



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2014-0107973
(43) 공개일자 2014년09월05일

(51) 국제특허분류(Int. Cl.)

H04N 13/00 (2006.01)

(21) 출원번호 10-2013-0022279

(22) 출원일자 2013년02월28일

심사청구일자 없음

(71) 출원인

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

설상훈

서울 강남구 언주로 117, 8동 402호 (도곡동, 우성4차아파트)

이훈재

경기 수원시 권선구 금곡로31번길 7, 104동 102호 (금곡동, LG빌리지)

박한제

경기 군포시 광정로 119, 733동 503호 (산본동, 대림솔거아파트)

(74) 대리인

특허법인 무한

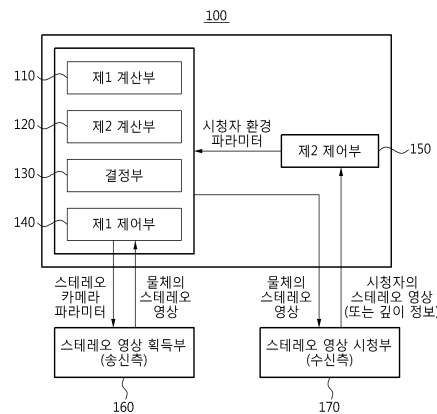
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 스테레오 영상 획득 및 시청 연동 시스템

(57) 요약

3D 영상으로 획득 할 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산하는 제1 계산부, 상기 3D 영상을 제공 받는 수신측에 연관되는 적어도 하나의 제2 파라미터를 이용하여, 상기 제1 포인트에 대응하여 상기 수신측에서 제공되는 적어도 하나의 제2 포인트의 제2 위치를 계산하는 제2 계산부, 및 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록, 상기 3D 영상을 획득하여 상기 수신측에 제공할 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정하는 결정부를 포함하는 영상 처리 장치가 제공된다.

대표도 - 도1



특허청구의 범위

청구항 1

3D 영상으로 획득 할 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산하는 제1 계산부;

상기 3D 영상을 제공 받는 수신측에 연관되는 적어도 하나의 제2 파라미터를 이용하여, 상기 제1 포인트에 대응하여 제공되는 적어도 하나의 상기 수신측에서의 제2 포인트의 제2 위치를 계산하는 제2 계산부; 및

상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록, 상기 3D 영상을 획득하여 상기 수신측에 제공할 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정하는 결정부

를 포함하는 영상 처리 장치.

청구항 2

제1항에 있어서,

상기 제1 위치 및 상기 제2 위치 중 적어도 하나는, 3D 공간상의 기준 위치에 대한 상대적인 위치인 영상 처리 장치.

청구항 3

제1항에 있어서,

상기 제1 파라미터는, 상기 송신측에 연관되는 베이스 라인, 초점 거리, 주시각 및 가상 베이스 라인 중 적어도 하나를 포함하는 영상 처리 장치.

청구항 4

제1항에 있어서,

상기 제2 파라미터는, 상기 수신측에 연관되는 화면 크기, 시청 거리, 시청자의 양안 사이의 거리 및 시청자의 위치 중 적어도 하나를 포함하는 영상 처리 장치.

청구항 5

제1항에 있어서,

상기 적어도 하나의 제1 파라미터에 기초하여 상기 송신측에 연관되는 카메라를 조정하여 상기 3D 영상을 획득하는 제1 제어부

를 더 포함하는 영상 처리 장치.

청구항 6

제1항에 있어서,

상기 제2 파라미터를 상기 수신측으로부터 전달 받아 상기 제2 계산부에 전달하는 제2 제어부

를 더 포함하는 영상 처리 장치.

청구항 7

제1항에 있어서,

상기 수신측으로부터 전달받은 스테레오 영상 및 깊이 정보 중 적어도 하나를 이용하여 상기 제2 파라미터를 측정하여 상기 제2 계산부에 전달하는 제2 제어부

를 더 포함하는 영상 처리 장치.

청구항 8

제1항에 있어서,

상기 결정부는 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록 하는 목적 함수의 해를 구하여 상기 제1 파라미터를 결정하는 영상 처리 장치.

청구항 9

제8항에 있어서,

상기 결정부는 상기 샘플링되는 적어도 하나의 제1 포인트의 개수가 상기 제1 파라미터 개수와 상기 제2 파라미터의 개수의 합보다 큰 경우 일부의 제1 포인트를 선택하여 상기 목적 함수의 해를 구하는 영상 처리 장치.

청구항 10

제9항에 있어서,

상기 결정부는 상기 선택 시 이상치(outlier)를 제외하고 선택하는 영상 처리 장치.

청구항 11

제1항에 있어서,

상기 제2 계산부는 상기 제2 위치를 계산하는 경우 기하학적 영상 보상 방법에 기초하여 상기 송신측에 연관되는 카메라의 주시각으로부터 발생하는 왜곡이 감소하도록 상기 제2 위치를 계산하는 영상 처리 장치.

청구항 12

제1항에 있어서,

상기 결정부가, 디스패리티 제약 항 및 파라미터 변화 제약 항 중 적어도 하나를 목적 함수에 추가하고 해를 구하여 상기 제1 파라미터를 결정하는 영상 처리 장치.

청구항 13

영상 처리 장치가 영상 처리를 수행하는 방법에 있어서,

3D 영상으로 획득 할 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산하는 단계;

상기 3D 영상을 제공 받는 수신측에 연관되는 적어도 하나의 제2 파라미터를 이용하여, 상기 제1 포인트에 대응하여 제공되는 적어도 하나의 상기 수신측에서의 제2 포인트의 제2 위치를 계산하는 단계; 및

상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록, 상기 3D 영상을 획득하여 상기 수신측에 제공할 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정하는 단계

를 포함하는 영상 처리 방법.

청구항 14

제13항에 있어서,

상기 제1 위치 및 상기 제2 위치 중 적어도 하나는, 3D 공간상의 기준 위치에 대한 상대적인 위치인 영상 처리 방법.

청구항 15

제13항에 있어서,

상기 적어도 하나의 제1 파라미터는, 상기 송신측에 연관되는 베이스 라인, 초점 거리, 주시각 및 가상 베이스 라인 중 적어도 하나를 포함하는 영상 처리 방법.

청구항 16

제13항에 있어서,

상기 적어도 하나의 제2 파라미터는, 상기 수신측에 연관되는 화면 크기, 시청 거리, 시청자의 양안 사이의 거리 및 시청자의 위치 중 적어도 하나를 포함하는 영상 처리 방법.

청구항 17

제13항에 있어서,

상기 적어도 하나의 제1 파라미터에 기초하여 상기 송신측에 연관되는 카메라를 조정하여 상기 3D 영상을 획득하는 단계

를 더 포함하는 영상 처리 방법.

청구항 18

제13항에 있어서,

상기 수신측으로부터 전달받은 스테레오 영상 및 깊이 정보 중 적어도 하나를 이용하여 상기 제2 파라미터를 측정하는 단계

를 더 포함하는 영상 처리 방법.

청구항 19

제13항에 있어서,

상기 적어도 하나의 제1 파라미터를 결정하는 단계는, 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록 하는 목적 함수의 해를 구하여 상기 적어도 하나의 제1 파라미터를 결정하는 영상 처리 방법.

청구항 20

제13항 내지 제19항 중 어느 한 항의 영상 처리 방법을 수행하는 프로그램을 수록한 컴퓨터 판독 가능 기록 매체.

명세서

기술분야

[0001] 스테레오 영상을 통한 실감 커뮤니케이션 (realistically communicating through stereo images)에 연관된다. 이를테면, 3D (3-Dimensional) 화상 통화, 환자의 환부를 3D 영상으로 제공받아 진행하는 의료 행위, 원격 폭발물 제거, 원격 쇼핑, 원격 기기 제어 등에 전반적으로 연관되나, 이에 한정되는 것은 아니다.

배경기술

[0002] 최근 3D TV 및 3D 카메라와 같은 3D 산업이 성장하고 있다. 이에 따라, 3D 스테레오 영상을 이용하는 내비게이션 및 의료 수술에 연관되는 연구들이 활성화되고 있으며, 상기 3D 스테레오 영상을 이용하는 분야에서 중요한 이슈 중 하나는 상기 3D 스테레오 영상을 시청하는 시청자 (viewer)에게 실감나는 3D 지각을 제공하는 것이다. 여기서, 3D 물체의 실감 원격 3D 지각 (a realistic remote 3D perception of a 3D object)은, 실제 3D 물체와 3D 모양 및/또는 크기가 같은 3D 물체를 지각할 수 있는 시각 능력을 의미한다.

[0003] 종래의 3D 스테레오 영상을 획득하는 장치 및 방법은 상기 시청자가 지각하는 3D 물체와 실제 3D 물체의 모양 및/또는 크기에서 차이가 발생하고, 이러한 차이로 인해 상기 시청자에게 실감나는 3D 지각을 제공할 수 없다.

[0004] 따라서, 상기 시청자에게 실감나는 3D 지각을 제공하기 위해 상기 시청자가 지각하는 3D 물체와 실제 3D 물체의 모양 및/또는 크기 사이의 차이를 최소화하는 방법이 요구된다.

[0005] 실감 원격 3D 지각을 위한 종래의 방법들은, 이를테면, 디스페리티 필드 예측에 기반하여 물체의 3D 재구성을 수행하고 3D 공간상에서 시청자가 지각하는 3D 물체를 조절하여 새로운 스테레오 영상을 합성하는 것이다. 이러한 기술은 N. Chang, and A. Zakhor, "View generation for three-dimensional scenes from video sequences," IEEE Trans. Image Process., vol. 6, no. 4, pp. 584-598, Apr. 1997. 및 R. Vasudevan, G.

Kurillo, E. Lobaton, T. Bernardin, O. Kreylos, R. Bajcsy, and K. Nahrstedt, "High-quality visualization for geographically distributed 3-D teleimmersive applications," IEEE Trans. Multimedia, vol. 13, no. 3, pp. 573-584, Jun. 2011. 에서 제시된 바 있다.

- [0006] 그러나, 디스패리티 필드 계산 자체에서 정확성이 높지 않아, 3D 점들의 재구성 정확성이 그리 높지 않을 수 있고, 따라서 불완전하게 재구성된 3D 물체들로부터 합성된 결과 스테레오 영상 또한, 실제 스테레오 영상 획득 결과에 비해 불완전할 수 있다.
- [0007] 현재의 3D TV 또는 3D 스마트폰과 같은 상용 3D 스테레오 디스플레이에서 널리 이용되는 3D깊이 조절 방법으로 parallax 조절 방법이 있다. 하지만, parallax 조절 방법은, 3D 스테레오 영상 전체 또는 일부의 수평 디스패리티를 같은 양만큼 조절하기 때문에, 시각적 피로가 발생하고 또한 사용자가 지각하는 3D 물체의 모양이 왜곡된다.
- [0008] 또한, F. Zilly, J. Kluger, and P. Kauff, "Production rules for stereo acquisition," Proc. IEEE, vol. 99, no. 44, pp. 590-606, Apr. 2011을 비롯한 다수의 기존 문헌에서는 시청 피로도를 저감하기 위한 목적에서 과도한 디스패리티를 줄이기 위해 영상 획득 시 스테레오 카메라 파라미터를 조절하는 방법이 제시된 바 있다.
- [0009] 그러나 의료적 응용, 정밀 기기 제어, 화상 회의, 원격 쇼핑 등, 3D 영상이 더 다양하고 위험성을 수반하는 분야에 활용됨에 따라, 상기 시청피로도 저감 이상으로 3D 물체의 모양 및/또는 크기의 왜곡 (distortion)을 없애는 것이 중요해지고 있다.

발명의 내용

과제의 해결 수단

- [0010] 일측에 따르면, 3D 영상으로 획득 할 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산하는 제1 계산부; 상기 3D 영상을 제공 받는 수신측에 연관되는 적어도 하나의 제2 파라미터를 이용하여, 상기 제1 포인트에 대응하여 제공되는 적어도 하나의 상기 수신측에서의 제2 포인트의 제2 위치를 계산하는 제2 계산부; 및 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록, 상기 3D 영상을 획득하여 상기 수신측에 제공할 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정하는 결정부를 포함하는 영상 처리 장치가 제공된다.
- [0011] 일실시예에 따르면, 상기 제1 위치 및 상기 제2 위치 중 적어도 하나는, 3D 공간상의 기준 위치에 대한 상대적인 위치일 수 있다.
- [0012] 일실시예에 따르면, 상기 제1 파라미터는, 상기 송신측에 연관되는 베이스 라인, 초점 거리, 주시각 및 가상 베이스 라인 중 적어도 하나를 포함할 수 있다.
- [0013] 일실시예에 따르면, 상기 제2 파라미터는, 상기 수신측에 연관되는 화면 크기, 시청 거리, 시청자의 양안 사이의 거리 및 시청자의 위치 중 적어도 하나를 포함할 수 있다.
- [0014] 한편, 일실시예에 따르면 상기 영상 처리 장치는 상기 적어도 하나의 제1 파라미터에 기초하여 상기 송신측에 연관되는 카메라를 조정하여 상기 3D 영상을 획득하는 제1 제어부를 더 포함할 수 있다.
- [0015] 또한, 일실시예에 따르면 상기 영상 처리 장치는 상기 제2 파라미터를 상기 수신측으로부터 전달 받아 상기 제2 계산부에 전달하는 제2 제어부를 더 포함할 수도 있다.
- [0016] 나아가, 일실시예에 따르면 상기 영상 처리 장치는, 상기 수신측으로부터 전달받은 스테레오 영상 및 깊이 정보 중 적어도 하나를 이용하여 상기 제2 파라미터를 측정하여 상기 제2 계산부에 전달하는 제2 제어부를 더 포함할 수도 있다.
- [0017] 일실시예에 따르면, 상기 결정부는 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록 하는 목적 함수의 해를 구하여 상기 제1 파라미터를 결정할 수 있다.
- [0018] 일실시예에 따르면, 상기 결정부는 상기 샘플링되는 적어도 하나의 제1 포인트의 개수가 상기 제1 파라미터 개수와 상기 제2 파라미터의 개수의 합보다 큰 경우 일부의 제1 포인트를 선택하여 상기 목적 함수의 해를 구할 수 있다.
- [0019] 이 경우, 상기 결정부는 상기 선택 시 이상치(outlier)를 제외하고 선택할 수도 있다.

- [0020] 일실시예에 따르면, 상기 제2 계산부는 상기 제2 위치를 계산하는 경우 기하학적 영상 보상 방법에 기초하여 상기 송신측에 연관되는 카메라의 주시각으로부터 발생하는 왜곡이 감소하도록 상기 제2 위치를 계산할 수 있다.
- [0021] 일실시예에 따르면, 상기 결정부는, 디스페리티 제약 항 및 파라미터 변화 제약 항 중 적어도 하나를 목적 함수에 추가하고 해를 구하여 상기 제1 파라미터를 결정할 수 있다.
- [0022] 다른 일측에 따르면, 영상 처리 장치가 영상 처리를 수행하는 방법에 있어서, 3D 영상으로 획득 할 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산하는 단계; 상기 3D 영상을 제공 받는 수신측에 연관되는 적어도 하나의 제2 파라미터를 이용하여, 상기 제1 포인트에 대응하여 제공되는 적어도 하나의 상기 수신측에서의 제2 포인트의 제2 위치를 계산하는 단계; 및 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록, 상기 3D 영상을 획득하여 상기 수신측에 제공할 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정하는 단계를 포함하는 영상 처리 방법이 제공된다.

도면의 간단한 설명

- [0023] 도 1은 일실시예에 따른 영상 처리 장치의 블록도이다.
- 도 2는 일실시예에 따른 상기 영상 처리 장치를 포함하는 송신측 및 수신측의 실시예를 도시한다.
- 도 3은 일실시예에 따른, 스테레오 영상을 획득하는 송신측 및 스테레오 영상을 시청하는 수신측의 연동 모델을 나타내는 도면이다.
- 도 4는 일실시예에 따른, 블록 디스페리티의 예측을 나타내는 도면이다.
- 도 5는 일실시예에 따른, 실제 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 3D 좌표 (제1 위치) 예측을 나타내는 도면이다.
- 도 6은 일실시예에 따른, 상기 스테레오 영상을 획득하는 송신측과 연관된 카메라 파라미터에 대해 시청자가 지각하는 3D 물체에서의 제2 포인트의 3D 좌표 (제2 위치) 계산을 나타내는 도면이다.
- 도 7은 일실시예에 따른, 주시각이 있는 스테레오 카메라를 이용하는 스테레오 영상 획득을 설명하는 도면이다.
- 도 8은 일실시예에 따른 영상 처리 방법의 흐름도이다.

발명을 실시하기 위한 구체적인 내용

- [0024] 이하에서, 일부 실시예들을, 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0025] 아래 설명에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어를 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다.
- [0026] 또한 특정한 경우는 이해를 돕거나 및/또는 설명의 편의를 위해 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 설명 부분에서 상세한 그 의미를 기재할 것이다. 따라서 아래 설명에서 사용되는 용어는 단순한 용어의 명칭이 아닌 그 용어가 가지는 의미와 명세서 전반에 걸친 내용을 토대로 이해되어야 한다.
- [0027] 도 1은 일실시예에 따른 영상 처리 장치(100)의 블록도이다. 스테레오 영상을 통해 시청자가 지각하는 3D 물체의 모양 및/또는 크기는 스테레오 카메라의 내부와 외부의 파라미터, 3D 스테레오 디스플레이 스크린의 크기, 시청 거리 등과 같은 다양한 스테레오 카메라 및 시청자 환경 파라미터에 의해 영향을 받을 수 있다.
- [0028] 따라서, 상기 스테레오 영상을 통해 3D 영상을 제공하기 위한 영상 처리 장치(100)는 상기 시청자가 지각하는 3D 물체의 모양 및/또는 크기가 실제 3D 물체의 모양 및/또는 크기로 유지되도록 할 수 있다. 이를 위해, 상기 영상 처리 장치(100)는 상기 실제 3D 물체에서 샘플링된 포인트들 (이하에서 제1 포인트라고 함)의 위치들 (이하에서, 제1 위치라고 함)과 제1 포인트에 대응되는 상기 시청자가 지각하는 3D 물체에서의 포인트들 (이하에서 제2 포인트라고 함)의 위치들 (이하에서, 제2 위치라고 함)의 차이를 최소화 하는 최적의 스테레오 카메라 파라미터 (이하에서, 제1 파라미터라고 함)들을 계산할 수 있다.
- [0029] 일실시예에 따르면, 상기 영상 처리 장치(100)는 제1 계산부(110), 제2 계산부(120), 결정부(130) 제1 제어부(140) 및 제2 제어부(150)를 포함할 수 있다. 상기 제1 계산부(100)는 상기 실제 3D 물체로부터 샘플링되는 적

어도 하나의 제1포인트의 제1 위치를 계산할 수 있다.

- [0030] 상기 영상 처리 장치(100)의 상기 제2 계산부(120)는 수신측(170)에 연관되는 적어도 하나의 시청자 환경 파라미터 (이하에서, 제2 파라미터라고 함)를 이용하여 상기 시청자가 지각하는 3D 물체에서 적어도 하나의 제1 포인트에 대응하는 제2 포인트의 제2 위치를 계산할 수 있다. 여기서, 상기 수신측(170)은 송신측(160)에서 획득되는 상기 실제 3D 물체의 스테레오 영상을 수신하여 상기 시청자에게 디스플레이할 수 있는 3D 스테레오 시청 시스템을 의미할 수 있으며, 이를 테면, 스크린, 디스플레이 장치를 포함할 수 있다. 또한, 상기 송신측(160)은 상기 실제 3D 물체의 스테레오 영상을 획득할 수 있는 이를 테면, 3D 스테레오 카메라를 포함할 수 있다.
- [0031] 상기 영상 처리 장치(100)의 상기 제1 제어부(140)는, 상기 최적의 스테레오 카메라 파라미터 (제1 파라미터)가 결정되면 송신측(160)에 연관된 스테레오 카메라 파라미터를 조절하여 상기 실제 3D 물체의 스테레오 영상을 획득할 수 있다.
- [0032] 상기 영상 처리 장치(100)의 상기 제2 제어부(150)는 수신측(170)에서의 화면 크기, 시청 거리, 상기 시청자의 두 눈 사이의 거리, 시청자의 위치 등을 포함할 수 있는 시청자 환경 파라미터 (제2 파라미터)를 수신하여 제2 계산부(120)로 전달할 수 있다.
- [0033] 또한, 상기 제2 제어부(150)는 스테레오 영상 카메라, 빛/또는 적외선 등을 이용한 깊이 센서 (depth sensor)에 기반하고, 얼굴 검출 (face detection) 및/또는 눈 검출 (eye detection) 기술 등을 이용하여, 상기 시청 거리 또는 상기 시청자의 두 눈 사이의 거리, 상기 시청자 위치 등과 같은 상기 시청자 환경 파라미터 (상기 제2 파라미터)를 측정할 수도 있다.
- [0034] 나아가 다른 실시예에서는, 상기 제2 제어부(150)는 상기 시청자 환경 파라미터 (상기 제2 파라미터)를 측정하지 않고, 상기 시청자 환경 파라미터의 기본값을 제2 계산부(120)로 전달할 수도 있다. 이를테면, 제2 제어부(150)는 상기 시청자의 두 눈 사이의 거리 정보가 수신되지 않고, 직접 측정되기 어렵거나, 기본 값을 사용하도록 하는 사용자의 지시가 있거나 및/또는 여하간의 이유로 기본 값을 사용하는 것이 적절하다고 판단되는 경우에, 상기 시청자의 두 눈 사이의 거리 기본 값인 75mm를 상기 제2 계산부(120)에 전달할 수 있다.
- [0035] 일실시예에 따르면, 상기 영상 처리 장치(100)의 상기 결정부(130)는 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록 상기 송신측에 연관되는 적어도 하나의 제1 파라미터를 결정할 수 있다. 또한, 상기 제1 파라미터를 결정하는 동안, 스테레오 카메라의 주시각으로부터 발생하는 상기 시청자가 지각하는 3D 물체의 왜곡을 줄이기 위한 기하학적 영상 보상이 수행될 수 있다. 이러한 왜곡은 평면 심도 왜곡으로 알려져 있다. 상기 기하학적 영상 보상의 수행에 대해서는 아래에서 도면을 참조하여 보다 상세하게 후술된다.
- [0036] 일실시예에 따르면, 상기 제1 파라미터는 상기 수신측(170)에 제공할 3D 영상을 획득하는 카메라에 연관되는 파라미터일 수 있다. 상기 제1 파라미터는 상기 송신측(160)에 연관되는 베이스 라인, 초점 거리, 주시각 및 가상 베이스 라인 중 적어도 하나를 포함할 수 있다.
- [0037] 일실시예에 따르면, 상기 제2 파라미터는 상기 시청자에게 3D 영상을 디스플레이하는 시청자 환경에 연관되는 파라미터일 수 있다. 이를 테면, 상기 제2 파라미터는 상기 시청자가 지각하는 3D 물체의 모양 및/또는 크기에 영향을 줄 수 있는 상기 수신측(170)에 연관되는 화면 크기 (스크린의 크기), 시청 거리, 상기 수신측에 연관되는 시청자의 눈 사이의 거리 및 시청자의 위치 중 적어도 하나를 포함할 수 있다.
- [0038] 일실시예에 따르면, 상기 영상 처리 장치(100)의 상기 결정부(130)는 상기 제1 위치와 상기 제2 위치의 차이가 최소화 되도록 하는 목적 함수의 해를 구하여 상기 송신측(160)에 연관되는 적어도 하나의 제1 파라미터를 결정할 수 있다. 상기 적어도 하나의 제1 파라미터는 상기 송신측에 포함되는 상기 스테레오 카메라에 연관되는 파라미터일 수 있다. 따라서, 상기 결정부(130)에 의해 결정되는 상기 제1 파라미터는 최적의 스테레오 카메라 파라미터일 수 있다.
- [0039] 상기 스테레오 영상을 통해 상기 시청자가 지각하는 3D 물체의 모양 및/또는 크기는 스테레오 카메라의 내부 또는 외부의 파라미터, 3D 스테레오 디스플레이 스크린의 크기 (상기 화면 크기), 시청 거리와 같은 다양한 스테레오 카메라 및 시청자 환경 파라미터에 의해 영향을 받는다. 따라서, 상기 스테레오 카메라에 연관되는 상기 제1 파라미터를 조정함으로써 실감나는 3D 지각을 제공할 수 있다. 위에 기술된 내용은, 상기 스테레오 카메라에 연관되는 상기 제1 파라미터에 연관되는 실시예들만 기술되나, 이에 제한되지 않고 경우에 따라서는 상기 사용자 환경 파라미터 (상기 제2 파라미터)를 조정함으로써 상기 시청자에게 실감나는 3D 지각을 제공할 수 있는 것으로 이해되어야 한다.

- [0040] 도 2는 상기 영상 처리 장치(100)를 포함하는 상기 송신측 및 수신측의 실시예를 도시한다. 도 2의 (a)는, 상기 스테레오 카메라에 연관되는 상기 제1 파라미터를 이용하여 획득한 스테레오 영상을 상기 수신측에 연관되는 시청자에게 전달하고 이를 상기 시청자가 3D 디스플레이 스크린을 통해 보는 예를 나타낸다.
- [0041] 도 2의 (b)는 상기 영상 처리 장치(100)를 이용하는 화상 통화의 실시예를 도시한다. 그러나, 상기 화상 통화에 이용되는 것으로 제한되지 않는다. 또한, 앞서 기술된 상기 송신측은 상기 스테레오 카메라와 같은 스테레오 영상을 촬영하는 카메라를 포함할 수 있으나 이에 제한되지 않으며 반대로 상기 수신측에서도 상기 송신측과 동일하게 상기 스테레오 카메라를 포함할 수 있다. 그러므로, 상기 영상 처리 장치(100)를 포함하는 상기 송신측이 상기 수신측일 수 있으며, 반대로 상기 수신측이 상기 송신측일 수 있다.
- [0042] 일실시예에 따르면, 상기 송신측 및 상기 수신측 양방향 모두 3D 물체로부터 샘플링되는 적어도 하나의 제1 포인트의 제1 위치를 계산할 수 있으며, 상대방의 시청자 환경에 연관되는 적어도 하나의 파라미터 (이를 테면, 제2 파라미터)를 이용하여 상기 상대방에게 지각되는 3D 물체 내에서 상기 적어도 하나의 제2 포인트에 대응하는 제2 위치를 계산할 수 있다.
- [0043] 또한, 상기 제1 위치 및 상기 제2 위치의 차이가 최소화 되도록 하는 목적 함수의 해를 구하여 상기 상대방의 카메라에 연관되는 적어도 하나의 파라미터를 결정함으로써, 서로에게 최적의 스테레오 카메라 파라미터를 제공할 수 있다. 이와 같이, 상기 송신측 및 상기 수신측 상관없이 상기 영상 처리 장치(100)를 포함할 수 있으나, 여기서는 설명의 편의를 위해 상기 송신측과 상기 수신측 두 가지의 경우를 구분하여 설명하도록 한다.
- [0044] 도 3은 일실시예에 따른, 스테레오 영상을 획득하는 송신측 및 스테레오 영상을 시청하는 수신측의 연동 모델을 나타내는 도면이다. 도 3의 (a)는 상기 3D 물체의 스테레오 영상을 획득하는 송신측을 도시하고, 도 3의 (b)는 상기 3D 물체의 스테레오 영상을 수신하는 수신측(이를 테면, 시청자 환경)을 도시한다. 일실시예에 따르면, 도 3의 (a)에서 도시되는 상기 송신측에서 이를 테면, 스테레오 카메라를 이용하여 3D 물체의 스테레오 영상을 획득할 수 있다. 또한, 상기 획득된 3D 물체의 스테레오 영상은 도 3의 (b)에 도시된 상기 수신측으로 전달되어 디스플레이 스크린(340)에 디스플레이 될 수 있다.
- [0045] 일실시예에 따르면, 상기 디스플레이 스크린(340)에 디스플레이 되는 경우, 상기 시청자가 지각하는 3D 물체에서의 포인트(302)는 상기 송신측의 스테레오 카메라에 의해 촬영되는 실제 3D 물체에서의 포인트(301)에 대응되는 포인트일 수 있다. 도 3에서 $x^{(L)}$ 및 $x^{(R)}$ 은 상기 포인트(301)에 대응되는 좌측 및 우측 영상에서의 2D 포인트를 각각 나타낸다. 여기서, 상기 송신측과 상기 수신측의 원점(O)은 좌측 및 우측 카메라 ($C^{(L)}$ (310) 및 $C^{(R)}$ (320))의 중심 포인트와 상기 시청자의 좌측 및 우측의 눈($E^{(L)}$ (350) 및 $E^{(R)}$ (360))의 중심 포인트(가운데 포인트)로 가정한다.
- [0046] 또한, 상기 수신측의 원점(O) 및 상기 디스플레이 스크린(340)의 중심 포인트도 z방향으로 정렬 (align)되어 있다고 가정한다. 하지만, 이에 제한되지 않고, 상기 수신측의 원점을 다른 위치로 설정할 수 있다.
- [0047] 일실시예에 따르면, 상기 수신측은, 이를 테면, 스테레오 카메라 (또는, 3D 깊이 센서) (330)를 이용하여 상기 수신측에 연관되는 파라미터(상기 제2 파라미터)를 획득할 수 있다. 또한, 상기 송신측은 상기 수신측에 연관되는 상기 제2 파라미터의 값을 다양하게 응용할 수 있다. 일실시예에 따르면, 상기 송신측이 상기 스테레오 영상을 획득하는 동안의 상기 수신측의 시청자 환경 파라미터(상기 제2 파라미터)를 알 수 있다고 가정할 수 있다.
- [0048] 상기 영상 처리 장치(100)는 상기 스테레오 영상을 획득하는 동안, 상기 실제 3D 물체의 깊이를 대략적으로 알고 있는 상태에서 상기 수신측에 연관되는 상기 제2 파라미터를 이용하여 상기 송신측의 스테레오 카메라의 최적의 파라미터를 예측할 수 있다. 이 경우, 상기 실제 3D 물체의 깊이를 대략적으로 알기 위해 상기 제1 계산부(110)는 상기 실제 3D 물체의 샘플링된 적어도 하나의 상기 제1 포인트의 상기 제1 위치를 계산할 수 있다. 여기서, 스테레오 카메라에 연관되는 상기 제1 파라미터를 획득하기 위해, 상기 실제 3D 물체의 상기 제1 포인트의 상기 제1 위치를 계산하더라도, 결과의 스테레오 영상은 상기 실제 3D 물체의 재구성된 점들로부터 합성되는 것은 아니며, 상기 영상 처리 장치(100)에 의해 결정된 상기 스테레오 카메라에 연관되는 제1 파라미터 값으로 스테레오 카메라 파라미터를 조절한 후, 상기 스테레오 카메라로부터 스테레오 영상이 획득될 수 있다.
- [0049] 일실시예에 따르면, 상기 영상 처리 장치(100)에 의해 결정되는 최적의 스테레오 카메라 파라미터는, 상기 실제 3D 물체에서 샘플링된 적어도 하나의 제1 포인트의 제1 위치와, 상기 시청자가 지각하는 3D 물체에서 제1 포인트에 대응되는 제2 포인트의 제2 위치의 차이로 정의된 목적 함수를 최소화함으로써 결정될 수 있다. 이로써,

상기 시청자에 의해 지각되는 상기 3D 물체 모양 및/또는 크기를 실제의 3D 물체의 그것으로 유지시킬 수 있다.

[0050] 일실시예에 따르면, 상기 스테레오 영상을 획득하기 위해 베이스 라인과 주시각이 고정되어 있는 상용화된 스테레오 카메라를 사용할 수 있으며, 이 경우에는, 가로 방향으로 넓은 영상 센서로부터 얻을 수 있는 넓은 영상에 기반하여 베이스 라인 변화와 유사한 효과를 나타내는 가상 베이스 라인 변화로 베이스 라인 변화를 근사하며, 최적의 가상 베이스 라인 및 초점 거리를 찾을 수 있다. 도 3의 (c)를 참고하여 가상 베이스라인 조절을 설명하면, 상기 가상의 베이스라인 b 는 좌측 카메라 내 좌측 영상 센서 위의 가로 방향으로 넓은 영상에서 획득되는 영역의 가로 방향 위치로 정의될 수 있으며, 우측 카메라의 가상 베이스 라인은 상기 좌측 카메라의 가상 베이스 라인에 대칭적으로 조절될 수 있다.

[0051] 상기 베이스 라인은 두 개의 카메라 중심 $C^{(L)}$ 과 $C^{(R)}$ 사이의 거리를 의미하며, 상기 베이스 라인의 조절은 상기 $C^{(L)}$ 과 $C^{(R)}$ 사이의 거리를 조절하는 것을 의미한다. 상기 베이스 라인을 작게 하여 상기 스테레오 영상을 촬영하는 경우 상기 스테레오 영상에서 오브젝트들이 더 멀리 보이며, 반대로 상기 베이스 라인을 크게 하여 상기 스테레오 영상을 촬영하는 경우, 상기 스테레오 영상에서 오브젝트들이 더 가까이 보인다.

[0052] 일실시예에 따르면, 상기 가상 베이스 라인 조절은 이미지 센서(image sensor) 위에서 스테레오 영상을 획득하는 영역을 가로 방향으로 움직이면서 수행될 수 있다. 상기 가상 베이스 라인 조절을 통해 획득되는 상기 스테레오 영상은, 실제로 베이스 라인을 조절하여 얻은 스테레오 영상과 동일하지는 않지만 유사할 수 있다.

[0053] 일실시예에 따르면, 상기 수신측의 제2 파라미터와 상기 실제 3D 물체의 깊이를 대략적으로 알고 있다는 가정하에, 상기 시청자가 지각하는 3D 물체의 위치, 크기, 모양을 상기 실제 3D 물체의 위치, 크기, 모양으로 유지하기 위해서, 상기 영상 처리 장치(100)는 수학식 1을 통해 상기 실제 3D 물체에서 샘플링 되는 적어도 하나의 제

1 포인트의 제1 위치 $\hat{\mathbf{A}}_n$ 과 제1 포인트에 대응되는 상기 시청자가 지각하는 3D 물체에서의 제2포인트의 제2 위치 $\mathbf{V}_{n,p}$ 의 차이로 정의된 목적함수 $J_1(\mathbf{p})$ 를 최소화하는 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다.

수학식 1

$$\begin{aligned} \hat{\mathbf{p}} &= \arg \min_{\mathbf{p}} J_1(\mathbf{p}) \\ &= \arg \min_{\mathbf{p}} \left[\frac{1}{N} \sum_{n=1}^N \left(\hat{\mathbf{A}}_n - \mathbf{V}_{n,p} \right)^2 \right], \end{aligned}$$

[0054]

[0055] 여기서, $\mathbf{p}$ 는 스테레오 카메라에 연관되는 제1 파라미터를 나타내며, N 은 상기 실제 3D 물체에서 샘플링된 상기 제1 포인트들의 개수를 나타낸다.

[0056] 다른 일실시예에 따르면, 상기 실제 3D 물체의 위치와 관계 없이 상기 시청자가 지각하는 물체의 크기 및 모양을 상기 실제 3D 물체의 크기 및 모양으로 유지하기 위해서, 상기 영상 처리 장치(100)는 수학식 2를 통해, 3D

공간상의 기준 위치 $\bar{\mathbf{A}}_n$ 에 대한 $\hat{\mathbf{A}}_n$ 의 상대적인 위치인 $\tilde{\mathbf{A}}_n$ 과, 3D 공간상의 기준 위치 $\bar{\mathbf{V}}_{n,p}$ 에 대한 $\mathbf{V}_{n,p}$ 의 상대적인 위치인 $\tilde{\mathbf{V}}_{n,p}$ 의 차이로 정의된 목적함수 $J_2(\mathbf{p})$ 를 최소화하는 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다.

수학식 2

$$\begin{aligned}\hat{\mathbf{p}} &= \arg \min_{\mathbf{p}} J_2(\mathbf{p}) \\ &= \arg \min_{\mathbf{p}} \left[\frac{1}{N} \sum_{n=1}^N \left(\tilde{\mathbf{A}}_n - \tilde{\mathbf{V}}_{n,\mathbf{p}} \right)^2 \right],\end{aligned}$$

[0057]

[0058] 여기서, 기준 위치 $\bar{\mathbf{A}}_n$ 과 $\bar{\mathbf{V}}_{n,\mathbf{p}}$ 는, 예를 들어 $\hat{\mathbf{A}}_n$ 과 $\tilde{\mathbf{V}}_{n,\mathbf{p}}$ 의 평균 위치를 나타낼 수 있으며, 이러한 경우 $\hat{\mathbf{A}}_n$ 과 $\tilde{\mathbf{V}}_{n,\mathbf{p}}$ 는 수학식 3을 통해 계산될 수 있다.

수학식 3

$$\begin{aligned}\tilde{\mathbf{A}}_n &= \hat{\mathbf{A}}_n - \bar{\mathbf{A}}_n \text{ where } \bar{\mathbf{A}}_n = \frac{1}{N} \sum_{n=1}^N \hat{\mathbf{A}}_n, \\ \tilde{\mathbf{V}}_{n,\mathbf{p}} &= \mathbf{V}_{n,\mathbf{p}} - \bar{\mathbf{V}}_{n,\mathbf{p}} \text{ where } \bar{\mathbf{V}}_{n,\mathbf{p}} = \frac{1}{N} \sum_{n=1}^N \mathbf{V}_{n,\mathbf{p}}.\end{aligned}$$

[0059]

[0060] 다른 일실시예에 따르면, 상기 실제 3D 물체의 위치 및 크기와 관계 없이 상기 시청자가 지각하는 물체의 모양을 상기 실제 3D 물체의 모양으로 유지하기 위해서, 상기 영상 처리 장치(100)는 수학식 4를 통해, $\tilde{\mathbf{A}}_n$ 과 $\tilde{\mathbf{V}}_{n,\mathbf{p}}$ 및 스케일 팩터 s 의 곱의 차이로 정의된 목적함수 $J_3(\mathbf{p}, s)$ 를 최소화하는 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다.

수학식 4

$$\begin{aligned}\hat{\mathbf{p}}, \hat{s} &= \arg \min_{\mathbf{p}, s} J_3(\mathbf{p}, s) \\ &= \arg \min_{\mathbf{p}, s} \left[\frac{1}{N} \sum_{n=1}^N \left(\tilde{\mathbf{A}}_n - s \cdot \tilde{\mathbf{V}}_{n,\mathbf{p}} \right)^2 \right],\end{aligned}$$

[0061]

[0062] 여기서, $\tilde{\mathbf{A}}_n$ 과 $\tilde{\mathbf{V}}_{n,\mathbf{p}}$ 은 수학식 3을 통해 계산될 수 있다.

[0063] 다른 일실시예에 따르면, 상기 영상 처리 장치(100)는 3D 스테레오 디스플레이 스크린으로부터 시청자가 지각하는 3D 물체까지의 과도한 거리로 인한 시각적 피로도가 감소하도록 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다. 이를 태면, 3D 스테레오 디스플레이 스크린으로부터 시청자가 지각하는 3D 물체까지의 과도한 거리는 상기 스테레오 영상 내에서 과도한 디스패리티를 유발하여 시청자에게 시각적 불편함을 줄 수 있

으며, 특히 3D 스테레오 디스플레이 스크린에 비해 물체가 시청자에게 가까울 때 더욱 그렇다.

[0064] 따라서, 상기 영상 처리 장치(100)는 수학적 식 5를 통해, 상기 목적 함수에 상기 3D 스테레오 디스플레이 스크린과 상기 시청자가 지각하는 3D 물체의 포인트들 사이의 거리들에 대한 가중치 합으로 정의된 추가 항 (이하, '디스페리티 제약 항'이라 함)을 더한 목적함수 $J_4(\mathbf{p})$ 를 최소화하는 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다:

수학적 식 5

$$\begin{aligned}\hat{\mathbf{p}} &= \arg \min_{\mathbf{p}} J_4(\mathbf{p}) \\ &= \arg \min_{\mathbf{p}} J(\mathbf{p}) + w_d \cdot \left[\frac{1}{N} \sum_{n=1}^N w_n \cdot \left(d_v - Z_{n,\mathbf{p}}^{(v)} \right)^2 \right]\end{aligned}$$

[0065]

[0066] 여기서, w_n 는 위의 상기 추가항에 대한 가중치를 나타내며, d_v 는 시청 거리 (상기 시청자로부터 상기 3D 스테레오 디스플레이 스크린까지의 거리)이며, w_n 은 상기 3D 스테레오 디스플레이 스크린으로부터 상기 시청자가 지각

하는 3D 물체에서의 n 번째 포인트 $\mathbf{V}_{n,\mathbf{p}} = [X_{n,\mathbf{p}}^{(v)}, Y_{n,\mathbf{p}}^{(v)}, Z_{n,\mathbf{p}}^{(v)}]^T$ 까지의 거리에 대한 가중치이다. $J(\mathbf{p})$ 는 수학적 식 1, 2, 4의 $J_1(\mathbf{p})$, $J_2(\mathbf{p})$, $J_3(\mathbf{p})$ 중 하나가 될 수 있다.

[0067] 가중치 w_n 은 상기 시청자가 지각하는 3D 물체에서의 포인트 $\mathbf{V}_{n,\mathbf{p}}$ 의 위치에 따라 다르게 설정될 수 있다. 예를 들면, 상기 시청자가 지각하는 3D 물체의 대부분이 3D 스테레오 디스플레이 스크린보다 멀게 보이도록 하기 위해, 3D 스테레오 디스플레이 스크린보다 $\mathbf{V}_{n,\mathbf{p}}$ 가 먼 경우 ($Z_{n,\mathbf{p}}^{(v)} > d_v$), w_n 을 작게 설정할 수 있다. 이는 3D 스테레오 디스플레이 스크린보다 가까운 물체에 의해 발생하는 시각적 피로도가 3D 스테레오 디스플레이 스크린보다 먼 물체에 의해 발생하는 시각적 피로도보다 심하기 때문이다.

[0068] 다른 일실시예에 따르면, 스테레오 비디오 획득의 경우, 상기 영상 처리 장치(100)는 상기 스테레오 카메라 파라미터의 급격한 변화로 인한 시각적 피로도가 감소하도록, 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다. 이를 태면, 상기 영상 처리 장치(100)를 이용하여 스테레오 영상을 획득하는 경우, 각각의 프레임 별로 최적의 $\mathbf{p}$ 를 찾을 수 있다. 이 경우, 스테레오 비디오 획득 시 시간에 따라 최적의 $\mathbf{p}$ 가 급격히 변화하여 상기 시청자의 시각적 피로도가 증가될 수 있다.

[0069] 따라서, 상기 영상 처리 장치(100)는 수학적 식 6을 통해, 상기 목적 함수에 시간에 따른 $\mathbf{p}$ 의 변화에 대한 비용(페널티)로 정의된 추가 항 (이하, '파라미터 변화 제약 항'이라 함)을 더한 목적함수 $J_5(\mathbf{p})$ 를 최소화하는 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 를 구할 수 있다:

수학적 식 6

$$\begin{aligned}\hat{\mathbf{p}}_t &= \arg \min_{\mathbf{p}_t} J_5(\mathbf{p}) \\ &= \arg \min_{\mathbf{p}_t} J(\mathbf{p}) + w_p \cdot (\mathbf{p}_t - \hat{\mathbf{p}}_{t-1})^2,\end{aligned}$$

[0070]

[0071] 여기서, w_p 는 위의 상기 추가 항에 대한 가중치를 나타내며, $\mathbf{p}_t$ 는 시간 t 에서의 스테레오 카메라 파라미터인 상기 제1 파라미터를 나타낸다. $J(\mathbf{p})$ 는 수학적 식 1, 2, 4의 $J_1(\mathbf{p})$, $J_2(\mathbf{p})$, $J_3(\mathbf{p})$ 중 하나가 될 수 있다.

[0072] 다른 일실시예에 따르면, 상기 영상 처리 장치(100)는 수학적 식 7을 통해, 수학적 식 1, 2, 4, 5, 6의 목적 함수 $J_1(\mathbf{p})$, $J_2(\mathbf{p})$, $J_3(\mathbf{p})$, $J_4(\mathbf{p})$, $J_5(\mathbf{p})$ 의 가중치 합으로 정의된 목적 함수를 최소화하는 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}_t$ 를 구할 수 있다:

수학적 식 7

$$\begin{aligned} \hat{\mathbf{p}}_t, s &= \arg \min_{\mathbf{p}_t, s} J(\mathbf{p}_t, s) \\ &= \arg \min_{\mathbf{p}_t, s} w_1 \cdot J_1(\mathbf{p}_t) + w_2 \cdot J_2(\mathbf{p}_t) + w_3 \cdot J_3(\mathbf{p}_t, s) \\ &\quad + w_d \cdot \left[\frac{1}{N} \sum_{n=1}^N w_n \cdot (d_v - Z_{n, \mathbf{p}_t}^{(V)})^2 \right] + w_p \cdot (\mathbf{p}_t - \hat{\mathbf{p}}_{t-1})^2. \end{aligned}$$

[0073]

[0074] 여기서, w_1 , w_2 , w_3 는 각각 $J_1(\mathbf{p}_t)$, $J_2(\mathbf{p}_t)$, $J_3(\mathbf{p}_t)$ 에 대한 가중치이다.

[0075] 수학적 식 7에서 가중치 w_1 , w_2 , w_3 , w_d , w_p 는 다양하게 조절될 수 있다. 이를 테면, $J_1(\mathbf{p}_t)$ 만을 이용하여 상기 스테레오 카메라에 연관되는 제1 파라미터를 구하기 위해서는 w_1 을 제외한 w_2 , w_3 , w_d , w_p 를 0으로 설정할 수 있다. 또 다른 예로, 상기 3D 스테레오 디스플레이 스크린과 상기 시청자가 지각하는 3D 물체 사이의 과도한 거리로 인해 발생하는 시각적 피로도가 적어야 한다면, 상기 w_d 는 상대적으로 크게 설정될 수 있다.

[0076] 일실시예에 따르면, 수학적 식 1에서 수학적 식 7까지의 상기 목적 함수의 최소화를 위해 최적화 방법을 사용할 수 있다. 여기서 최적화는, $\mathbf{p}$ 의 이산 검색 공간(discrete search space)에서의 완결 또는 부분 검색 방법(exhaustive or partial search method), Newton's method와 같은 비선형 최적화 방법 또는 위의 식을 근사하여 최적화하는 등 다양한 방법을 통해 수행될 수 있다. 또한, 상기 목적 함수를 다르게 정의할 경우 목적 함수의 최대화를 위해 최적화 방법을 사용할 수도 있다.

[0077] 일실시예에 따르면, 수학적 식 1에서 수학적 식 7까지의 상기 목적 함수의 해를 구하는 경우, 상기 실제 3D 물체의 샘플링된 제1 포인트의 개수 N 을 $\mathbf{p}$ 의 파라미터 개수보다 충분히 많도록 설정하여, 최소화 문제가 결정되지 못하는 문제(underdetermined problem)가 되는 것을 막을 수 있다. 이를 테면, $\mathbf{p}$ 가 수심측과 송신측에 연관되는 파라미터 7개의 파라미터(제1 파라미터 및 제2 파라미터를 의미하며, 도 3에 도시된 실시예에서는 d_c , θ , b , f , w_i , w_s , 및 d_v)를 포함하는 경우, 최소 7개 이상의 실제 3D 물체의 샘플링 포인트들의 좌표들을 이용하여 수학적 식 1에서 수학적 식 7까지의 최소화 문제를 풀 수 있다.

[0078] 또한, N 이 $\mathbf{p}$ 의 파라미터 개수 보다 충분히 큰 경우, 상기 실제 3D 물체 내 표본 점 중 일부만을 이용하여 수학적 식 1에서 수학적 식 7까지의 상기 목적 함수의 해를 구하는 것이 가능하다. 이 경우, 이상치(outlier)를 제거하기 위해 RANSAC (RANDOM SAMPLE CONSENSUS)과 같은 방법을 이용하여 상기 실제 3D 물체에서 샘플링된 제1 포인트들의 제1 위치들 중 신뢰할 수 있는 결과만을 선택하여 사용할 수 있다.

[0079] 일반적으로, 수학적 식 1에서 수학적 식 7까지의 상기 목적 함수의 최소화 과정에서 스테레오 카메라 파라미터인 상기 제1 파라미터는 베이스 라인(d_c), 초점 거리(f), 주시각(θ), 가상 베이스 라인(b)을 포함할 수 있다($\mathbf{p} = \{d_c, f, \theta, b\}$). 상기 베이스 라인과 상기 주시각이 고정되어 있는 스테레오 카메라를 사용하는 경우, 상기 $\mathbf{p}$ 는 상기 가상 베이스 라인과 상기 초점 거리만을 포함할 수 있다($\mathbf{p} = \{b, f\}$).

[0080] 일실시예에 따르면, 수학적 식 1에서 수학적 식 7까지의 상기 목적 함수를 최소화하는 해를 구하면(최적의 스테레오 카메라 파라미터인 제1 파라미터를 결정하면), 상기 송신측의 상기 스테레오 카메라에 연관되는 상기 제1 파라미터를 상기 최적의 스테레오 카메라 파라미터 값으로 조절하여 스테레오 영상을 획득할 수 있다. 이로써, 상

기 시청자가 지각하는 3D 물체 모양 및/또는 크기들이 상기 실제 3D 물체의 그것으로 유지될 수 있다.

[0081] 일실시예에 따르면, 확대되거나 축소된 3D 지각을 위해, 상기 시청자 또는 상기 송신측과 연관된 스테레오 카메라 사용자가 특정 초점거리(zoom 수준)를 정하는 경우가 있다. 이 경우, 상기 영상 처리 장치(100)는 정해진 초점 거리 근처의 제한된 검색 공간 안에서 초점 거리를 결정할 수 있다. 또한, 상기 스테레오 영상에서 하나의 3D 물체가 정해지고 기존의 물체 분할 방법 또는 수동으로 물체의 영역을 선택할 수 있다면, 수학식 1에서

수학식 7까지의 최소화 과정에서 $\hat{\mathbf{A}}_n$ 과 $\mathbf{V}_{n,p}$ 를 정해진 물체에 대해서만 계산함으로써, 상기 3D 물체의 모양 및/또는 크기만을 고려하도록 변경할 수 있다.

[0082] 아래에서, 상기 영상 처리 장치(100)에서 최적의 스테레오 카메라 파라미터를 결정하는 계산 과정을 보다 상세히 설명한다. 또한, 여기에서 2D 및 3D 공간 내 점들의 좌표는 등차좌표계(homogeneous coordinates)를 이용하여 표현된다.

[0083] 도 4은 일실시예에 따른, 블록 디스패리티의 예측을 나타내는 도면이다. 일실시예에 따르면, 상기 실제 3D 물체 내 샘플링된 제1 포인트의 제1 위치 i_1 을 계산하기 위해, 먼저 입력된 미리 보기 스테레오 영상 내 디스패리티들을 영상 블록 쌍 단위로 예측할 수 있다. 그런 다음, 이를 이용하여 상기 실제 3D 물체 내 점의 제1 포인트의 제1 위치를 계산할 수 있다.

[0084] 또한, 특징 점 추출(feature point extraction)에 기반하여 도 4의 (b)에 도시되는 좌측 및 우측 영상에서의 대응점 (corresponding points)들을 찾아 이들의 디스패리티로부터 상기 실제 3D 물체 내 제1 포인트들의 제1 위치를 예측하는 방법을 사용하는 것도 가능할 수 있다.

[0085] 일실시예에 따르면, 도 4의 (b)와 같은 상기 입력된 미리 보기 스테레오 영상의 좌측 및 우측 영상이 도 4의 (a)와 같이 N 개의 영상 블록들로 나뉜다고 가정할 수 있다. 이 경우, $B_n^{(L)}$ 과 $B_n^{(R)}$ 은 좌측 및 우측 영상의 n 번째 영상 블록 내 픽셀들의 집합을 각각 나타낸다. 수평 블록 정합(block matching)에 기반하여 n 번째 영상 블록 쌍에 대응하는 블록 디스패리티 d_n 을 예측하는 것은 아래의 수학식 8과 같이 표현될 수 있다.

수학식 8

$$d_n = \arg \min_{K_{\min} \leq k \leq K_{\max}} \left[\sum_{x,y \in B_n^{(L)}} \left(f_{x,y}^{(L)} - f_{x-k,y}^{(R)} \right)^2 \right]$$

[0086]

[0087] 이 경우, $f_{x,y}^{(L)}$ 와 $f_{x,y}^{(R)}$ 는 좌측 및 우측의 각각 영상에서 $[x, y, 1]^T$ 에서의 픽셀 값을 나타내며, $K_{\min}$ 와 $K_{\max}$ 은 검색 범위를 나타낸다. 도 4의 (c)는 블록 디스패리티 예측 결과의 실시예를 나타낸다. 앞서 기술된 블록 디스패리티 예측 방법은 텍스처가 적은 영상 블록에 대해서는 효과적이지 않을 수 있다. 따라서, 도 4의 (c)에서 "-"로 표시된 블록은 텍스처가 적은 영상 블록 쌍이기 때문에 이에 대해서는 블록 디스패리티 예측을 수행하지 않을 수 있다.

[0088] 도 5는 일실시예에 따른, 상기 실제 3D 물체에서 샘플링되는 적어도 하나의 제1 포인트의 제1 위치 예측을 나타내는 도면이다. 도 5을 참조하여 앞서 기술된 상기 스테레오 영상의 영상 블록 쌍에 대한 블록 디스패리티 d_n 이

예측되면, 도 5에서 영상 블록 쌍에 대응되는 상기 실제 3D 물체에서 샘플링되는 제1 포인트의 제1 위치 $\hat{\mathbf{A}}_n$ 의 3D 위치를 다음의 수학식 9와 같이 계산할 수 있다.

수학식 9

$$\hat{\mathbf{A}}_n = \arg \max_{\mathbf{A}_n} \left[\prod_{j \in \{L, R\}} \Pr(\mathbf{x}_n^{(j)} | \mathbf{A}_n, \mathbf{\Lambda}^{(j)}, \mathbf{\Omega}^{(j)}, \mathbf{\tau}^{(j)}) \right]$$

[0089]

[0090] 여기서, j 는 카메라 인덱스(좌측 또는 우측)를 나타내고, $\mathbf{x}_n^{(j)} = [\mathbf{x}_n^{(j)}, \mathbf{y}_n^{(j)}, 1]^T$ 는 좌측 또는 우측 영상에서 n 번째 영상 블록의 2D 좌표를 나타내며(이 경우, $\mathbf{x}_n^{(R)} = \mathbf{x}_n^{(L)} - \mathbf{d}_n$), $\mathbf{\Lambda}^{(j)}$ 은 좌측 또는 우측 카메라의 내부 행렬(intrinsic matrix)을 나타내고, $\mathbf{\Omega}^{(j)}$ 과 $\mathbf{\tau}^{(j)}$ 는 각각 좌측 또는 우측 카메라의 회전 및 이동 행렬(외부 행렬, extrinsic matrix)을 나타낸다.

[0091] 수학식 9에서, 내부 및 외부 행렬 $\{\mathbf{\Lambda}^{(j)}, \mathbf{\Omega}^{(j)}, \mathbf{\tau}^{(j)}\}$ 과 $\mathbf{A}_n$ 이 주어지는 경우, 영상 위의 좌표 $\mathbf{x}_n^{(j)}$ 를 관찰할 우

도(likelihood) $\Pr(\mathbf{x}_n^{(j)} | \mathbf{A}_n, \mathbf{\Lambda}^{(j)}, \mathbf{\Omega}^{(j)}, \mathbf{\tau}^{(j)})$ 는, 스페리칼 공분산(spherical covariance)에 의한 정규 분포로 표현되는 가산 잡음(additive noise)을 포함하는 핀홀(pinhole) 카메라 모델을 이용하여 수학식 10과 같이 표현될 수 있다:

수학식 10

$$\begin{aligned} & \Pr(\mathbf{x}_n^{(j)} | \mathbf{A}_n, \mathbf{\Lambda}^{(j)}, \mathbf{\Omega}^{(j)}, \mathbf{\tau}^{(j)}) \\ &= \text{Norm}_{\mathbf{x}_n^{(j)}} [\text{pinhole}[\mathbf{A}_n, \mathbf{\Lambda}^{(j)}, \mathbf{\Omega}^{(j)}, \mathbf{\tau}^{(j)}], \sigma^2 \mathbf{I}] \end{aligned}$$

[0092]

[0093] 여기서, $\text{Norm}_{\mathbf{x}} [\boldsymbol{\mu}, \boldsymbol{\Sigma}]$ 는 평균 $\boldsymbol{\mu}$ 및 공분산 $\boldsymbol{\Sigma}$ 을 포함하는 다변수 정규 분포를 나타내며, σ^2 는 잡음의 분산을 나타낸다. 상기 핀홀(pinhole) 카메라 모델은 수학식 11로 표현될 수 있다.

수학식 11

$$\begin{aligned} & \text{pinhole}[\mathbf{A}_n, \mathbf{\Lambda}, \mathbf{\Omega}, \mathbf{\tau}] = \mathbf{\Lambda} [\mathbf{\Omega} \quad \mathbf{\tau}] \mathbf{A}_n \\ &= \begin{bmatrix} r_1 f & \gamma & \delta_x \\ 0 & r_1 f & \delta_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta & -d_c \cos \theta \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & d_c \sin \theta \end{bmatrix} \mathbf{A}_n \end{aligned}$$

[0094]

[0095] 이 경우, r_1 은 3D 공간 좌표를 영상 좌표로 변환하기 위한 영상 센서에서의 축소 인자(down-scaling factor)를 나타낸다. 계산을 단순화하기 위해, 스큐(skew) 파라미터 γ 와 x 및 y 방향에 대한 영상 오프셋 파라미터 δ_x

및 δ_y 는 0으로 가정할 수 있다.

[0096] $\hat{\mathbf{A}}_n$ 의 3D 위치가 계산된 후에는, 수학식 3을 이용하여 3D 공간상의 기준 위치에 대한 $\hat{\mathbf{A}}_n$ 의 상대적인 위치를 나타내는 $\tilde{\mathbf{A}}_n$ 를 계산할 수 있다.

[0097] 도 6는 일실시예에 따른, 카메라 파라미터에 대해 시청자가 지각하는 3D 물체 내 제2 포인트의 제2 위치 계산을 나타내는 도면이다.

[0098] 일실시예에 따르면, 상기 실제 3D 물체에서 샘플링된 포인트의 3D 위치 $\hat{\mathbf{A}}_n$ 이 계산되고, 이에 해당하는 $\tilde{\mathbf{A}}_n$ 이 계산되면, 상기 시청자가 지각하는 3D 물체의 모양 및/또는 크기를 상기 실제 3D 물체의 모양 및/또는 크기로 유지하도록, 수학식 1에서 수학식 7까지의 목적 함수를 최소화 하는 해를 구할 수 있다. 상기 목적 함수를 최소화 하는 해를 구함으로써, 상기 스테레오 카메라에 연관되는 제1 파라미터($\mathbf{p}$)를 결정할 수 있다. 이를 위해, 주어진 $\mathbf{p}$ 에 대한 $\tilde{\mathbf{V}}_{n,\mathbf{p}}$ 를 계산하는 방법을 도 6를 참조하여 설명한다

[0099] 도 6의 (a)는 상기 송신측에서의 상기 실제 3D 물체를 나타내고, 도 6의 (b)는 상기 수신측에서 상기 시청자가 지각하는 3D 물체를 나타낸다. 도 6의 (a)에서, 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 가 주어지는 경우, $\hat{\mathbf{A}}_n$ 가 좌측 영상(610) 및 우측 영상(620)에 투영된 점을 각각 $\mathbf{x}_{n,\mathbf{p}}^{(L)}$ 와 $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 라고 한다. 수학식 9를 이용하여 계산된 $\hat{\mathbf{A}}_n$ 으로부터, $\mathbf{x}_{n,\mathbf{p}}^{(L)}$ 와 $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 는 수학식 11의 핀홀(pinhole) 카메라 모델을 이용하여 수학식 12과 같이 나타낼 수 있다.

수학식 12

$$\begin{aligned}\mathbf{x}_{n,\mathbf{p}}^{(L)} &= \mathbf{T}_b^{(L)} \left[\text{pinhole} \left[\hat{\mathbf{A}}_n, \Lambda_{\mathbf{p}}^{(L)}, \Omega_{\mathbf{p}}^{(L)}, \tau_{\mathbf{p}}^{(L)} \right] \right] \\ &= \mathbf{T}_b^{(L)} \Lambda_{\mathbf{p}}^{(L)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(L)} & \tau_{\mathbf{p}}^{(L)} \end{bmatrix} \hat{\mathbf{A}}_n, \\ \mathbf{x}_{n,\mathbf{p}}^{(R)} &= \mathbf{T}_b^{(R)} \left[\text{pinhole} \left[\hat{\mathbf{A}}_n, \Lambda_{\mathbf{p}}^{(R)}, \Omega_{\mathbf{p}}^{(R)}, \tau_{\mathbf{p}}^{(R)} \right] \right] \\ &= \mathbf{T}_b^{(R)} \Lambda_{\mathbf{p}}^{(R)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(R)} & \tau_{\mathbf{p}}^{(R)} \end{bmatrix} \hat{\mathbf{A}}_n,\end{aligned}$$

$$\left(\mathbf{T}_b^{(L)} = \begin{bmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{T}_b^{(R)} = \begin{bmatrix} 1 & 0 & -b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

[0100]

[0101] 이 경우, $\Lambda_{\mathbf{p}}^{(j)}$, $\Omega_{\mathbf{p}}^{(j)}$ 및 $\tau_{\mathbf{p}}^{(j)}$ 는 주어진, 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 에 대한 좌측 또는 우측 카메라의 내부 행렬과 회전 및 이동 행렬 (외부 행렬)을 나타내고, $\mathbf{T}_b^{(j)}$ 는 스테레오 영상의 가상 베이스라인 조절을 위한 변환 행렬을 나타낸다.

[0102] $\mathbf{x}_{n,\mathbf{p}}^{(L)}$ 와 $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 을 계산한 후, 스테레오 카메라의 주시각에 의해 발생하는 상기 시청자가 지각하는 3D 물체의

왜곡을 감소하기 위한 기하학적 영상 보상이 아래의 수학식 13와 같이 수행될 수 있다.

수학식 13

$$\begin{aligned}\mathbf{x}_{n,\mathbf{p}}^{(cL)} &= \mathbf{T}_c^{(L)} \mathbf{x}_{n,\mathbf{p}}^{(L)}, \\ &= \mathbf{T}_c^{(L)} \mathbf{T}_b^{(L)} \Lambda_{\mathbf{p}}^{(L)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(L)} & \tau_{\mathbf{p}}^{(L)} \end{bmatrix} \hat{\mathbf{A}}_n, \\ \mathbf{x}_{n,\mathbf{p}}^{(cR)} &= \mathbf{T}_c^{(R)} \mathbf{x}_{n,\mathbf{p}}^{(R)}, \\ &= \mathbf{T}_c^{(R)} \mathbf{T}_b^{(R)} \Lambda_{\mathbf{p}}^{(R)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(R)} & \tau_{\mathbf{p}}^{(R)} \end{bmatrix} \hat{\mathbf{A}}_n,\end{aligned}$$

$$\left(\mathbf{T}_c^{(L)} = \begin{bmatrix} c^{(L)} \Big|_{(-\theta), x_{n,\mathbf{p}}^{(L)}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{T}_c^{(R)} = \begin{bmatrix} c^{(R)} \Big|_{\theta, x_{n,\mathbf{p}}^{(R)}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

[0103]

여기서, $\mathbf{T}_c^{(j)}$ 는 스테레오 영상에서의 기하학적 영상 보상을 위한 변환 행렬을 나타내며, $c^{(j)} \Big|_{\theta, x_{n,\mathbf{p}}^{(j)}}$ 는 주시각 (θ) 및 $\mathbf{x}_{n,\mathbf{p}}^{(j)}$ 의 x 좌표 $x_{n,\mathbf{p}}^{(j)}$ 에 따라 결정되는 보상 변수를 나타낸다. 도 6의 (b)에서 스테레오 영상이 3D 스테레오 디스플레이 스크린에 디스플레이 되었을 경우, $\mathbf{x}_{n,\mathbf{p}}^{(cL)}$ 와 $\mathbf{x}_{n,\mathbf{p}}^{(cR)}$ 에 대응되는 3D 스테레오 디스플레이 스크린 위의 3D 포인트들 $\mathbf{S}_{n,\mathbf{p}}^{(L)}$ 와 $\mathbf{S}_{n,\mathbf{p}}^{(R)}$ 를 수학식 14와 같이 계산할 수 있다:

수학식 14

$$\begin{aligned}\mathbf{S}_{n,\mathbf{p}}^{(L)} &= \lambda \begin{bmatrix} X_{n,\mathbf{p}}^{(SL)}, Y_{n,\mathbf{p}}^{(SL)}, Z_{n,\mathbf{p}}^{(SL)}, 1 \end{bmatrix} = \mathbf{T}_s \mathbf{x}_{n,\mathbf{p}}^{(cL)} \\ &= \mathbf{T}_s \mathbf{T}_c^{(L)} \mathbf{T}_b^{(L)} \Lambda_{\mathbf{p}}^{(L)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(L)} & \tau_{\mathbf{p}}^{(L)} \end{bmatrix} \hat{\mathbf{A}}_n, \\ \mathbf{S}_{n,\mathbf{p}}^{(R)} &= \lambda \begin{bmatrix} X_{n,\mathbf{p}}^{(SR)}, Y_{n,\mathbf{p}}^{(SR)}, Z_{n,\mathbf{p}}^{(SR)}, 1 \end{bmatrix} = \mathbf{T}_s \mathbf{x}_{n,\mathbf{p}}^{(cR)} \\ &= \mathbf{T}_s \mathbf{T}_c^{(R)} \mathbf{T}_b^{(R)} \Lambda_{\mathbf{p}}^{(R)} \begin{bmatrix} \Omega_{\mathbf{p}}^{(R)} & \tau_{\mathbf{p}}^{(R)} \end{bmatrix} \hat{\mathbf{A}}_n,\end{aligned}$$

$$\left(\mathbf{T}_s = \begin{bmatrix} r_2 & 0 & r_2 \cdot \delta_x \\ 0 & r_2 & r_2 \cdot \delta_y \\ 0 & 0 & d_v \\ 0 & 0 & 1 \end{bmatrix} \right)$$

[0105]

[0106]

여기서, r_2 와 $\mathbf{T}_s^{(j)}$ 는 영상 좌표를 3D 공간 좌표로 변환하기 위한 스크린 확대 인자 및 변환 행렬을 각각 나타낸다. 영상 오프셋 파라미터 δ_x 및 δ_y 는 0으로 가정할 수 있다. 도 6의 (b)에서, 상기 시청자의 좌측 및 우측 눈의 3D 위치를 $\mathbf{E}^{(L)} = [-d_e, 0, 0, 1]^T$ 과 $\mathbf{E}^{(R)} = [d_e, 0, 0, 1]^T$ 로 나타낼 수 있다. 그 다음, $\mathbf{S}_{n,p}^{(L)}$ 로부터 $\mathbf{E}^{(L)}$ 을 지나는 선과 $\mathbf{S}_{n,p}^{(R)}$ 로부터 $\mathbf{E}^{(R)}$ 을 지나는 선의 교차점을 구함으로써, $\hat{\mathbf{A}}_n$ 에 대응되는 시청자가 지각하는 3D 물체에서의 상기 제2 포인트의 상기 제2 위치 $\mathbf{V}_{n,p}$ 를 수학적 식 15와 같이 계산할 수 있다.

수학적 식 15

$$\begin{aligned} \mathbf{V}_{n,p} &= \lambda \left[X_{n,p}^{(r)}, Y_{n,p}^{(r)}, Z_{n,p}^{(r)}, 1 \right] = \mathbf{T}_v^{(L)} \mathbf{S}_{n,p}^{(L)} + \mathbf{T}_v^{(R)} \mathbf{S}_{n,p}^{(R)}, \\ &= \left(\begin{array}{c} \mathbf{T}_v^{(L)} \mathbf{T}_s \mathbf{T}_c^{(L)} \mathbf{T}_b^{(L)} \Lambda_p^{(L)} \left[\Omega_p^{(L)} \quad \tau_p^{(L)} \right] \\ + \mathbf{T}_v^{(R)} \mathbf{T}_s \mathbf{T}_c^{(R)} \mathbf{T}_b^{(R)} \Lambda_p^{(R)} \left[\Omega_p^{(R)} \quad \tau_p^{(R)} \right] \end{array} \right) \hat{\mathbf{A}}_n, \\ &= \mathbf{T}_p \hat{\mathbf{A}}_n. \end{aligned}$$

$$\left(\begin{array}{c} \mathbf{T}_v^{(L)} = \begin{bmatrix} d_e & 0 & 0 & 0 \\ 0 & d_e & 0 & 0 \\ 0 & 0 & d_e & 0 \\ 1 & 0 & 0 & d_e \end{bmatrix}, \mathbf{T}_v^{(R)} = \begin{bmatrix} d_e & 0 & 0 & 0 \\ 0 & d_e & 0 & 0 \\ 0 & 0 & d_e & 0 \\ -1 & 0 & 0 & d_e \end{bmatrix}, \\ \mathbf{T}_p = \mathbf{T}_v^{(L)} \mathbf{T}_s \mathbf{T}_c^{(L)} \mathbf{T}_b^{(L)} \Lambda_p^{(L)} \left[\Omega_p^{(L)} \quad \tau_p^{(L)} \right] \\ + \mathbf{T}_v^{(R)} \mathbf{T}_s \mathbf{T}_c^{(R)} \mathbf{T}_b^{(R)} \Lambda_p^{(R)} \left[\Omega_p^{(R)} \quad \tau_p^{(R)} \right]. \end{array} \right)$$

[0107]

[0108]

여기서, $\mathbf{T}_v^{(j)}$ 은 $\mathbf{S}_{n,p}^{(j)}$ 으로부터 $\mathbf{V}_{n,p}$ 을 구하기 위한 변환 행렬을 나타낸다. 수학적 식 15에서 주어진 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 에 대해 $\mathbf{T}_p$ 가 한 번 계산되면, 이를 이용해 모든 n 에 대한 $\mathbf{V}_{n,p}$ 을 계산할 수 있다. 이 후, 수학적 식 3을 이용하여 3D 공간상의 기준 위치 $\bar{\mathbf{V}}_{n,p}$ 에 대한 $\mathbf{V}_{n,p}$ 의 상대적인 위치를 나타내는 $\tilde{\mathbf{V}}_{n,p}$ 을 계산할 수 있다. 이 때, 수학적 식 3에서의 $\bar{\mathbf{V}}_{n,p}$ 는 $\bar{\mathbf{A}}_n$ 과 $\mathbf{T}_p$ 를 이용해 수학적 식 16과 같이 표현될 수 있다.

수학식 16

$$\bar{\mathbf{V}}_{n,\mathbf{p}} = \frac{1}{N} \sum_{n=1}^N \mathbf{V}_{n,\mathbf{p}} = \mathbf{T}_{\mathbf{p}} \bar{\mathbf{A}}_n$$

[0109]

[0110] 한 번 $\hat{\mathbf{A}}_n$ 이 계산이 되면, 수학식 1에서 수학식 7까지의 최소화를 수행하는 동안 주어진 상기 스테레오 카메라에 연관되는 제1 파라미터 $\mathbf{p}$ 에 대한 $\hat{\mathbf{V}}_{n,\mathbf{p}}$ 를 계산함으로써 목적 함수를 최소화 하는 상기 제2 파라미터($\hat{\mathbf{p}}$)를 찾을 수 있다.

[0111] 도 7은 일실시예에 따른, 주시각이 있는 스테레오 카메라를 이용하는 스테레오 영상 획득을 설명하는 도면이다.

[0112] 일실시예에 따르면, 주시각이 있는 스테레오 카메라를 이용하여 스테레오 영상을 획득할 경우, 상기 획득된 스테레오 영상으로부터 표현되는 상기 시청자가 지각하는 3D 물체는 평면 심도 왜곡(depth plane curvature)을 갖는다. 밀집한 디스패리티 예측이나 3D 재구성 없이 상기 평면 심도 왜곡을 감소시키기 위해, 이를 테면, 기하학적 영상 보상 방법이 사용될 수 있다.

[0113] 상기 기하학적 영상 보상 방법은 수학식 1에서 수학식 7까지의 상기 목적 함수를 최소화 하는 과정에서 상기 시청자가 지각하는 3D 물체 내 제2 포인트 $\mathbf{V}_{n,\mathbf{p}}$ 의 위치(상기 적어도 하나의 제2 위치)를 계산할 경우, 수학식 13을 통해 수행될 수 있다. 또한, 상기 기하학적 영상 보상 방법을 이미 획득한 스테레오 영상의 전체 픽셀에 적용하여 스테레오 영상을 통해 시청자가 지각하는 3D 물체에서의 왜곡을 줄일 수 있다.

[0114] 상기 기하학적 영상 보상 방법의 일실시예를 설명한다. 도 7은 주시각이 있는 스테레오 카메라를 이용한 스테레오 영상 획득을 도시한다. 우측 카메라의 3D 위치를 $\mathbf{C}^{(R)} = [d_c, 0, 0, 1]^T$ 이라 한다. 상기 스테레오 영상 획득시, 주어지는 스테레오 카메라의 파라미터(상기 제1 파라미터) $\mathbf{p}$ 에 대해서 상기 실제 3D 물체의 포인트 $\mathbf{A}_n = [X_n^{(A)}, Y_n^{(A)}, Z_n^{(A)}, 1]^T$ 이 우측 영상(720) 위 2D 점 $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 로 투영되었다고 가정할 때, $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 의 좌표는 핀홀(pinhole) 카메라 모델을 이용하여 수학식 17과 같이 나타낼 수 있다.

수학식 17

$$\begin{aligned} \mathbf{x}_{n,\mathbf{p}}^{(R)} &= \lambda \left[x_{n,\mathbf{p}}^{(R)}, y_{n,\mathbf{p}}^{(R)}, 1 \right]^T \\ &= \text{pinhole} \left[\mathbf{A}_n, \mathbf{\Lambda}^{(R)}, \mathbf{\Omega}^{(R)}, \boldsymbol{\tau}^{(R)} \right] = \mathbf{\Lambda}^{(R)} \left[\mathbf{\Omega}^{(R)} \quad \boldsymbol{\tau}^{(R)} \right] \mathbf{A}_n \\ &= \begin{bmatrix} r_1 f & \gamma & \delta_x \\ 0 & r_1 f & \delta_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta & -d_c \cos \theta \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & d_c \sin \theta \end{bmatrix} \mathbf{A}_n, \end{aligned}$$

[0115]

[0116] 또한, 수학식 17로부터 $\mathbf{x}_{n,\mathbf{p}}^{(R)}$ 의 x좌표 $x_{n,\mathbf{p}}^{(R)}$ 를 수학식 18과 같이 계산할 수 있다.

수학식 18

$$x_{n,p}^{(R)} = \frac{r_1 f \left(\left(X_{n,p}^{(A)} - d_c \right) \cos \theta + Z_{n,p}^{(A)} \sin \theta \right)}{Z_{n,p}^{(A)} \cos \theta - \left(X_{n,p}^{(A)} - d_c \right) \sin \theta}$$

[0117]

[0118] 이 경우, $\Lambda^{(R)}$ 는 우측 카메라의 내부 행렬을 나타내며, $\Omega^{(R)}$ 과 $\tau^{(R)}$ 는 우측 카메라의 회전 및 이동 행렬 (외부 행렬)을 각각 나타낸다. 내부 행렬에서, 스큐(skew) 파라미터 y 와 x 및 y 방향에 대한 영상 오프셋 파라미터 δ_x , δ_y 는 0으로 가정할 수 있다.

[0119] 주시각에 의해 발생하는 왜곡을 줄이기 위한 스테레오 영상에서의 기하학적 영상 보상을 위한 변환 행렬을 $T_c^{(j)}$ 라 할 때, 우측 영상(720)에서의 기하학적 영상 보상은 수학식 19와 같이 $T_c^{(R)}$ 을 통해 $x_{n,p}^{(R)}$ 의 좌표를 $x_{n,p}^{(cR)}$ 로 변환함으로써 수행된다.

수학식 19

$$x_{n,p}^{(cR)} = T_c^{(R)} x_{n,p}^{(R)},$$

$$T_c^{(R)} = \begin{pmatrix} c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

[0120]

[0121] 이 경우, $c^{(j)} \Big|_{\theta, x_{n,p}^{(j)}}$ 는 주시각 θ 및 $x_{n,p}^{(R)}$ 의 x좌표 $x_{n,p}^{(R)}$ 에 따라 결정되는 오른쪽 영상에서의 보상 변수를 나타내며, $\left(x_{n,p}^{(R)} \Big|_{\theta=0} \right) / \left(x_{n,p}^{(R)} \right)$ 로 정의된다. 여기서, $x_{n,p}^{(R)} \Big|_{\theta=0} = r_1 f \left(X_{n,p}^{(A)} - d_c \right) / Z_{n,p}^{(A)}$ 은 주시각 θ 가 0일 때의 $x_{n,p}^{(R)}$ 의 x좌표를 나타낸다.

[0122] 수학식 19에서 기하학적 영상 보상은 상기 우측 영상(720) 위 각각의 2D 점 $x_{n,p}^{(R)} = \lambda \left[x_{n,p}^{(R)}, y_{n,p}^{(R)}, 1 \right]^T$ 의 $x_{n,p}^{(R)}$ 에 $c^{(R)} \Big|_{x_{n,p}^{(R)}, \theta}$ 를 곱하여 새로운 좌표 $x_{n,p}^{(cR)} = \lambda \left[\left(c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}} \right) \cdot x_{n,p}^{(R)}, y_{n,p}^{(R)}, 1 \right]^T$ 를 계산하고 $x_{n,p}^{(R)}$ 를 $x_{n,p}^{(cR)}$ 로

이동시킴으로써 수행될 수 있다. 이를 통해 얻은 $\mathbf{x}_{n,p}^{(cR)}$ 는, 주시각이 0인 스테레오 카메라로 스테레오 영상을 획득한 경우의 $\mathbf{x}_{n,p}^{(R)}$ 을 근사한다.

[0123] 수학식 18에서, 테일러 급수(Taylor series)에 기초하여 $\theta \approx 0$ 일 때 $\sin \theta$ 와 $\cos \theta$ 를 θ 와 1로 근사하고, $\left| X_{n,p}^{(R)} \right| \square \left| d_c \right|$ 를 가정함으로써 보상 변수 $c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}}$ 는 수학식 20과 같이 계산될 수 있다.

수학식 20

$$c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}} \triangleq \left(x_{n,p}^{(R)} \Big|_{\theta=0} \right) / \left(x_{n,p}^{(R)} \right),$$

$$\approx \left(1 - \frac{X_{n,p}^{(A)}}{Z_{n,p}^{(A)}} \cdot \theta \right) / \left(1 + \frac{Z_{n,p}^{(A)}}{X_{n,p}^{(A)}} \cdot \theta \right)$$

[0124]

[0125] 나아가, $\theta \approx 0$ 일 경우, $X_{n,p}^{(A)} / Z_{n,p}^{(A)}$ 는 $x_{n,p}^{(A)} / f$ 로 근사화 될 수 있다. 따라서, $c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}}$ 는 수학식 21과 같이 나타낼 수 있다.

수학식 21

$$c^{(R)} \Big|_{\theta, x_{n,p}^{(R)}} \approx \left(1 - \frac{x_{n,p}^{(R)}}{f} \cdot \theta \right) / \left(1 + \frac{f}{x_{n,p}^{(R)}} \cdot \theta \right)$$

[0126]

[0127] 유사하게, 왼쪽 카메라의 주시각이 $-\theta$ 일 경우, 좌측 영상(710)에서의 기하학적 영상 보상은 수학식 22와 같이 수행된다.

수학식 22

$$\mathbf{x}_{n,p}^{(cL)} = \mathbf{T}_c^{(L)} \mathbf{x}_{n,p}^{(L)}$$

$$\mathbf{T}_c^{(L)} = \begin{pmatrix} \mathbf{c}^{(L)} \Big|_{(-\theta), x_{n,p}^{(L)}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{c}^{(L)} \Big|_{(-\theta), x_{n,p}^{(L)}} = \left(1 - \frac{x_{n,p}^{(L)}}{f} \cdot (-\theta) \right) / \left(1 + \frac{f}{x_{n,p}^{(R)}} \cdot (-\theta) \right).$$

[0128]

[0129] 일실시예에 따르면, 상기 스테레오 카메라의 주시각에 의해 발생하는 왜곡을 줄이기 위한 기하학적 영상 보상 방법에서는, 상기 스테레오 영상에서의 각 포인트의 새로운 좌표를 주시각 θ 및 영상 위 점의 x좌표에 따라 간단히 계산할 수 있으며, 밀집한 디스패리티 예측이나 3D 재구성이 필요하지 않다.

[0130] 도 8은 일실시예에 따른 영상 처리 방법(800)의 흐름도이다. 일실시예에 따르면, 상기 제1 계산부(110)는 상기 실제 3D 물체에서의 적어도 하나의 제1 포인트들의 3D 위치(상기 제1 위치)들을 영상 블록 쌍 단위로 수평 블록 정합 방법을 이용하여 계산할 수 있다(810).

[0131] 일실시예에 따르면, 상기 결정부(130)는 상기 제1 위치와 상기 제2 위치의 차이가 최소화되도록 하는 상기 송신 측에 연관되는 적어도 하나의 제1 파라미터(이를 테면, 최적의 스테레오 카메라 파라미터)를 계산하여 결정할 수 있다(820).

[0132] 일실시예에 따르면, 상기 영상 처리 장치의 제2 계산부(120)는 제2 제어부(150)로부터 시청자 환경 파라미터(이를테면, 제2 파라미터)를 전달받아, 주어진 제1 파라미터에 따라 상기 시청자가 지각하는 3D 물체에서 상기 제1 포인트에 대응하는 제2 포인트의 제2 위치를 계산할 수 있다(830).

[0133] 그리고, 결정부(130)는 상기 제1 파라미터가 최적의 파라미터인지를 판단하여, 최적의 파라미터가 아닌 경우에는 단계(820) 이하의 과정을 반복할 수 있다(840). 이러한 과정을 통해, 상기 제1 위치와 상기 제2 위치의 차이가 최소화되도록 제1 파라미터를 최적 값으로 결정할 수 있다.

[0134] 상기 최적의 스테레오 카메라 파라미터가 결정되면, 제1 제어부(140)는 상기 스테레오 카메라의 파라미터들을 최적의 파라미터 값으로 제어한 후 스테레오 영상을 촬영하고, 상기 수신측으로 상기 스테레오 영상을 전송할 수 있다(850).

[0135] 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, FPA(field programmable 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.

[0136] 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(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.

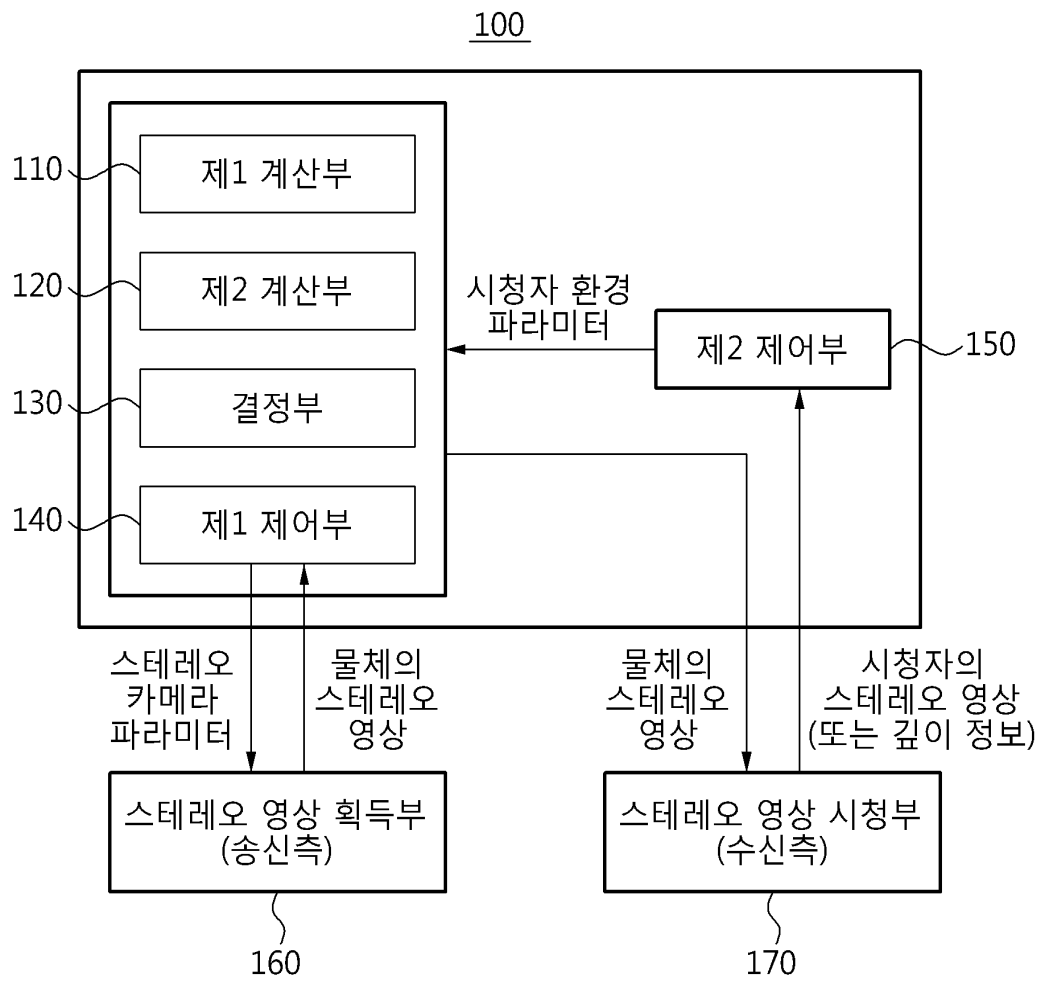
[0137] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic 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. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

[0138] 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.

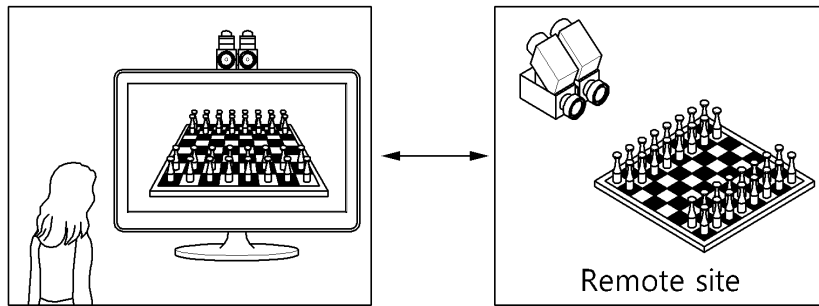
[0139] 그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

도면

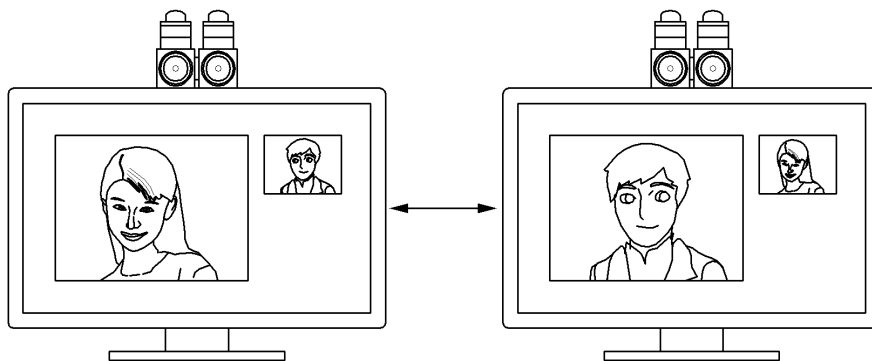
도면1



도면2

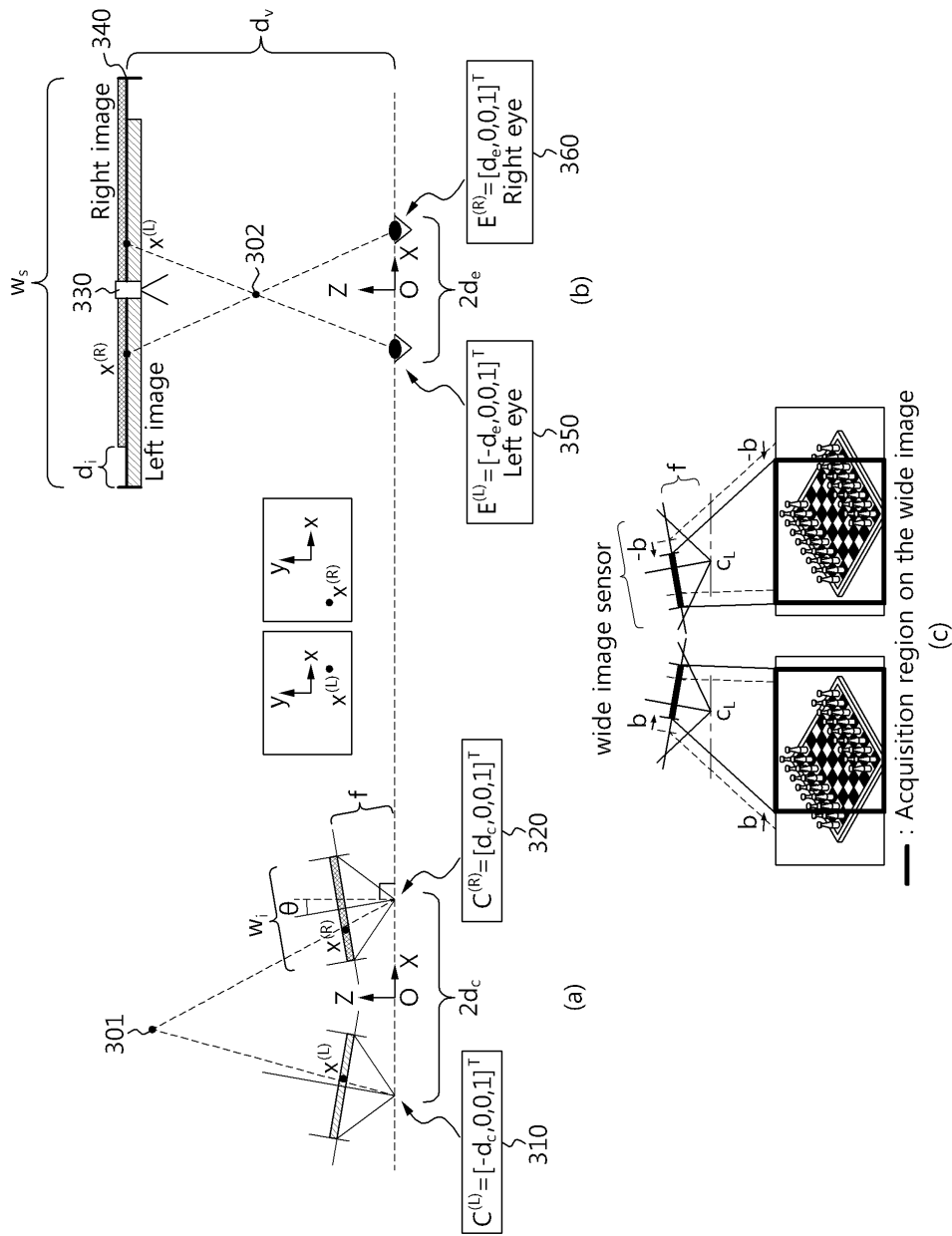


(a)

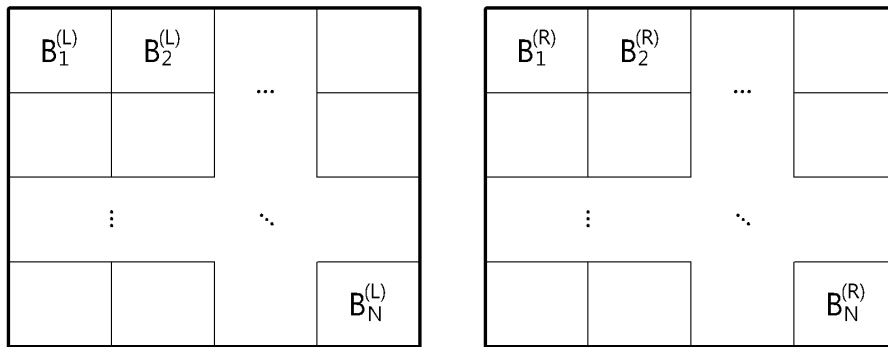


(b)

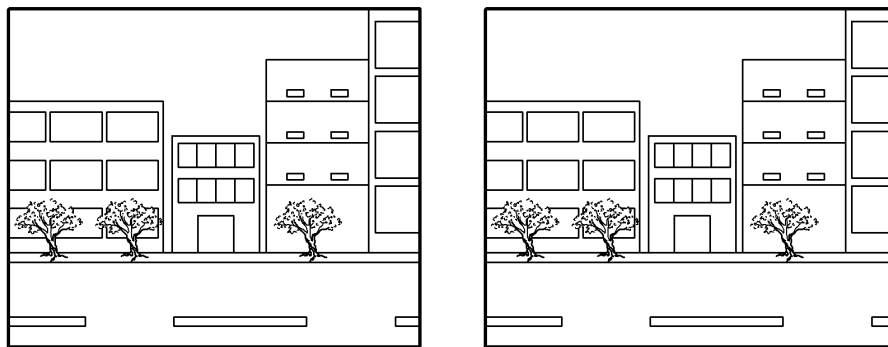
도면3



도면4



(a)

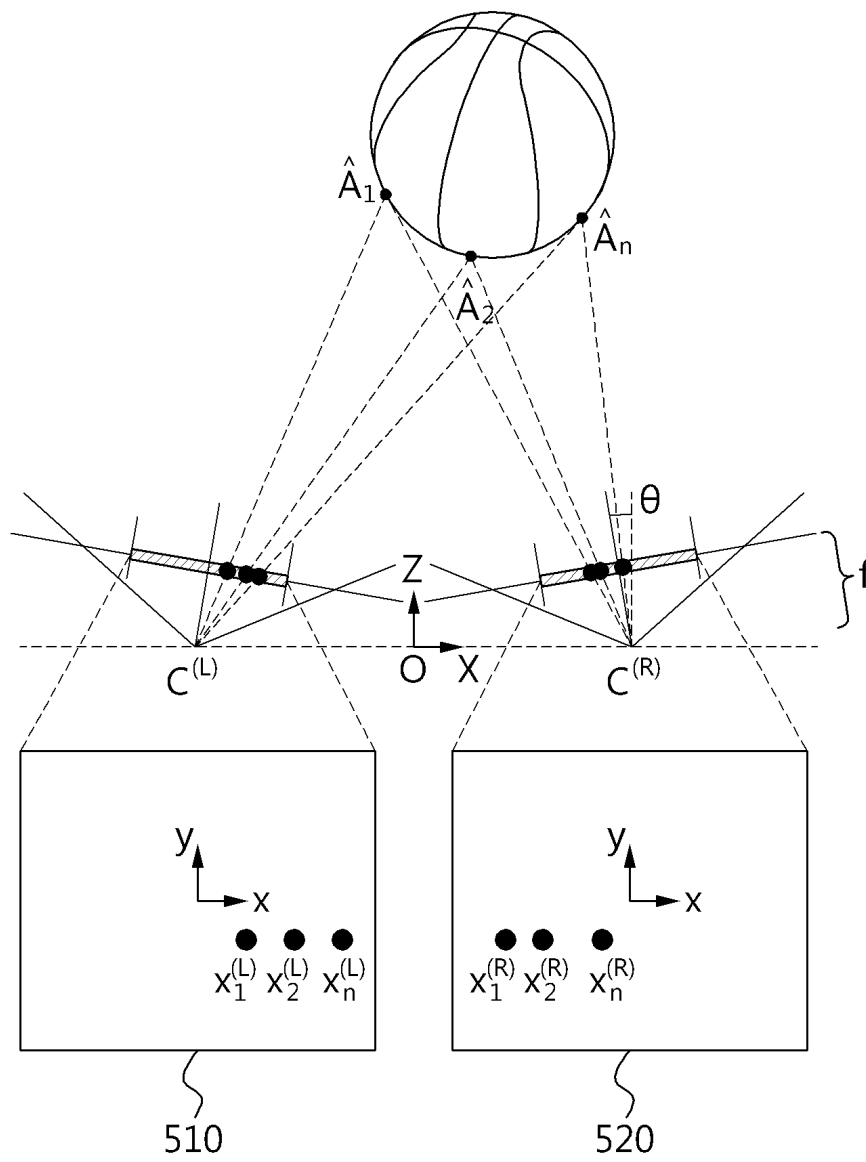


(b)

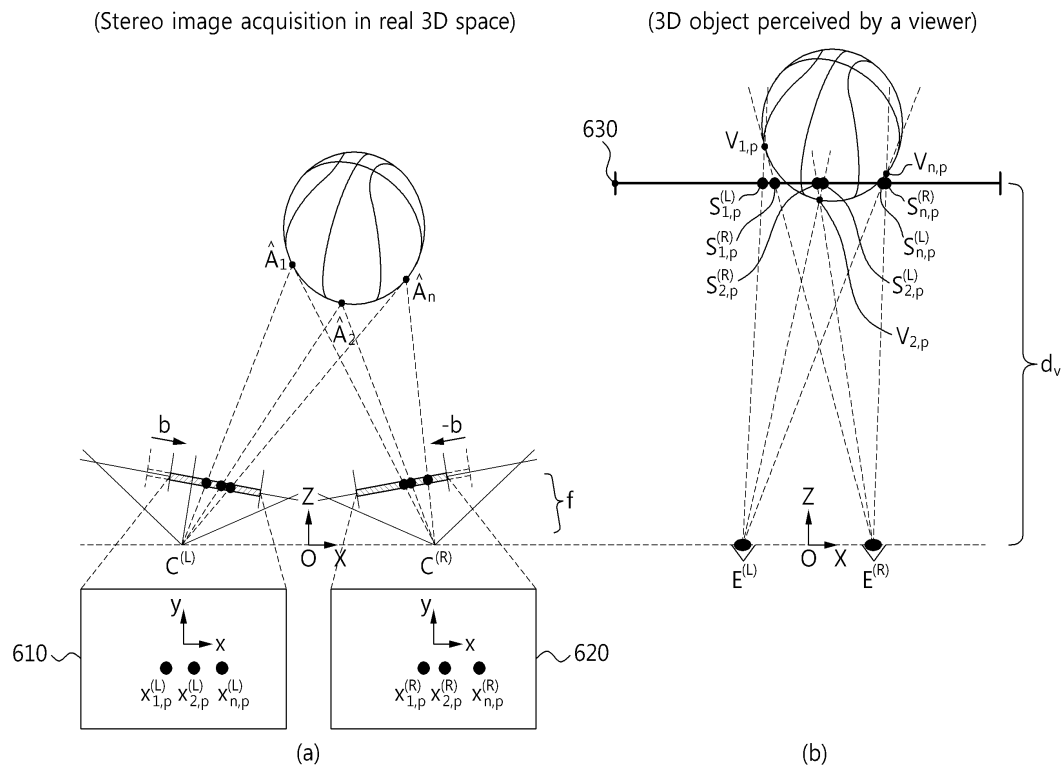
-	-	-	-	-	-	-	10
-	-	-	-	-	9	12	10
13	14	12	-	-	12	13	13
13	13	13	13	13	11	12	10
13	14	13	13	12	12	12	10
13	13	14	13	12	14	12	10
-	-	-	-	-	-	-	-
13	12	-	16	15	16	-	14

(c)

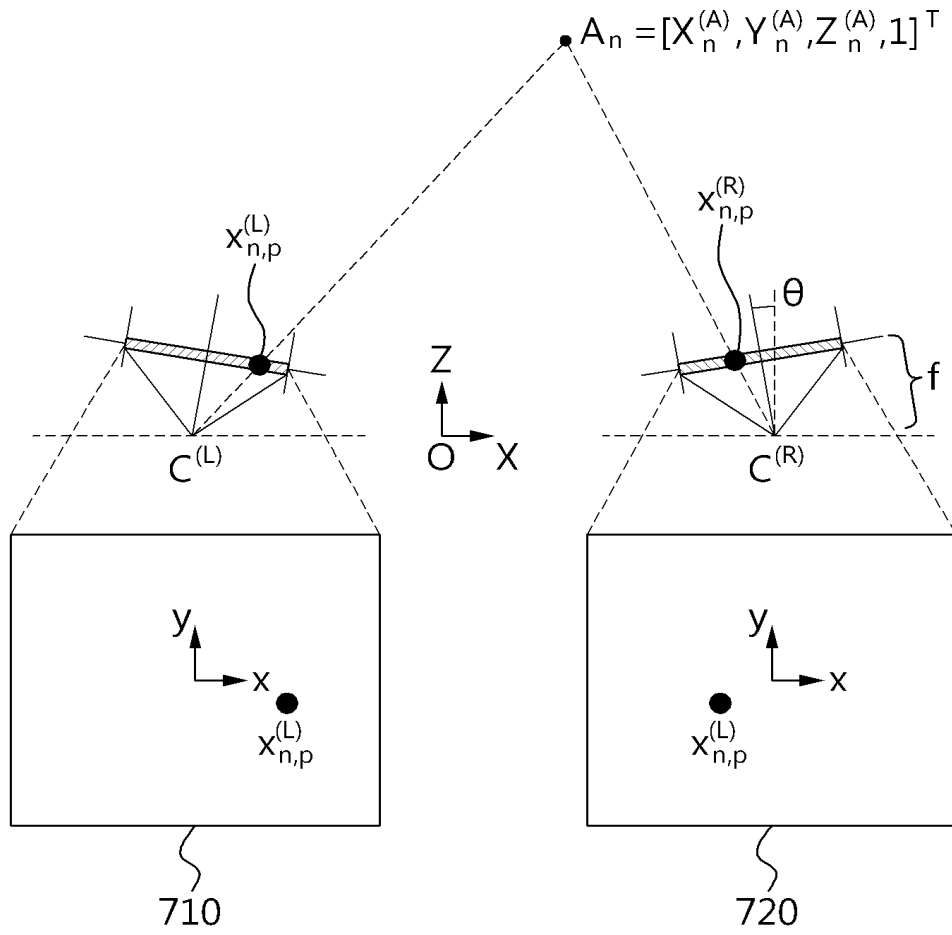
도면5



도면6



도면7



도면8

