

논문 2011-48SP-3-2

다시점 카메라와 깊이 카메라를 이용한 3차원 장면의 깊이 정보 생성 방법

(Depth Generation Method
Using Multiple Color and Depth Cameras)

강 윤 석*, 호 요 성**

(Yun-Suk Kang and Yo-Sung Ho)

요 약

본 논문에서는 다시점 색상 카메라와 다시점 깊이 카메라를 이용하여 촬영한 영상의 후처리 방법과 3차원 장면의 깊이 정보를 생성하는 방법을 제안한다. 깊이 카메라는 장면의 깊이 정보를 실시간으로 측정할 수 있는 장점이 있지만, 잡음과 왜곡이 발생하고 색상 영상과의 상관도도 떨어진다. 따라서 다시점 깊이 영상에 후처리 작업을 수행한 후, 이를 다시점 색상 영상과 조합하여 3차원 깊이 정보를 생성한다. 깊이 카메라로부터 얻은 각 시점에서의 초기 변이 정보를 기반으로 한 스테레오 정합의 결과는 기존 방법의 결과 보다 우수한 성능을 나타내었음을 볼 수 있었다.

Abstract

In this paper, we explain capturing, postprocessing, and depth generation methods using multiple color and depth cameras. Although the time-of-flight (TOF) depth camera measures the scene's depth in real-time, there are noises and lens distortion in the output depth images. The correlation between the multi-view color images and depth images is also low. Therefore, it is essential to correct the depth images and then we use them to generate the depth information of the scene. The results of stereo matching based on the disparity information from the depth cameras showed the better performance than the previous method. Moreover, we obtained the accurate depth information even at the occluded or textureless regions which are the weaknesses of stereo matching.

Keywords : Depth generation, Multi-view camera, TOF camera, 3DTV

I. 서 론

TV 방송이 1930년대 처음 시작된 이래로 방송에 관련된 기술들은 나날이 발전해 왔다. 영상의 촬영과 처리 방법, 부호화하는 방법, 효과적으로 전송하고 받아서 재생하는 방법과 장치에 이르기까지 눈부신 발전을 이룩해 왔다.

최근에는 전 세계적으로 폭발적인 관심을 받고 있는

3D 영화와 맞물려, 기존의 방송과 달리 입체감을 느낄 수 있는 3DTV의 상용화가 본격적으로 시작되었다. 많은 사람들이 전시회에서나 볼 수 있었던 입체영상을 이제는 안방에서 시청하기에 이르렀다. 3DTV는 이제 차세대 방송 시스템의 대표 주자로 인식되고 있으며, 관련된 활발한 연구가 지속되고 있다¹⁾.

이러한 3DTV와 같은 입체 디스플레이 장치를 이용하여 장면을 입체로 보기 위해서는 동일한 장면을 두 시점 이상에서 촬영한 영상이 필요하다. 특히 다시점 디스플레이의 경우는 3개 이상의 시점에서의 영상이 필요로 한다. 카메라의 개수 및 크기의 제약으로 인해 카메라 간격을 조밀하게 하여 많은 시점의 영상을 촬영하

* 학생회원, ** 정회원, 광주과학기술원 실감방송연구센터 (Realistic Broadcasting Research Center, Gwangju Institute of Science and Technology)
접수일자: 2010년9월13일, 수정완료일: 2011년1월26일

는 것은 불가능하지만, 다시점 영상과 장면의 깊이 영상을 이용하여 중간 시점에서의 영상을 합성하면 다시점 입체영상 콘텐츠를 생성할 수 있다.

따라서 촬영하는 장면에 대한 정확한 깊이 정보를 얻는 것은 입체영상 콘텐츠 제작에 있어서 무엇보다 중요하다. 일반적으로 장면의 깊이 정보를 얻는 방법은 크게 두 가지로 나눌 수 있는데, 수동 센서 방식과 능동 센서 방식이 그것이다. 수동 센서 방식은 촬영된 영상을 기반으로 하여 깊이 정보를 예측한다. 대표적으로 스테레오 정합(stereo matching)을 예로 들 수 있다. 이 방법은 데이터의 획득이 쉽고, 추가적인 장비가 필요하지 않다는 장점이 있지만, 가려짐 문제나 빛의 영향 등에 의해 고품질의 깊이 영상을 얻기가 힘들다.

능동 센서 방식은 측정 장비를 이용하여 장면의 깊이 정보를 직접 획득한다. 대표적 장비인 Time-of-flight (TOF) 깊이 카메라는 카메라에서 나온 빛이 물체에 도달하였다가 돌아오는 시간을 계산하여 깊이 영상을 출력한다. 이 방법은 실제의 깊이 정보를 실시간으로 측정할 수 있는 장점을 가지고 있지만, 장비의 사용에 의한 제약이나 잡음, 왜곡 문제 등이 따른다.

한편, 수동과 능동 센서 방식의 장점들을 혼합하여 사용하는 연구도 최근 활발히 진행되고 있다. 각각의 방식이 서로의 단점들을 잘 보완해줄 수 있기 때문이다. 이러한 혼합형 방식에서는 주로 양안식(stereo) 카메라 혹은 다시점 카메라와 함께 한 대의 깊이 카메라를 사용하여 장면의 깊이 정보를 생성하는 방식을 취하고 있다^[2-3]. 그러나 이러한 기존의 방법들은 오차가 많은 TOF 카메라 영상을 그대로 사용하였을 뿐 아니라, 한 대의 TOF 카메라를 사용하였기 때문에 장면의 깊이 정보를 효율적으로 획득하지 못하는 단점이 있다.

본 논문에서는 다시점 카메라와 다시점의 깊이 카메라를 이용하여 장면의 깊이 정보를 생성하는 방법을 소개한다. 카메라 시스템의 구성과 촬영, 영상의 후처리 방법에 대해 논한 후, 다시점의 깊이를 생성하는 방법을 소개한다.

II. 제안하는 깊이 생성 방법

1. 영상 획득 및 후처리

가. 다시점 색상 및 깊이 영상 획득

그림 1은 다시점 카메라와 다시점 깊이 카메라의 배열을 보여 준다. 두 종류의 카메라는 각각 평행하게 배

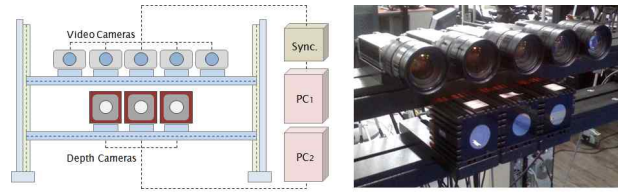
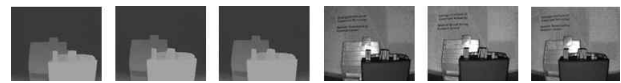


그림 1. 다시점 카메라와 다시점 깊이 카메라 배열
Fig. 1. Multi-view color and depth cameras.



다시점 색상 영상



다시점 깊이 영상

다시점 강도(intensity) 영상

그림 2. 촬영된 영상

Fig. 2. Captured images.

열되어 있으며, 카메라 간의 수평 및 수직 거리는 각각 약 68mm와 65mm이다. 다시점 카메라는 동기화장치를 통해 저장소에 연결되며, 다시점 깊이 카메라는 개별적으로 하나의 저장소에 연결된다.

이와 같은 카메라 배열에서 획득된 영상이 그림 2에 나타나 있다. 깊이 카메라의 출력 영상은 두 종류로써, 장면의 깊이를 0부터 255사이의 8비트 흑백 영상으로 나타낸 깊이 영상과 장면의 강도(intensity) 영상을 각각 획득할 수 있다. 그러나 그림 2에서 보듯이 해상도가 매우 작고(176x144), 렌즈 왜곡이 발생하며, 객체의 경계부 등에 잡음이 존재한다. 또한, 깊이 카메라의 변조 주파수의 제한으로 인하여 동시에 최대 세 대의 카메라밖에 사용하지 못하는 단점도 있다.

한편 다섯 시점에서 촬영된 다시점 색상 영상은 카메라 배열의 오차에 의해 각각의 에피폴라 선(epipolar line)이 서로 평행을 이루지 못하고 있으며, 카메라 간 특성 차이와 위치에 따른 조명의 영향, 카메라의 특성 차이 등으로 인하여 색상의 차이도 존재한다.

나. 획득한 영상의 후처리 과정

앞에서 언급한 다시점 색상 및 깊이 영상이 가지고 있는 문제점들을 해결하기 위해서는 몇 단계의 후처리 과정이 필요하다. 일반적으로 영상을 촬영한 후에는 카메라 보정(camera calibration)을 수행하여 카메라 변수를 얻는다^[4]. 하지만, 깊이 영상의 경우에는 뒤에서 설명할 렌즈 왜곡 보정 후에 카메라 보정을 수행한다. 색

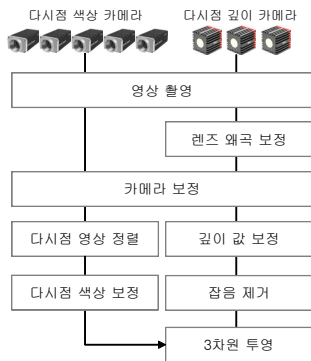


그림 3. 획득한 영상의 후처리 과정
Fig. 3. Postprocessing procedure for captured images.

상과 깊이 영상 각각에 적용해야 할 후처리 작업은 그림 3에 나타나 있다.

(1) 색상 영상 후처리

다시점 카메라를 통해서 획득한 다시점 색상 영상은 다시점 영상 정렬(multi-view image rectification)을 통하여 그림 4와 같은 형태로 변환된다^[5]. 다시점 영상 정렬의 결과로 모든 영상에서의 에피폴라 선이 평행을 이루게 되며, 동일한 특성을 갖춘 카메라를 수평 축을 따라 일정한 간격으로 배열한 특성을 가진 다시점 영상과 카메라 변수를 얻을 수 있다.

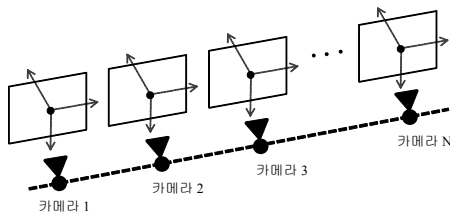


그림 4. 다시점 영상 정렬의 결과
Fig. 4. Results of multi-view image rectification.

영상 정렬 후에는 각 시점 영상 사이의 색상 차이를 보정하는 다시점 색상 보정(multi-view color correction)을 수행한다^[6]. 촬영 전 각 시점에서 미리 획득한 색상 차트를 기반으로 하여 시점과 시점 사이에 존재하는 색상 특성의 차이를 최소화한다.

(2) 깊이 영상 후처리

획득된 다시점의 깊이 영상은 색상 영상보다 더 많은 후처리 작업을 필요로 한다. 촬영 시 발생한 영상의 왜곡과 잡음 등을 최소화하고, 함께 촬영된 다시점 색상 영상과 함께 사용하기 위하여 깊이 값을 보정하는 과정



그림 5. 렌즈 왜곡 보정
Fig. 5. Lens distortion correction.

이 필요하다.

깊이 영상에 대해서 가장 먼저 수행해야 할 것은 렌즈 왜곡 보정이다. 그림 2에 나타난 것과 같이 깊이 카메라에서 획득된 영상은 장면 내의 직선 성분이 촬영을 통해 원형으로 왜곡되는 현상이 발생한다. 이 왜곡은 색상 카메라에서 획득된 영상과 깊이 영상의 상관도를 떨어뜨리므로 보정을 통해 최소화하여야 한다.

렌즈 왜곡 보정은 원형으로 왜곡이 발생한 최소한 세계의 직선 성분을 깊이와 함께 촬영된 강도 영상에서 찾은 후, 왜곡의 중심과 왜곡 변수를 찾음으로써 그림 5와 같이 보정할 수 있다^[7].

렌즈 왜곡 보정 후, 깊이 영상에 대하여 카메라 보정을 수행하여 카메라 변수를 획득한다. 영상 정렬된 다시점 색상 카메라와 왜곡 보정 후 획득한 깊이 카메라의 카메라 변수를 비교해 보면, 카메라의 광축(optical axis) 방향으로의 위치가 일치하지 않음을 알 수 있다. 이것은 깊이의 각 카메라에서의 측정된 장면의 깊이가 다를 수도 있음을 보여준다. 이 차이를 최소화하기 위해서 그림 6과 같이 격자무늬의 패턴을 장면의 가장 뒤쪽으로부터 카메라의 앞쪽까지 일정한 간격으로 이동하며 영상을 획득한다. 이 때, 다시점 카메라로부터는 패턴의 각 위치에 대한 변이(disparity)정보를 얻을 수 있고, 깊이 카메라로부터는 0부터 255사이의 깊이 값(depth index)을 얻을 수 있다.

다시점 색상 영상으로부터 얻은 변이와 깊이 카메라로부터 얻은 깊이 값을 바탕으로 둘 사이의 관계를 그림 7과 같은 함수로 만들 수 있다. 이 함수를 이용하여

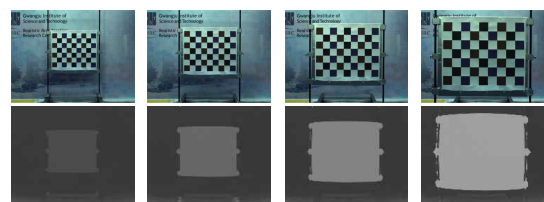


그림 6. 깊이 값 보정을 위한 촬영
Fig. 6. Images for depth value correction.

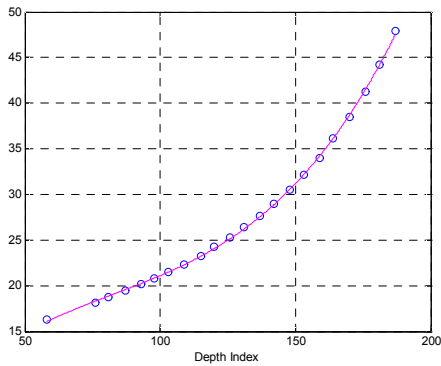


그림 7. 깊이 값과 변이 정보 간의 대응 함수
Fig. 7. Mapping curve between depth index and disparity.

깊이 카메라의 깊이 값을 다시점 카메라에서의 변이 정보로 각각 대응시킬 수 있고, 카메라의 위치 차이에 의해 발생하는 깊이의 오차를 보정할 수 있다.

깊이 영상을 위한 후처리의 다음 단계는 영상 내에 존재하는 잡음을 제거하는 것이다. 이상적인 깊이 영상은 일반적인 객체의 경계에서 객체와 배경의 깊이 값이 정확하게 구분되어야 한다. 하지만 깊이 카메라를 통해 획득된 깊이 영상은 객체 경계부에서 객체와 배경, 혹은 객체와 다른 객체의 깊이 값 사이에 중간 값을 가지는 화소들을 포함한다. 이러한 중간 값들은 3차원 투영(3D warping)시에 잘못된 위치의 화소로 투영시키는 오차를 유발할 수 있다.

따라서 객체 경계부의 중간 값들을 깨끗하게 정리하면서 주변의 평탄한 영역에 대한 필터링을 수행할 수 있는 방법으로 충격 필터(shock filter)를 사용할 수 있다^[8]. 그림 8(a)에 나타난 것처럼, 충격 필터는 점진적으로 변화하는 깊이의 불연속 구간을 계단 함수의 형태로

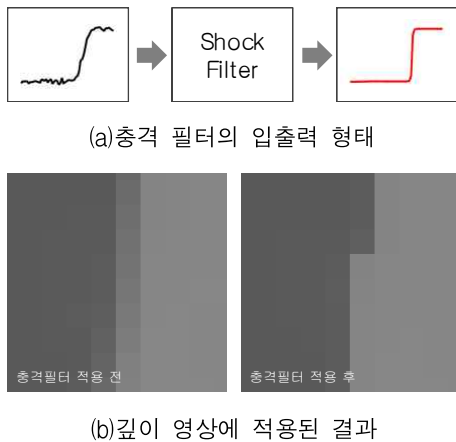


그림 8. 충격 필터의 적용 결과
Fig. 8. Result of shock filter.

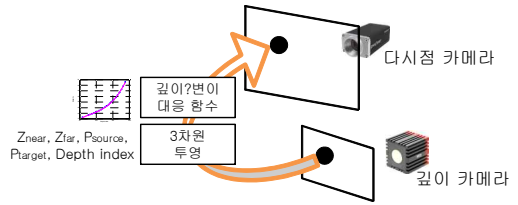


그림 9. 깊이 영상의 3차원 투영
Fig. 9. 3D warping of depth image.

변화시킨다. 또한 평탄한 영역에 존재하던 잡음도 깨끗하게 정리할 수 있다. 그림 8(b)는 충격 필터를 깊이 영상에 적용한 결과의 일부분을 보여준다.

충격 필터를 적용한 후, 다시점 깊이 영상에 대해 마지막으로 수행할 후처리는 3차원 투영이다. 깊이 카메라에서 촬영된 깊이 영상은 3차원 투영을 통해서 다시점 카메라의 위치로 투영한 후 사용되어야 한다.

그림 9는 3차원 투영의 개념을 보여준다. 다시점 카메라의 위치로 이동된 깊이 영상의 화소는 앞에서 구했던 깊이 값과 변이 정보의 대응 함수를 통해서 해당 화소에서의 변이 값으로 나타나게 된다. 깊이 카메라의 해상도는 매우 낮기 때문에 색상 영상의 위치로 투영된 깊이 영상, 즉 색상 영상 위치에서의 변이 영상은 깊이 투영되지 않은 영역을 많이 가지고 있다. 이러한 영역에 대해서는 값을 가지고 있는 주변 화소들로부터 현재 화소에서의 초기 변이 값을 예측하여 사용하여야 한다.

제안하는 방법에서는 각 시점마다 깊이 카메라를 사용하여 한 대의 깊이 카메라를 이용했던 기존의 방법들보다 좀 더 정확한 초기 깊이 정보를 획득할 수 있다.

2. 다시점 깊이 생성

깊이 영상의 3차원 투영을 통하여 다시점 영상의 위치에서 얻어진 변이 영상에는 정보가 없는 화소들이 다수 존재한다. 이 화소들의 변이 값 예측을 위하여 해당 화소를 기준으로 정사각형 형태의 탐색 영역을 사용한다. 탐색 영역 내부에 있는 화소에 따라서 다음과 같은 세 가지 방법으로 분류할 수 있다^[9].

첫 번째 경우는 탐색 영역 내에 변이 정보를 가지고 있는 모든 화소들의 변이 값이 전부 동일한 값을 가질 때, 두 번째 경우는 영역 내의 변이 값들이 서로 다른 값을 가질 때, 마지막 경우는 영역 내에 변이 값을 가진 화소가 존재하지 않을 때이다. 첫 번째 경우에는 동일하게 가지는 값으로, 두 번째 경우는 평균값으로 초기 변이를 결정하고, 세 번째 경우는 참조할만한 값이 없

으므로 탐색을 통해 탐색 블록 내의 화소간 차이의 절대합이 가장 작은 위치를 초기 변이로 결정한다.

이제 스테레오 정합의 비용 계산(cost calculation)을 위한 탐색 범위를 경우에 따라 각각 지정해 준다. 변이 값이 있는 화소의 경우 변이 값을 초기치로 설정한 후 매우 작은 탐색 범위 내에서 비용을 계산한다. 변이 값을 가지지 않는 화소의 경우에는 앞에서 언급한 세 가지 경우에 각각 다른 방법을 적용한다.

첫 번째 경우에는 탐색 영역 내의 동일한 변이 값이 현재 화소에도 적용된다고 판단하고 변이가 있는 화소와 같은 방법을 적용한다. 두 번째 경우에는 탐색 영역 내의 변이 값들의 평균을 현재 화소의 변이 값으로 판단하고, 첫 번째 경우보다는 긴 탐색 범위 내에서 비용을 계산한다. 세 번째 경우에는 참조할 수 있는 변이 값이 없으므로 장면의 최소 및 최대 변이 값을 탐색 범위로 설정하여 비용 계산을 수행한다.

또한 첫 번째 경우와 두 번째 경우에는 초기 깊이 값과 현재 탐색 위치에서 가지는 변이 값의 차를 식 (1)의 형태로 만들어 에너지 함수에 TOF 항으로 포함시킨다. 그러면 기존에 계산된 비용 뿐만 아니라 깊이 카메라로부터 온 정보를 함께 이용하여 에너지 함수를 식 (2)와 같이 정의하고, 신뢰 확산(belief propagation) 최적화 방법을 이용하여 장면의 최종 깊이 정보를 구할 수 있다^[10]. 식에서 p, q, ξ 는 각각 현재 위치, 주변 위치, 탐색 범위 내에서의 변이 정보를 의미하며, $d_{p,i}$ 는 p 화소에서의 초기 변이 값이다.

$$T_p(f_p) = |f_p - d_{p,i}| \quad (1)$$

$$E(f) = \sum_p D_p(f_p) + \sum_{p,q} S(f_p, f_q) + \sum_p T_p(f_p) \quad (2)$$

III. 실험 결과

제안한 방법의 실험을 위해서 그림 1과 같이 배열된 카메라를 사용하였으며, 그림 2, 그림 10과 같이 촬영된 영상을 실험에 사용하였다. 색상 영상의 해상도는 640x480이며, 카메라로부터 장면의 가장 먼 곳까지의 직선거리는 약 2.5m 이다. 실험에 사용된 변수들로 영역 탐색에 5x5 크기의 탐색 윈도우를 사용하였고, 짧은 탐색구간과 긴 탐색구간은 각각 현재 화소가 투영된 위치로부터 각각 전후로 두 개의 화소와 초기 변이 값의 차에 해당하는 화소만큼을 이용하였다.



그림 10. 실험에 사용된 영상
Fig. 10. Images for testing.



(a)깊이 카메라 정보 없이 얻은 결과



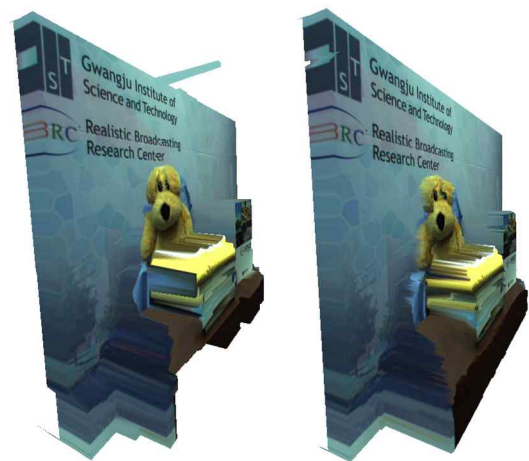
(b)제안한 방법으로 얻은 결과



(c)깊이 카메라 정보 없이 얻은 결과



(d)제안한 방법으로 얻은 결과



(e)3차원 모델 생성: 기존의 방법(왼쪽)과 제안한 방법(오른쪽)

그림 11. 깊이 생성 결과
Fig. 11. Results of depth generation.

이제 앞에서 언급한 후처리 방법과 깊이 생성 방법을 적용하여 얻은 결과가 그림 11에 나타나 있다. 깊이 카메라의 영상 없이 신뢰 확산 방법을 이용하여 장면의 깊이 정보를 구했을 때에는(그림 11(a)와 그림 11(c)) 무늬가 없는 영역이나 물체에 의해 가려진 부분 등에서 정확한 깊이 정보를 획득할 수 없었지만, 제안한 방법의 결과(그림 11(b)와 그림 11(d))에서는 장면의 정확한 깊이 정보를 얻을 수 있었다.

그림 11(e)는 기존의 방법과 제안한 방법을 사용하여 구해진 깊이 영상과 색상 영상을 이용하여 3차원 장면을 만든 것을 보여준다. 제안한 방법의 결과로 얻어진 모델이 기존의 방법보다 객체와 객체 간 깊이를 잘 표현해주고 있으며, 색상이 일정한 배경 영역 등에서도 안정적으로 깊이를 표현하는 것을 볼 수 있다.

IV. 결 론

본 논문에서는 다시점 카메라와 다시점 깊이 카메라를 이용하여 획득된 영상에 대한 후처리 방법과 깊이 생성 방법을 설명하였다. 깊이 카메라의 영상은 잡음과 왜곡이 심하고, 색상 카메라와의 상관도도 매우 떨어지기 때문에 촬영된 영상의 후처리 과정을 통해서 그 품질을 개선할 필요성이 있다. 본 논문에서는 촬영된 깊이 영상의 왜곡과 잡음을 제거하였고, 색상 영상 위치에서의 변이 정보로 변환하였다. 이 정보를 기반으로 구해진 장면의 깊이 영상은 깊이 카메라에서 얻은 정보 없이 깊이를 획득하는 방법보다 우수한 성능을 나타내었음을 볼 수 있었다. 제안한 방법은 다시점 영상 및 깊이 영상의 촬영과 처리, 그리고 3차원 입체 영상의 제작에 활용될 수 있을 것이다.

참 고 문 헌

- [1] A. Smolic and P. Kauff, "Interactive 3D Video Representation and Coding Technologies," Proceedings of the IEEE, Spatial Issue on Advances in Video Coding and Delivery, vol. 93, no. 1, pp. 99-110, Jan. 2005.
- [2] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 231-236, June 2008.
- [3] B. Bartczak and R. Koch, "Dense Depth Maps from Low Resolution Time-of-Flight Depth and High Resolution Color Views," Proc. of 5th International Symposium on Visual Computing, pp. 1-12, Nov. 2009.
- [4] <http://www.vision.caltech.edu/bouguetj>, Camera Calibration Toolbox for MATLAB.
- [5] Y.S. Kang and Y.S. Ho, "Geometrical Compensation for Multi-view Video in Multiple Camera Array," Proc. of Int'l Symposium on Electronics and Marine, pp. 83-86, Sept. 2008.
- [6] A. Ilie and G. Welch, "Ensuring color consistency across multiple cameras," Proc. of IEEE international Conference on Computer Vision, pp. II: 1268-1275, Oct. 2005.
- [7] A. Wang, T. Qiu, and L. Shao, "A Simple Method of Radial Distortion Correction with Centre of Distortion Estimation," Journal of Mathematical Imaging and Vision, vol. 35, no. 3, pp. 165-172, July 2009.
- [8] G. Gilboa, N. Sochen, and Y.Y. Zeevi, "Regularized Shock Filters and Complex Diffusion", ECCV 2002, LNCS 2350, pp. 399-313, May 2002.
- [9] Y.S. Ho and Y.S. Kang, "Multi-view Depth Generation using Multi-Depth Camera System," Proc. of International Conference on 3D Systems and Application (3DSA), pp. 1-4, May 2010.
- [10] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," International Journal of Computer Vision, vol. 70, no. 1, pp. 41-54, Oct. 2006.

저 자 소 개



강 운 석(학생회원)
2007년 한국항공대학교
전자공학과 학사
2008년 광주과학기술원 정보통신
공학과 석사
2008년~현재 광주과학기술원
정보통신공학부 박사과정
<주관심분야 : 디지털 영상처리, 다시점 영상
획득 및 처리, 3차원 TV, 실감방송>



호 요 성(정회원)
대한전자공학회 논문지
제46권, SP편, 제1호 참조