. 455-500, 2009.

(비특허문헌 0005) J. Douglas Carroll, Jih-Jie Chang, Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition, Psychometrika, 35(3), pp. 283-319, 1970.

(비특허문헌 0006) Ledyard R Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika, 31(3), pp. 279-311, 1966.

(비특허문헌 0007) Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, Attention Is All you Need, NeurIPS, pp. 5998-6008, 2017.

(비특허문헌 0008) Nelson Max, Optical Models for Direct Volume Rendering, IEEE Transactions on Visualization and Computer Graphics, vol. 1, no. 2, pp. 99-108, 1995.

(비특허문헌 0009) Vincent Sitzmann, Michael Zollhoefer, Gordon Wetzstein, Scene Representation Networks: Continuous 3D-Structure-Aware Neural Scene Representation, NeurIPS, 2019.

(비특허문헌 0010) Lingjie Liu, Jaitao Gu, Kyaw Zaw Lin, Tat-Seng Chau, Christian Theobalt, Neural Sparse Voxel Fields, NeurIPS, 2020.

(비특허문헌 0011) Peter Hedman, Pratul P. Srinivasan, Ben Mildenhall, Jonathan T. Barron, Paul Debevec, Baking Neural Radiance Fields for Real-Time View Synthesis, International Conference on Computer Vision, pp. 5875-5884, 2021.

(비특허문헌 0012) Alex Yu, Ruilong Li, Matthew Tancik, Hao Li, Ren Ng, Angjoo Kanazawa, PlenOctrees: For Real-time Rendering of Neural Radiance Fields, International Conference on Computer Vision, pp. 5752-5761, 2021.

(비특허문헌 0013) Sara Fridovich-Keil, Alex Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, Angjoo Kanazawa, Plenoxels: Radiance Fields without Neural Networks, Computer Vision and Pattern Recognition, pp. 5491-5500, 2022.

(비특허문헌 0014) Cheng Sun, Min Sun, Hwann-Tzong Chen, Direct Voxel Grid Optimization: Super-fast Convergence for Radiance Fields Reconstruction, Computer Vision and Pattern Recognition, pp. 5449-5459, 2022.

발명의 내용

해결하려는 과제

[0012] 본 발명은 기존에 제안된 방법들의 상기와 같은 문제점들을 해결하기 위해 제안된 것으로서, 복셀(Voxel) 기반의 NeRF를 이용하되, 고차원의 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하되 벡터 없이 행렬만을 사용하는 블록-텀 분해를 적용함으로써, 기존의 NeRF보다 학습 속도와 렌더링 속도가 빠르고 렌더링 성능이 우수하며, 다른 텐서 분해 기법들보다 상대적으로 적은 수의 파라미터로 목표 텐서에 근사할 수 있어서 계산량을 줄이고 성능을 개선할 수 있는, 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법을 제공하는 것을 그 목적으로 한다.

[0014] 다만, 본 발명이 이루고자 하는 기술적 과제는 상기한 바와 같은 기술적 과제로 한정되지 않으며, 또 다른 기술적 과제들이 존재할 수 있고, 명시적으로 언급하지 않더라도 과제의 해결수단이나 실시 형태로부터 파악될 수 있는 목적이나 효과도 이에 포함됨은 물론이다.

과제의 해결 수단

[0015] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치는,

[0016] 3차원 객체 렌더링 장치로서,

[0017] 객체의 3차원 위치와 레이(ray) 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 특징 추출부;

[0018] 상기 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 블록-텀 분해부;

[0019] 상기 결합한 결괏값을 MLP(Multi-Layer Perceptron)에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 MLP 출력부; 및

[0020] 상기 특징 추출부, 블록-텀 분해부 및 MLP 출력부의 실행을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링하는 객체 렌더링부를 포함하는 것을 그 구성상의 특징으로 한다.

[0022] 바람직하게는, 상기 특징 추출부는,

[0023] 양선형 보간법(bilinear interpolation)을 이용해 Q, K 및 V에 해당하는 특징을 추출할 수 있다.

[0025] 바람직하게는, 상기 블록은,

[0026] 각각 CP 분해 또는 터커(Tucker) 분해로 표현될 수 있다.

[0028] 더욱 바람직하게는, 상기 블록-텀 분해부는,

[0029] 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 터커 분해부;

[0030] 상기 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 멀티헤드 어텐션부; 및

[0031] 상기 P개의 결괏값을 결합하는 연결부를 포함할 수 있다.

[0033] 더더욱 바람직하게는, 상기 터커 분해부는,

[0034] 코어 텐서와 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행할 수 있다.

[0036] 더욱더 바람직하게는, 상기 코어 텐서는,

[0037] 가우시안 표준정규분포를 따르는 랜덤값으로 설정될 수 있다.

[0039] 바람직하게는, 상기 객체 렌더링부는,

[0040] 모든 레이에 대해 생성된 출력값을 이용해 각각의 레이에 해당하는 픽셀값을 계산하여, 볼륨 렌더링을 수행할 수 있다.

[0042] 더욱 바람직하게는, 상기 객체 렌더링부는,

[0043] 디퍼런셔블(Differentiable) 볼륨 렌더링을 사용해 렌더링할 수 있다.

[0045] 상기한 목적을 달성하기 위한 본 발명의 특징에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법은,

[0046] 컴퓨터에서 각 단계가 수행되는 3차원 객체 렌더링 방법으로서,

[0047] (1) 객체의 3차원 위치와 레이 정보를 이용해, 상기 레이가 객체를 통과할 때 상기 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출하는 단계;

[0048] (2) 상기 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해를 수행해 상기 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 단계;

[0049] (3) 상기 결합한 결괏값을 MLP에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 단계; 및

[0050] (4) 상기 단계 (1) 내지 단계 (3)을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링을 통해 3차원 객체를 렌더링하는 단계를 포함하는 것을 그 구성상의 특징으로 한다.

[0052] 바람직하게는, 상기 단계 (1)에서는,

[0053] 양선형 보간법을 이용해 Q, K 및 V에 해당하는 특징을 추출할 수 있다.

[0055] 바람직하게는, 상기 단계 (2)에서 상기 블록은,

[0056] 각각 CP 분해 또는 터커 분해로 표현될 수 있다.

[0058] 더욱 바람직하게는, 상기 단계 (2)는,

[0059] (2-1) 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 단계;

[0060] (2-2) 상기 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 단계; 및

[0061] (2-3) 상기 P개의 결괏값을 결합하는 단계를 포함할 수 있다.

[0063] 더더욱 바람직하게는, 상기 단계 (2-1)에서는,

[0064] 코어 텐서와 상기 Q, K 및 V 행렬을 이용해 터커 분해를 수행할 수 있다.

[0066] 더욱더 바람직하게는, 상기 코어 텐서는,

[0067] 가우시안 표준정규분포를 따르는 랜덤값으로 설정될 수 있다.

[0069] 바람직하게는, 상기 단계 (4)에서는,

[0070] 모든 레이에 대해 생성된 출력값을 이용해 각각의 레이에 해당하는 픽셀값을 계산하여, 볼륨 렌더링을 수행할 수 있다.

[0072] 더욱 바람직하게는, 상기 단계 (4)에서는,

[0073] 디퍼런셔블 볼륨 렌더링을 사용해 렌더링할 수 있다.

발명의 효과

[0074] 본 발명에서 제안하고 있는 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법에 따르면, 복셀 기반의 NeRF를 이용하되, 고차원의 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하되 벡터 없이 행렬만을 사용하는 블록-텀 분해를 적용함으로써, 기존의 NeRF보다 학습 속도와 렌더링 속도가 빠르고 렌더링 성능이 우수하며, 다른 텐서 분해 기법들보다 상대적으로 적은 수의 파라미터로 목표 텐서에 근사할 수 있어서 계산량을 줄이고 성능을 개선할 수 있다.

[0076] 더불어, 본 발명의 다양하면서도 유익한 장점과 효과는 상술한 내용에 한정되지 않으며, 본 발명의 구체적인 실시 형태를 설명하는 과정에서 보다 쉽게 이해될 수 있을 것이다.

도면의 간단한 설명

[0077] 도 1은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치의 구성을 도시한 도면.

도 2는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치의 전체적인 구조를 도시한 도면.

도 3은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치에서 사용되는

블록-텀 분해를 설명하기 위해 도시한 도면.

도 4는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치의 특징 추출부에서 추출하는 Q, K, V 행렬을 설명하기 위해 도시한 도면.

도 5는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치에서, 블록-텀 분해부의 세부적인 구성을 도시한 도면.

도 6은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치의 블록-텀 분해부가 결괏값을 얻는 과정을 도시한 도면.

도 7은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법의 흐름을 도시한 도면.

도 8은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법에서, 단계 S120의 세부적인 흐름을 도시한 도면.

도 9는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법과 다른 방법에 의한 렌더링 수행 결과를 비교하여 나타낸 도면.

도 10은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치 및 방법에 의한 다양한 시점에서의 렌더링 결과는 타나낸 도면.

발명을 실시하기 위한 구체적인 내용

[0078] 아래에서는 첨부한 도면을 참조하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명의 실시예를 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0080] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "간접적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미하며, 하나 또는 그 이상의 다른 특징이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0082] 이하의 실시예는 본 발명의 이해를 돕기 위한 상세한 설명이며, 본 발명의 권리 범위를 제한하는 것이 아니다. 따라서 본 발명과 동일한 기능을 수행하는 동일 범위의 발명 역시 본 발명의 권리 범위에 속할 것이다.

[0084] 또한, 본 발명의 각 실시예에 포함된 각 구성, 과정, 공정 또는 방법 등은 기술적으로 상호간 모순되지 않는 범위 내에서 공유될 수 있다.

[0086] 또한, 본 발명에 있어서 단말, 장치 또는 디바이스가 수행하는 것으로 기술된 동작이나 기능 중 일부는 해당 단말, 장치 또는 디바이스와 연결된 서버에서 대신 수행될 수 있다. 마찬가지로, 서버가 수행하는 것으로 기술된 동작이나 기능 중 일부도 해당 서버와 연결된 단말, 장치 또는 디바이스에서 수행될 수도 있다.

[0088] 특히, 본 발명의 각 실시예에 따른 시스템을 실행시키기 위한 수단으로는 애플리케이션(Application), 또는 웹 서버일 수 있으며, 이 애플리케이션, 또는 웹 서버를 기록한 기록매체를 읽을 수 있는 수단인 단말로는, 일반적인 데스크톱이나 노트북 등의 일반 PC뿐만 아니라, 스마트 폰, 태블릿 PC, 등의 모바일 단말기를 포함할 수 있

다.

[0090] 이하, 첨부된 도면을 참고하여 본 발명의 실시예들을 상세히 설명하도록 한다.

[0092] 도 1은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 구성을 도시한 도면이다. 도 1에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)는, 3차원 객체 렌더링 장치(100)로서, 특징 추출부(110), 블록-텀 분해부(120), MLP 출력부(130) 및 객체 렌더링부(140)를 포함하여 구성될 수 있다.

[0094] 도 2는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 전체적인 구조를 도시한 도면이다. 도 2에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)에서는, 첫 번째 그림과 같이 레이가 객체를 통과할 때 객체의 3차원 위치 x=(x,y,z)에 대해서, 두 번째 그림과 같이 Q, K, V 행렬을 추출하고 양선형 보간법(bilinear interpolation)을 이용해 동일한 크기의 특징을 추출할 수 있다. 코어 텐서인 $\mathcal{G}_i$는 가우시안 표준정규분포를 따르는 랜덤값으로 설정할 수 있다.

[0096] 기존의 NeRF는 객체의 위치 정보인 (x,y,z)와 물체를 바라보는 방향(viewing direction, 카메라의 한 점과 각도) 정보인 (θ,φ)를 포함한 5차원 데이터를 입력으로 받아 RGB 값과 물체의 밀도를 예측해 (r,g,b,σ)를 출력하는 완전 연결 계층을 학습하는데, 총 9개의 레이어와 256개의 채널로 구성된다. NeRF는 학습 속도와 렌더링 속도가 느리고 렌더링의 성능이 좋지 않은 한계가 있으므로, 본 발명에서는 복셀 기반의 NeRF로 이러한 한계를 해결하고자 한다.

[0098] 한편, TensoRF(비특허문헌 0003)는 텐서 분해 기법(비특허문헌 0004)을 활용하여 텐서를 벡터와 행렬로 분해하고 렌더링에 사용하는 방법인 Vector-Matrix(VM) 분해를 제안하였다. TensoRF에서 제안하는 VM 분해는 분해된 벡터와 행렬에서 선형/양선형 보간법을 이용하여 특징을 계산하기 때문에 기존 복셀 기반의 방법들보다 적은 메모리를 사용하면서도 우수한 성능을 보였다. 따라서, 본 발명에서는 TensoRF를 기반으로 하여 NeRF의 계산량 문제를 해결하고 성능을 개선하는 것을 목표로 한다. 하지만 TensoRF는 벡터와 행렬을 사용하기 때문에 학습해야 할 파라미터가 많다는 단점이 있다. 따라서 본 발명에서는 벡터 없이 행렬만을 사용하는 블록-텀 분해를 적용하여 다양한 특징을 학습할 수 있는 새로운 모델을 제안한다. 블록-텀 분해는 고차원의 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하는 기법으로 다른 텐서 분해 기법들보다 상대적으로 적은 수의 파라미터로 목표 텐서에 근사할 수 있다는 장점이 있다.

[0100] 도 3은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)에서 사용되는 블록-텀 분해를 설명하기 위해 도시한 도면이다. 블록-텀 분해는 고차원 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하는 기법이다. N차원 텐서 $\mathcal{A} \in \mathbb{R}^{d_1 \times \cdots \times d_n}$는 P개의 블록으로 분해되고, 각각의 블록은 여러 개의 분해 기법 중에서도 CP 분해(비특허문헌 0005)나 터커 분해(비특허문헌 0006)로 표현할 수 있다. 블록-텀 분해는 다음과 같은 수학식 1로 정의될 수 있다.

수학식 1

[0102] $$\mathcal{A} = \sum_{i=1}^{P} \mathcal{G}_i \times_1 \mathcal{X}_i^{(1)} \times_2 \mathcal{X}_i^{(2)} \times_3 \cdots \times_n \mathcal{X}_i^{(n)}$$

[0104] 여기서, $\mathcal{X}$ 는 요인(factor) 행렬을 나타내며 n은 코어(core) 텐서 $\mathcal{G}$ 의 차원을 의미한다. 따라서 $\mathcal{X}$ 는 총 n개만큼 있는데, 본 발명에서는 3차원을 사용한다. 블록-텀 분해를 위해 P번의 터커 분해를 결합하는 방법을 사용할 수 있다.

[0106] 이하에서는, 도 1 내지 도 3을 참조하여, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 각 구성에 대해 상세히 설명하도록 한다.

[0108] 특징 추출부(110)는, 객체의 3차원 위치와 레이 정보를 이용해, 레이가 객체를 통과할 때 위치 x에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬에 해당하는 특징을 추출할 수 있다. 특징 추출부(110)는, 양선형 보간법을 이용해 Q, K 및 V에 해당하는 특징을 추출할 수 있다.

[0110] 도 4는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 특징 추출부(110)에서 추출하는 Q, K, V 행렬을 설명하기 위해 도시한 도면이다. 도 4의 좌측은 TensoRF에서 사용되는 텐서 분해를 나타낸 도면으로, 텐서를 벡터-행렬 외적의 결합(sum)으로 표현할 수 있다. 반면에, 도 4의 우측은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)에서 사용된 방법으로, 위치 x(녹색별 표시)에 대해 Q(X,Y), K(X,Z) 및 V(Y,Z) 행렬을 추출하고, 이를 이용해 양선형 보간법으로 동일한 크기의 특징을 추출함으로써 파라미터를 줄일 수 있다.

[0112] 블록-텀 분해부(120)는, 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해(Block-term decomposition)를 수행해 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합할 수 있다. 여기서, 블록은, 각각 CP 분해 또는 터커 분해로 표현될 수 있다.

[0114] 도 5는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)에서, 블록-텀 분해부(120)의 세부적인 구성을 도시한 도면이고, 도 6은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 블록-텀 분해부(120)가 결괏값을 얻는 과정을 도시한 도면이다. 도 5에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)의 블록-텀 분해부(120)는, 터커 분해부(121), 멀티헤드 어텐션부(122) 및 연결부(123)를 포함하여 구성될 수 있다.

[0116] 터커 분해부(121)는, Q, K 및 V 행렬을 이용해 터커(Tucker) 분해를 수행하되, P만큼 터커 분해를 반복 수행할 수 있다. 보다 구체적으로, 터커 분해부(121)는, 코어 텐서와 Q, K 및 V 행렬을 이용해 터커 분해를 수행할 수 있다. 이는 다음과 같은 수학식 2로 정의될 수 있다.

수학식 2

[0118]
$$\mathcal{X} \approx \mathcal{G} \times_1 Q \times_2 K \times_3 V = \sum_{a=1}^{A} \sum_{b=1}^{B} \sum_{c=1}^{C} g_{abc} q_a \circ k_b \circ v_c$$

[0120] 여기서, $\mathcal{G}$는 코어 텐서이고, a, b, c는 코어 텐서의 인덱스를 의미한다. 코어 텐서는, 가우시안 표준정규분포를 따르는 랜덤 값으로 설정될 수 있는데, 실험에서는 코어 텐서를 대각 텐서로 정의하였다.

[0122] 멀티헤드 어텐션부(122)는, 터커 분해를 한 결과 P개를 트랜스포머(비특허논문 0007)의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻을 수 있다. 멀티헤드 어텐션 과정은 다음 수학식 3과 같이 표현할 수 있다.

수학식 3

[0124] $$Head_p = Attention(\mathcal{G}_p; Q, K, V)$$

[0126] 연결부(123)는, P개의 결괏값을 결합할 수 있다. 즉, 터커 분해를 헤드의 코어 텐서의 개수 P개만큼 반복한 결괏값 $Head_p$를 모두 결합할 수 있다.

[0128] MLP 출력부(130)는, 결합한 결괏값을 MLP에 입력해 밀도와 색상으로 구성되는 출력값을 생성할 수 있다. 즉, 도 2에 도시된 바와 같이, 연결부(123)에서 결합된 결괏값을 모두 더해주고(Summation), Flatten 하여 MLP에 입력할 수 있다. MLP를 거치고 나면 출력값 (σ_i,c_i)을 생성할 수 있다.

[0130] 전술한 특징 추출부(110), 블록-텀 분해부(120) 및 MLP 출력부(130)의 실행은 모든 레이에 대해 반복적으로 적용되며, 생성된 각 출력값은 볼륨 렌더링을 통해 최종적으로 이미지 렌더링에 사용될 수 있다.

[0132] 객체 렌더링부(140)는, 특징 추출부(110), 블록-텀 분해부(120) 및 MLP 출력부(130)의 실행을 모든 레이에 대해 반복하고, 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링(Volume Rendering)을 통해 3차원 객체를 렌더링할 수 있다. 보다 구체적으로, 객체 렌더링부(140)는, 모든 레이에 대해 생성된 출력값을 이용해 각각의 레이에 해당하는 픽셀값을 계산하여, 볼륨 렌더링을 수행할 수 있다.

[0134] 이미지 렌더링을 위해, 객체 렌더링부(140)는, NeRF에서 사용하는 디퍼런셔블 볼륨 렌더링(Differentiable Volume Rendering: DVR)(비특허문헌 0008)을 사용해 렌더링할 수 있다. 각각의 레이에 해당하는 픽셀 값을 계산하는 수식은 다음 수학식 4 및 5와 같다.

수학식 4

[0136] $$C = \sum_{i=1}^{N} T_i(1 - \exp(-\sigma_i \delta_i))c_i$$

수학식 5

[0138] $$T_i = \exp\left(-\sum_{j=1}^{n-1} \sigma_j \delta_j\right)$$

[0140] 여기서, (σ_i, c_i)는 모델에서 생성된 출력값인데, 각각 밀도와 색상 값을 의미한다. T_i는 투명도, δ_i는 간격을 의미한다.

[0142] 도 7은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법의 흐름을 도시한 도면이다. 도 7에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법은, 컴퓨터에서 각 단계가 수행되는 3차원 객체 렌더링 방법으로서, 객체의 3차원 위치와 레이 정보를 이용해 Q, K 및 V 행렬에 해당하는 특징을 추출하는 단계(S110), 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해를 수행해 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 단계(S120), 결합한 결괏값을 MLP에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 단계(S130) 및 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링을 통해 3차원 객체를 렌더링하는 단계(S140)를 포함하여 구현될 수 있다.

[0144] 도 8은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법에서, 단계 S120의 세부적인 흐름을 도시한 도면이다. 도 8에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 방법의 단계 S120은, Q, K 및 V 행렬을 이용해 터커 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 단계(S121), 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 단계(S122) 및 P개의 결괏값을 결합하는 단계(S123)를 포함하여 구현될 수 있다.

[0146] 각각의 단계들과 관련된 상세한 내용들은, 앞서 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100)와 관련하여 충분히 설명되었으므로, 상세한 설명은 생략하기로 한다.

[0148] 실험

[0150] Synthetic-NeRF 데이터셋을 사용하여 모델 성능을 평가했다. 실험 세부 설정은 batch size는 4096, Optimizer는 Adam, Q, K, V의 Learning rate는 0.02, 그리고 코어 텐서와 MLP의 Learning rate는 0.001로 설정하였으며, 1개의 RTX3090 GPU 환경에서 실험을 수행하였다.

[0152] 도 9는 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100) 및 방법과 다른 방법에 의한 렌더링 수행 결과를 비교하여 나타낸 도면이다. 방법 명칭 옆의 “[숫자]”는 선행 비특허문헌의 일련번호를 나타낸다. Synthetic-NeRF 데이터셋에서, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100) 및 방법(“Ours”)을 기존 NeRF 모델들과 비교하였을 때, PSNR은 가장 성능이 좋은 모델인 TensoRF-VM 보다 약 1.51dB이 낮다. 하지만 SSIM은 0.006의 차이로 유사한 성능을 보였으며, 파라미터 크기는 약 6MB 줄어들었다. 따라서, 제안하는 모델이 효율적인 파라미터 크기를 가지며 성능 면에서도 다른 모델과 유사함을 입증했다.

[0154] 도 10은 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100) 및 방법에 의한 다양한 시점에서의 렌더링 결과는 타나낸 도면이다. 도 10에 도시된 바와 같이, 본 발명의 일실시예에 따른 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100) 및 방법은, Synthetic-NeRF 데이터셋에서 여러 객체에 대한 다양한 관점에서의 렌더링 결과, 복잡한 기하와 재료를 가진 레고 및 의자 카테고리에서 우수한 렌더링 결과를 보였다.

[0156] 이와 같이, 본 발명에서는 TensoRF와 블록-텀 분해를 결합하여 분해된 행렬 간의 관계성을 학습할 수 있는 새로운 모델을 제안하였다. 제안한 모델은 실험에서 파라미터 크기가 다른 복셀 기반 NeRF 모델들과 비교하였을 때 효율적이지만 성능은 유사함을 입증할 수 있었다.

[0158] 전술한 바와 같이, 본 발명에서 제안하고 있는 블록-텀 분해 기반의 신경 방사 필드 3차원 객체 렌더링 장치(100) 및 방법에 따르면, 복셀 기반의 NeRF를 이용하되, 고차원의 텐서를 낮은 차원의 텐서 블록들의 합으로 분해하되 벡터 없이 행렬만을 사용하는 블록-텀 분해를 적용함으로써, 기존의 NeRF보다 학습 속도와 렌더링 속도가 빠르고 렌더링 성능이 우수하며, 다른 텐서 분해 기법들보다 상대적으로 적은 수의 파라미터로 목표 텐서에 근사할 수 있어서 계산량을 줄이고 성능을 개선할 수 있다.

[0160] 한편, 본 발명은 다양한 통신 단말기로 구현되는 동작을 수행하기 위한 프로그램 명령을 포함하는 컴퓨터에서 판독 가능한 매체를 포함할 수 있다. 예를 들어, 컴퓨터에서 판독 가능한 매체는, 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 장치를 포함할 수 있다.

[0162] 이와 같은 컴퓨터에서 판독 가능한 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 이때, 컴퓨터에서 판독 가능한 매체에 기록되는 프로그램 명령은 본 발명을 구현하기 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 예를 들어, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해 실행될 수 있는 고급 언어 코드를 포함할 수 있다.

[0164] 전술한 본 발명의 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명된 각 구성요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명된 구성 요소들도 결합된 형태로 실시될 수 있다.

[0166] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.

부호의 설명

[0167] 100: 3차원 객체 렌더링 장치

110: 특징 추출부

120: 블록-텀 분해부

121: 터커 분해부

122: 멀티헤드 어텐션부

123: 연결부

130: MLP 출력부

140: 객체 렌더링부

S110: 객체의 3차원 위치와 레이 정보를 이용해 Q, K 및 V 행렬에 해당하는 특징을 추출하는 단계

S120: 추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해를 수행해 객체의 텐서를 P개의 블록으로 분해하고, 각 블록에 의한 결괏값 P개를 결합하는 단계

S121: Q, K 및 V 행렬을 이용해 터커 분해를 수행하되, P만큼 터커 분해를 반복 수행하는 단계

S122: 터커 분해를 한 결과 P개를 트랜스포머의 멀티헤드 어텐션과 같은 과정을 거쳐 P개의 결괏값을 얻는 단계

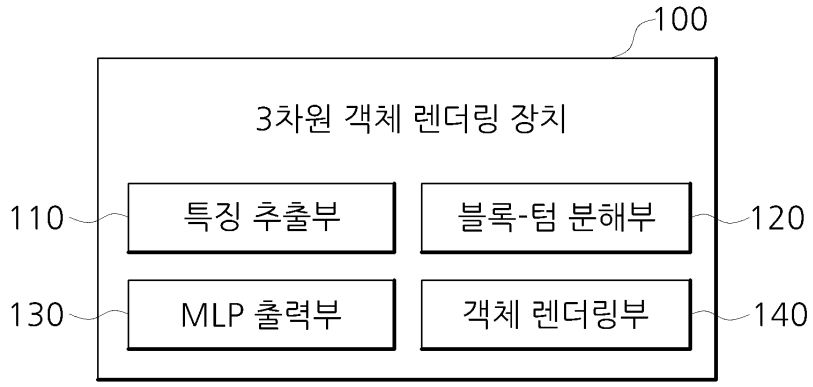
S123: P개의 결괏값을 결합하는 단계

S130: 결합한 결괏값을 MLP에 입력해 밀도와 색상으로 구성되는 출력값을 생성하는 단계

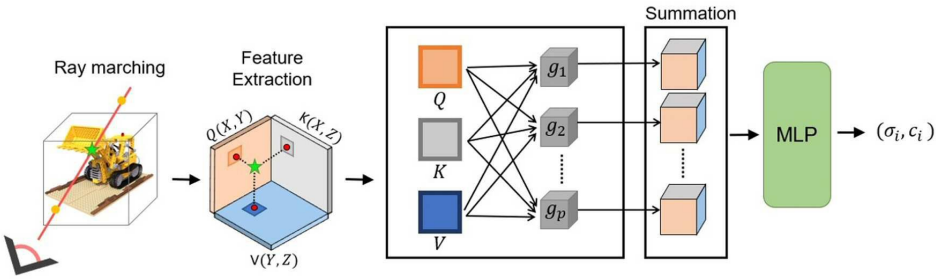
S140: 모든 레이에 대해 생성된 출력값을 이용해 볼륨 렌더링을 통해 3차원 객체를 렌더링하는 단계

도면

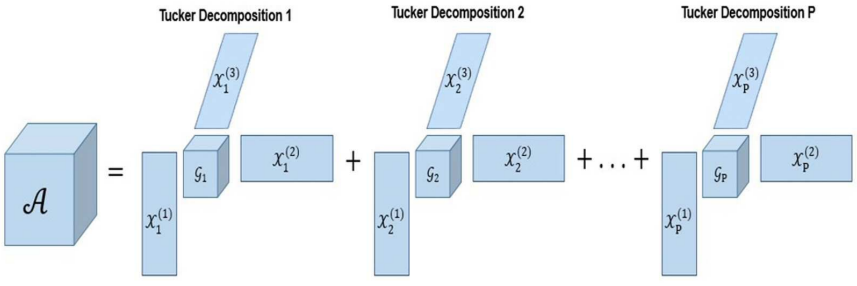
도면1



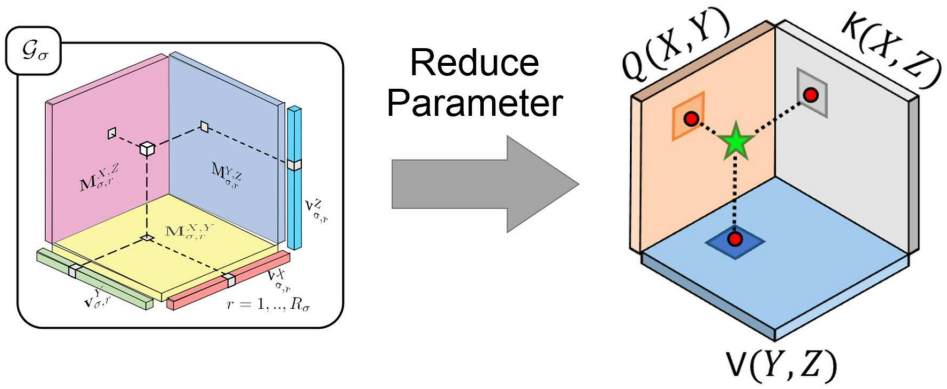
도면2



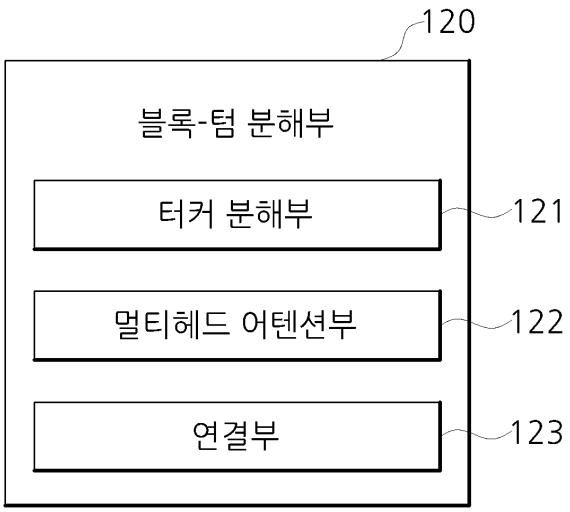
도면3



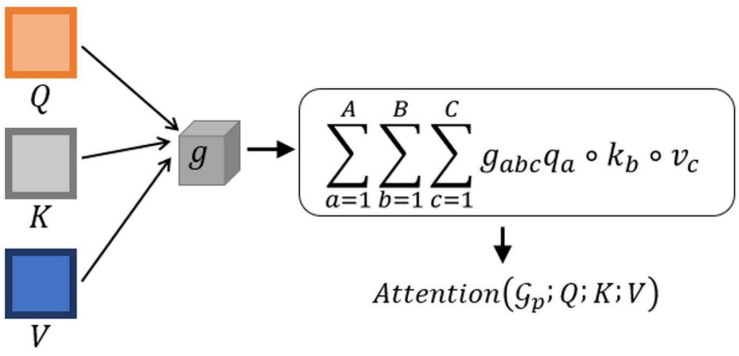
도면4


Reduce
Parameter
Q(X,Y)
K(X,Z)
V(Y,Z)

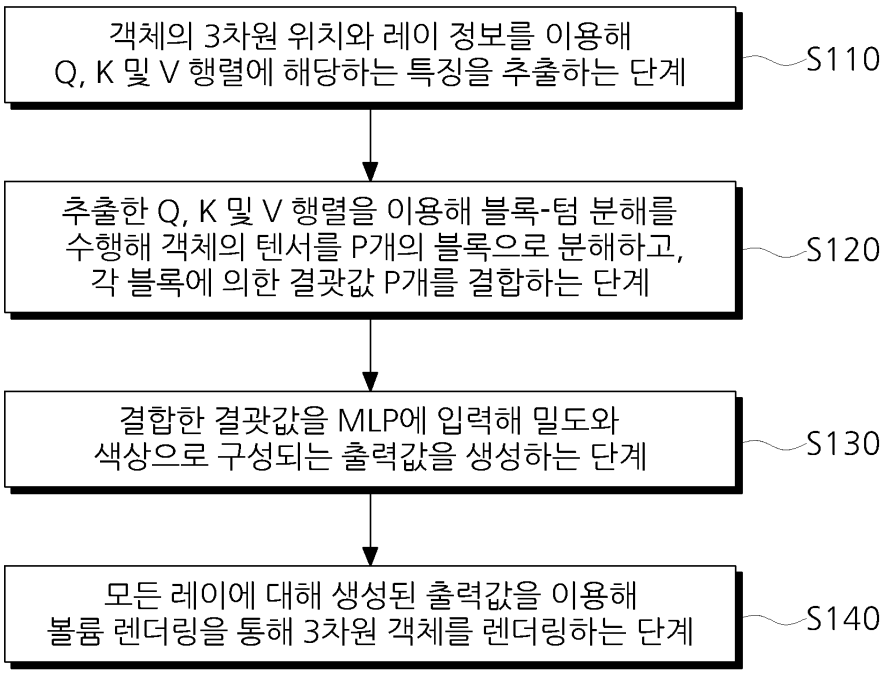
도면5


120
블록-텀 분해부
터커 분해부
121
멀티헤드 어텐션부
122
연결부
123

도면6

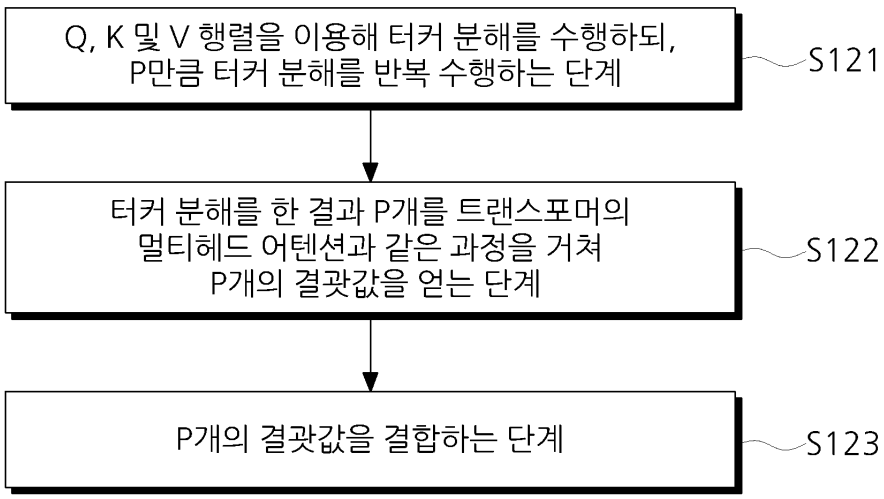

Q
K
V
g
$\sum_{a=1}^{A}\sum_{b=1}^{B}\sum_{c=1}^{C} g_{abc}q_a \circ k_b \circ v_c$
$Attention(\mathcal{G}_p; Q; K; V)$

도면7

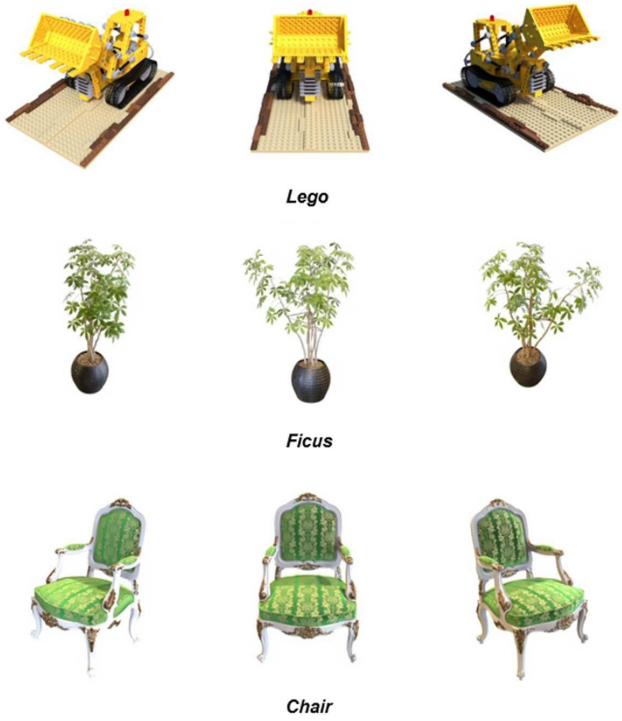

객체의 3차원 위치와 레이 정보를 이용해
Q, K 및 V 행렬에 해당하는 특징을 추출하는 단계
S110
추출한 Q, K 및 V 행렬을 이용해 블록-텀 분해를
수행해 객체의 텐서를 P개의 블록으로 분해하고,
각 블록에 의한 결괏값 P개를 결합하는 단계
S120
결합한 결괏값을 MLP에 입력해 밀도와
색상으로 구성되는 출력값을 생성하는 단계
S130
모든 레이에 대해 생성된 출력값을 이용해
볼륨 렌더링을 통해 3차원 객체를 렌더링하는 단계
S140

도면8

S120


Q, K 및 V 행렬을 이용해 터커 분해를 수행하되,
P만큼 터커 분해를 반복 수행하는 단계
S121
터커 분해를 한 결과 P개를 트랜스포머의
멀티헤드 어텐션과 같은 과정을 거쳐
P개의 결괏값을 얻는 단계
S122
P개의 결괏값을 결합하는 단계
S123

도면9

Method	Synthetic-NeRF		
	PSNR ↑	SSIM ↑	Size(MB) ↓
SRN[9]	22.26	0.846	-
NSVF[10]	31.75	0.953	-
NeRF[1]	31.01	0.947	5.00
SNeRG[11]	30.38	0.950	1771.5
PlenOctrees[12]	31.71	0.958	1976.3
Plenoxels[13]	31.71	0.958	778.1
DVGO[14]	31.95	0.957	612.1
TensoRF-CP[2]	31.56	0.949	3.9
TensoRF-VM[2]	33.14	0.963	71.8
Ours	**31.63**	**0.957**	**65.8**

도면10