

특집논문-08-13-5-04

복합형 카메라 시스템에서 관심영역이 향상된 고해상도 깊이맵 생성 방법

김 성 열^{a)}, 호 요 성^{a)‡}Generation of ROI Enhanced High-resolution Depth Maps
in Hybrid Camera SystemSung-Yeol Kim^{a)} and Yo-Sung Ho^{a)‡}

요 약

본 논문은 저해상도의 깊이 카메라와 고해상도의 양안식 카메라를 결합한 복합형 카메라 시스템에서 관심영역(region of interest, ROI)이 향상된 깊이맵을 생성하는 새로운 방법을 제안한다. 제안하는 방법은 깊이 카메라로 획득한 깊이 정보를 3차원 워핑(warping)하여 좌영상의 ROI 깊이맵을 생성한다. 그런 다음, 양안식 카메라로 획득한 좌우영상의 배경 영역을 스테레오 정합하여 좌영상의 배경 깊이맵을 생성한다. 최종적으로, ROI 깊이맵과 배경 깊이맵을 결합하여 최종 깊이맵을 생성한다. 제안하는 방법으로 생성한 고해상도 깊이맵은 기존의 스테레오 정합 방법보다 ROI에 정확한 깊이 정보를 제공한다.

Abstract

In this paper, we propose a new scheme to generate region-of-interest (ROI) enhanced depth maps in the hybrid camera system, which is composed of a low-resolution depth camera and a high-resolution stereoscopic camera. The proposed method creates an ROI depth map for the left image by carrying out a three-dimensional (3-D) warping operation onto the depth information obtained from the depth camera. Then, we generate a background depth map for the left image by applying a stereo matching algorithm onto the left and right images captured by the stereoscopic camera. Finally, we merge the ROI map with the background one to create the final depth map. The proposed method provides higher quality depth information on ROI than the previous methods.

Keyword : Depth Map Generation, Depth Camera, Stereo Matching, Hybrid Camera System

1. 서 론

최근에 화소 단위로 색상 영상과 깊이 정보를 결합한 깊이 비디오(video-plus-depth)^[1]가 3차원 비디오를 표현하는 방법으로 각광 받고 있다. 차세대 3차원 응용에서 고품질의

3차원 비디오를 제공하기 위해, 정확한 깊이 정보를 측정하는 것은 매우 중요하다. 깊이 측정법은 크게 간접 깊이 측정법과 직접 깊이 측정법으로 나뉜다. 간접 깊이 측정법은 양안식 또는 다시점 카메라로 획득한 2차원 영상으로부터 깊이 정보를 간접적으로 측정한다. 스테레오 정합^[2]은 간접 깊이 측정법의 대표적인 예이다.

반면에, 직접 깊이 측정법은 물리적인 센서를 이용하여 깊이 정보를 직접적으로 측정한다. 구조광 패턴 방법^[3]과 깊이 카메라 방법^[4]이 직접 깊이 측정법에 포함된다. 특히,

a) 광주과학기술원

Gwangju Institute of Science and Technology

‡ 교신저자: 호요성(hoyo@gist.ac.kr)

※ 본 연구는 광주과학기술원 실감방송연구센터(RBRC)를 통한 정보통신 부대학IT연구센터(ITRC)의 지원에 의한 것입니다.

비행시간(time-of-flight, TOF) 원리에 기반 한 깊이 카메라는 실내에서 깊이 정보를 측정해야 하는 제약이 있지만, 실제 장면의 깊이맵을 실시간으로 제공한다.

최근에 간접 깊이 측정법과 TOF 깊이 카메라를 동시에 이용한 새로운 깊이 측정법^[5,6]이 제안되었다. 그러나 이 복합형 카메라 시스템은 저해상도의 깊이 카메라에 의존하여 깊이 정보를 획득하였기 때문에, 고해상도의 깊이맵을 제공하지 못했다. 본 논문에서는 고해상도 양안식 카메라와 저해상도 TOF 깊이 카메라를 결합하여 고해상도의 깊이맵을 생성하는 방법을 제안한다.

일반적으로 사람들은 3차원 장면의 배경보다 관심영역(region-of-interest, ROI)에 초점을 맞추기 때문에, ROI에 보다 정확한 깊이 정보가 요구된다. 현재의 깊이 카메라는 실물의 깊이 정보를 측정할 수 있는 측정 거리가 짧기 때문에, 깊이 카메라로 획득한 깊이 정보를 ROI의 깊이 정보로 간주하여 ROI에 대한 깊이 정보의 정확도를 높인다.

본 논문의 2장에서는 제안하는 복합형 카메라 시스템을 소개하고, 3장에서는 복합형 카메라를 이용하여 깊이맵을 생성하는 방법을 설명한다. 4장에서는 깊이맵 생성에 대한 실험 결과를 보여주고, 5장에서 본 논문의 결론을 맺는다.

II. 복합형 카메라 시스템

그림 1에서 보듯이, 제안하는 복합형 카메라 시스템은 고해상도의 양안식 카메라와 저해상도의 깊이 카메라로 이루어진다. 제안하는 복합형 카메라 시스템은 프레임마다 양안식 카메라에서 획득한 좌우영상과 깊이 카메라로부터 획득한 색상영상 및 깊이맵을 제공한다.

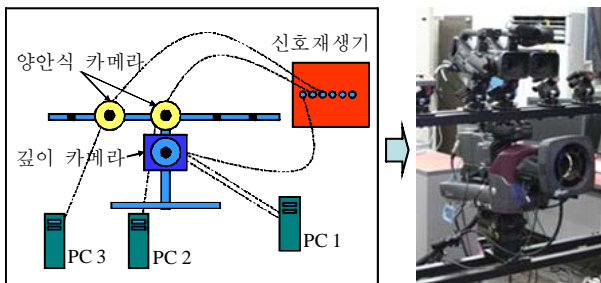


그림 1. 복합형 카메라 시스템
Fig 1. Hybrid camera system

본 논문은 복합형 카메라 시스템을 이용하여 양안식 카메라의 좌영상에 대한 깊이맵을 생성한다. 그림 2는 제안하는 깊이맵 생성 방법에 대한 전체 흐름을 보여준다. 전처리 단계로서 깊이 카메라 정보에 대한 양안식 카메라의 상대적인 카메라 정보를 계산하고, 좌우영상을 영상 정렬화(image rectification)한다. 깊이맵 생성단계에서는 깊이 카메라로 획득한 깊이 정보를 3차원 워핑(3-D warping)하여 세계 좌표로 이동시키고, 워핑된 깊이 정보를 다시 3차원 워핑하여 좌측 카메라로 투영해 ROI 깊이맵을 생성한다. 또한, 좌영상의 배경 영역을 우영상과 스테레오 정합하여 좌영상의 배경 깊이맵을 생성한다. 최종적으로, ROI 깊이맵과 배경 깊이맵을 결합하여 최종 깊이맵을 생성한다.

III. ROI 향상 깊이맵 생성

1. 전처리 과정

깊이맵을 생성하기 위해, 우선 카메라 보정 알고리즘^[7]을 이용하여 깊이 카메라, 좌측 카메라, 우측 카메라의 내부 행렬, 회전 행렬, 이동 행렬을 각각 추정한다. 그런 다음, 영상 정렬화 알고리즘^[8]을 이용하여 좌우영상의 공선이 서로 평행하도록 교정한다. 양안식 카메라의 내부 행렬 K'_l 과 K'_r , 회전행렬 R'_l 와 R'_r , 이동행렬 t_l, t_r 을 이용하여 양안식 카메라의 투영 행렬 \tilde{P}_l, \tilde{P}_r 은 식 1과 같이 구한다.

$$\begin{aligned} \tilde{P}_l &= K'_l [R'_l | t_l] \\ \tilde{P}_r &= K'_r [R'_r | t_r] \end{aligned} \quad (1)$$

그런 다음, 깊이 카메라의 회전 행렬의 역행렬 R_s^{-1} 과 이동 행렬 t_s 를 이용하여 깊이 카메라에 대한 양안식 카메라의 상대적 카메라 정보를 식 2와 같이 구한다.

$$\begin{aligned} P'_s &= K_s [I | O] \\ \tilde{P}'_l &= K'_l [R'_l R_s^{-1} | t_l - t_s] \\ \tilde{P}'_r &= K'_r [R'_r R_s^{-1} | t_r - t_s] \end{aligned} \quad (2)$$

여기서, $P'_s, \tilde{P}'_l, \tilde{P}'_r$ 는 깊이 카메라, 좌측 카메라, 우측

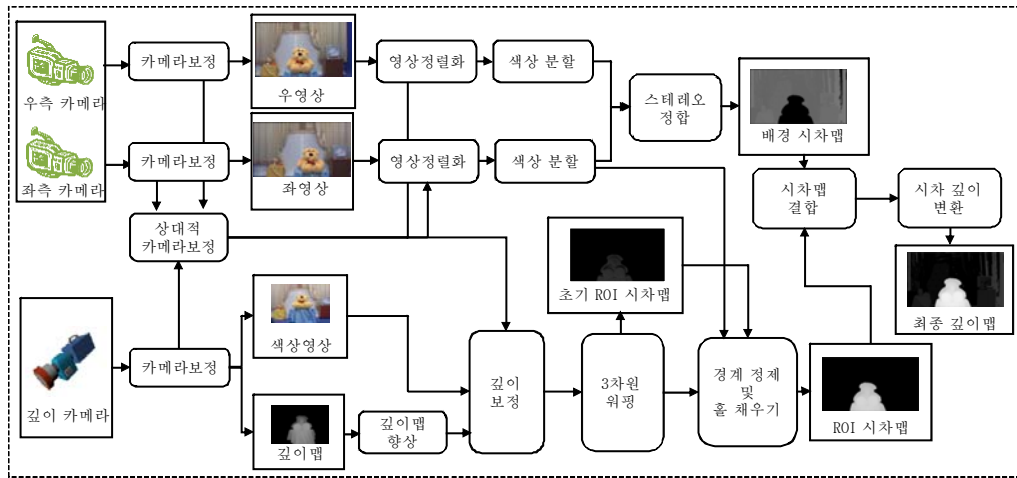


그림 2. 제안하는 깊이맵 생성에 대한 전체 흐름도
 Fig 2. Overall framework of the proposed depth map creation

카메라의 상대적 투영 행렬이다.

일반적으로, 깊이 카메라로 측정된 깊이 거리와 실제 거리에 오차가 있다. 이 오차를 보정하기 위해, 우선 깊이맵 향상법^[9]을 이용하여 깊이맵에 포함되어 있는 잡음을 제거한다. 그런 다음, 그림 3에서 보듯이, 좌우영상의 ROI에 대한 최대 시차 d_{Max} 와 최소 시차 d_{Min} 을 영상 편집 도구를 이용하여 구하고, 깊이 카메라에서 ROI까지의 최소 거리 D_{Min} 와 최대 거리 D_{Max} 를 식 3과 같이 구한다.

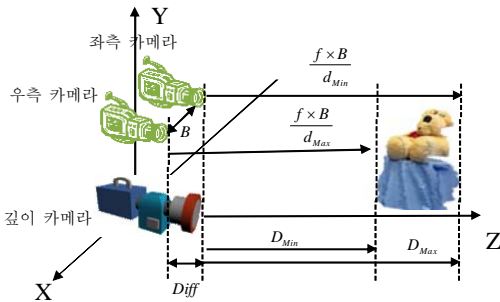


그림 3. 깊이 카메라의 깊이 정보 보정
 Fig 3. Depth data correction of depth camera

$$D_{Max} = \frac{f \times B}{d_{Min}} - Diff$$

$$D_{Min} = \frac{f \times B}{d_{Max}} - Diff$$

(3)

여기서, f 와 B 는 좌측 카메라의 초점거리와 수평축 거리

이고, $Diff$ 는 좌측 카메라와 깊이 카메라의 거리 차이이다.

깊이맵 화소값과 실제 거리 사이의 관계는 식 4와 같다.

$$Z(i, j) = D_{Min} + ((255 - D(i, j)) - H_{Min}) \times \frac{D_{Max} - D_{Min}}{H_{Max} - H_{Min}} \quad (4)$$

여기서, $Z(i, j)$ 는 깊이맵에서 화소 위치 (i, j) 의 화소값 $D(i, j)$ 에 대해 보정된 실제 거리이다. H_{Max} 와 H_{Min} 은 깊이맵의 화소값에 대한 히스토그램 최대값과 최소값이다. 결과적으로 $(D_{Max} - D_{Min}) / (H_{Max} - H_{Min})$ 는 깊이맵의 화소값 단위 변화량에 따른 거리 단위 변화량이다.

2. ROI 깊이맵 생성

깊이맵 정보를 3차원 위평하여 좌영상의 ROI에 대한 깊이 정보를 얻는다. 깊이맵의 화소 위치 (p_{sx}, p_{sy}) 에 상응하는 화소값이 $D_s(p_{sx}, p_{sy})$ 일 때, 3차원 화소 $p_s(p_{sx}, p_{sy}, D_s(p_{sx}, p_{sy}))$ 를 3차원 위평하여 세계 좌표로 보낸다. p_s 에 상응하는 좌영상의 화소 위치 (p_{lx}, p_{ly}) 는 식 5에서 얻는다.

$$p_l = \tilde{P}_l^T \cdot P_s^{-1} \cdot p_s \quad (5)$$

또한, 좌영상의 화소 위치 (p_{lx}, p_{ly}) 에 대한 깊이 정보 D_l

(p_{lx}, p_{ly}) 는 식 6에 의해 계산한다.

$$D_l(p_{lx}, p_{ly}) = Diff + D_s(p_{sx}, p_{sy}) \quad (6)$$

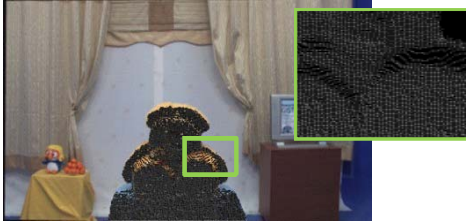


그림 4. 3차원 워핑하여 생성한 초기 ROI 시차맵
Fig 4. Initial ROI depth map created by 3-D warping

3차원 워핑으로 획득한 좌영상의 깊이 정보를 시차로 변환하여 초기 ROI 시차맵을 생성한다. 그림 4는 초기 ROI 시차맵을 좌영상과 포갠 영상을 보여준다. 3차원 워핑 후, 저해상도의 깊이맵이 고해상도의 좌영상으로 확장되었고, 홀(hole)이 생겼다. 또한, 초기 ROI 깊이맵의 깊이 정보와 좌영상의 색상 정보가 ROI의 경계에서 일치하지 않았다.

본 논문은 좌영상과 초기 ROI 깊이맵의 불일치를 보정하기 위해서 식 7을 이용하여 경계를 정제한다.

$$Y(s_i) = \begin{cases} 1, & \text{if } \frac{n(A(s_i))}{n(s_i)} \geq 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

여기서, $Y(s_i)$ 는 i 번째 분할영상 s_i 가 ROI에 포함되는지에 대한 여부를 나타낸다. $n(s_i)$ 는 분할영상 s_i 에 포함된 화소 수이고, $n(A(s_i))$ 는 분할영상 s_i 에 포함된 초기 ROI 깊이맵 영역 $A(s_i)$ 의 화소 수이다.

ROI에 포함된 분할영상에서 초기 ROI 깊이맵에 발생한 홀을 식 8을 이용하여 채운다.

$$R(x, y)_k = \frac{1}{n} \cdot \sum_{i=-\lfloor W/2 \rfloor}^{\lfloor W/2 \rfloor} \sum_{j=-\lfloor W/2 \rfloor}^{\lfloor W/2 \rfloor} R(x+i, y+j)_k \quad (8)$$

여기서, $R(x, y)_k$ 는 k 번째 분할영상에 속한 이웃 화소 위치 $(x+i, y+j)$ 의 화소 값 $R(x+i, y+j)_k$ 를 이용하여 k 번

째 분할영상의 홀 위치 (x, y) 에서 채울 깊이 정보이다. n 은 $W \times W$ 창안에서 k 번째 분할영상에 속한 홀이 아닌 유효한 이웃 화소의 개수이다. 홀 채우기 알고리즘은 분할영상 단위로 이루어져, 이웃 분할영상의 깊이 정보가 현재 분할영상에 존재하는 홀을 채우는데 영향을 주지 않는다.

ROI 시차맵을 생성한 후, 스테레오 정합 기술을 이용하여 배경영역에 대한 배경 시차맵을 생성한다. 배경영역은 좌영상의 색상 분할 결과에서 ROI에 대한 분할영상 집합을 제거한 영역이다. 배경 시차맵을 획득한 후, ROI 시차맵과 배경 시차맵을 결합하여 최종 시차맵을 생성하고, 시차맵 깊이 정보로 변환하여 최종 깊이맵을 얻는다.

IV. 실험 결과

제안한 방법을 평가하기 위해, HD 카메라 2대와 Z-Cam을 결합한 그림 1의 복합형 카메라 시스템을 구성하였다. 깊이 카메라의 깊이 정보 측정 거리는 175cm에서 215cm이었고, 양안식 카메라의 수평축 거리는 20cm이었다.

복합형 카메라 시스템으로 실험 영상인 BEAR 영상과 ACTOR 영상을 획득했다. 실험 영상의 좌우영상 해상도는 1920×1080이었고, 깊이 카메라의 색상영상과 깊이맵 해상도는 720×486이었다. 그림 5는 실험 영상을 보여준다.

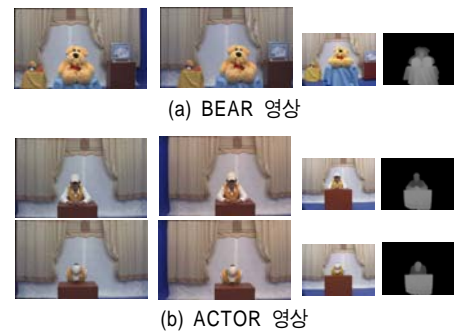


그림 5. 실험 영상
Fig 5. Test images

본 실험에서 스테레오 정합 알고리즘인 SAD, 신뢰 전달(belief propagation, BP)^[10], 그래프 컷(graph cut, GC)^[2], 동적 프로그래밍(dynamic programming, DP)^[2], 주사선 최적화(scan-line optimization, SO)^[2]방법을 이용하여 깊이맵을 생

성하였다. 또한, 제안한 방법의 성능을 객관적으로 평가하기 위해, 3차원 스캐너^[11]를 이용하여 BEAR 영상의 ROI에 대한 실제 깊이맵을 제작하였다.

그림 6은 제안한 방법과 기존의 스테레오 정합 방법을 이용하여 생성한 최종 깊이맵을 보여준다. 표 1은 제안한 방법으로 실험영상의 깊이맵을 생성하는데 소요된 시간을 보여준다. 깊이맵 품질을 객관적으로 평가하기 위해, 실제 ROI 깊이맵과 생성한 깊이맵을 비교하여, 근 평균 제곱 (root-mean square, RMS) 오류 R_E ^[2]와 정합 오류율 (bad matching percentage) B_A 를 분석하였다. 표 2는 RMS 오류 R_E 와 정합 오류율 B_A 를 보여준다.

표 1. 깊이맵 생성 처리 속도
Table 1. Processing time for depth map creation

구분	실험영상		
	BEAR	ACTOR 1	ACTOR 2
색상분할	4.81초	4.97 초	5.24 초
ROI 시차맵	9.57 초	9.09 초	8.74 초
배경 시차맵	65.61 초	27.13 초	24.78 초
총 처리시간	79.99 초	41.19 초	38.76 초

표 2. BEAR 영상에 대한 깊이 정보 품질 비교
Table 2. Comparison of depth quality for BEAR image

구분	R_E	R_{Diff}	B_A	B_{Diff}
SAD	60.5	36.1	85.3%	46.4%
BP	26.5	2.1	50.1%	11.2%
GC	62.1	37.7	83.3%	44.4%
DP	46.1	21.7	76.7%	37.8%
SO	67.7	43.3	93.5%	40.6%
제안한 방법	24.4	-	38.9%	-

표 2에서 R_{Diff} 는 제안한 방법과 스테레오 정합 방법들 간의 RMS 오류 R_E 차이이고, B_{Diff} 는 제안한 방법과 스테레오 정합 방법들 간의 정합 오류율 B_A 차이이다. BEAR 영상에서 제안한 방법으로 생성한 ROI 깊이 정보가 스테레오 정합 방법 중 성능이 가장 우수한 BP 방법보다 RMS 오류 R_E 에서 약 2.1, 정합 오류율 B_A 에서 약 11.2% 더 정확한 것으로 판명되었다.

깊이맵의 품질을 주관적으로 평가하기 위해, 그림 7에서 보듯이, BEAR 영상의 ROI 깊이 정보를 이용하여 3차원 장면을 모델링하였다. 제안한 방법의 ROI를 3차원 장면 모델링한 결과가 스테레오 정합을 이용하여 얻어진 ROI를 3차원 장면 모델링한 결과보다 실제 ROI를 모델링한 결과와 유사하였다. 또한, 그림 8에서 보듯이, 제안한 방법으로 생성한 깊이 정보를 이용하여 가상 뷰를 생성하면 기존의 스테레오 정합 방법보다 품질이 좋은 가상 뷰를 얻었다.

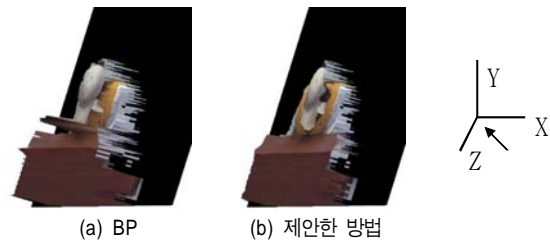


그림 8. 가상 뷰 생성
Fig 8. Virtual view generation

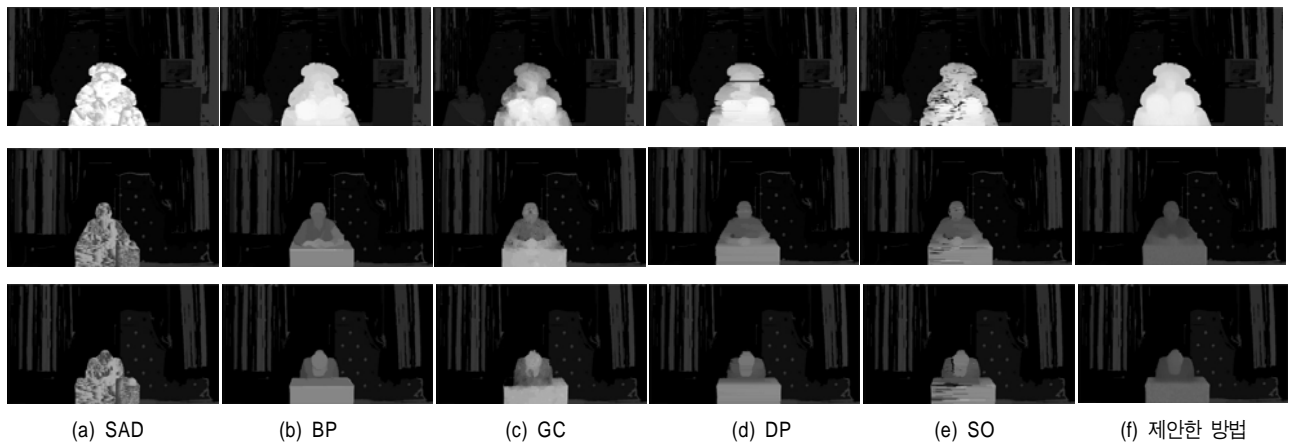


그림 6. 깊이맵 생성 결과
Fig 6. Result of depth map creation

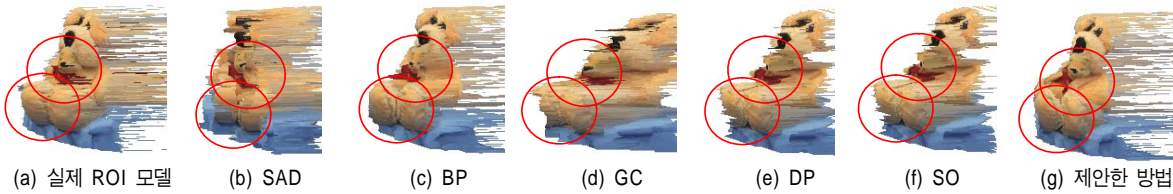


그림 7. 3차원 장면 모델링 결과
Fig 7. Result of 3-D scene modeling

V. 결론

본 논문은 저해상도의 TOF 깊이가 카메라와 고해상도의 양안식 카메라를 결합하여 ROI가 향상된 깊이맵을 생성하는 방법을 제안하였다. 실험 결과에서 알 수 있듯, 제안한 방법이 객관적 평가와 주관적 평가에서 기존의 스테레오 정합 방법보다 정확한 ROI 깊이를 제공하였다.

참고 문헌

[1] C. Fehn, "A 3D-TV System Based on Video Plus Depth Information," Proc. of Asilomar Conference on Signals, Systems and Computers, vol. 2, pp. 1529-1533, 2003.

[2] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms," International Jour. of Computer Visioin, vol. 47, no. 1-3, pp. 7-42, 2002.

[3] M. Waschbüsch, S. Würmlin, D. Cotting, and M. Gross, "Point-sampled 3D Video of Real-world Scenes," Signal Processing Image Communication, vol. 22, no. 2, pp. 203-216, 2007.

[4] G.J. Iddan and G. Yahav, "3D Imaging in the Studio and Elsewhere," Proc. of SPUE Videometrics and Optical Methods for 3D Shape Measurements, pp. 48-55, 2001.

[5] G. Um, K.Y. Kim, C. Ahn, and K.H. Lee, "Three-dimensional Scene Reconstruction Using Multi-view Images and Depth Camera," Proc. of SPIE Stereoscopic Displays and Virtual Reality Systems XII, vol. 5664, pp. 271-280, 2005.

[6] J. Diebel and S. Thrun, "An Application of Markov Random Fields to Range Sensing," Proc. of Advances in Neural Information Processing systems, pp. 291-298, 2005.

[7] Camera calibration toolbox program for Matlab provided by Caltech, http://www.vision.caltech.edu/bouguetj/calib_doc/.

[8] Y.S. Kang, C. Lee, Y.S. Ho, "An Efficient Rectification Algorithm for Multi-view Images in Parallel Camera Array," Proc. of 3DTV Conference, pp. 61-64, 2008.

[9] J.H. Cho, I.Y. Chang, S.M Kim, and K.H. Lee, "Depth Image Processing Technique for Representing Human Actors in 3DTV Using Single Depth Camera," Proc. of 3DTV conference, paper no. 15, 2007.

[10] P. F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," International Jour. of Computer Vision, vol. 70, no. 1, pp. 41-54, 2006.

[11] LMS-Z390i, <http://www.riegl.com/>.

저 자 소 개



김 성 열

- 2001년 2월 : 강원대학교 정보통신공학과 (학사)
- 2003년 2월 : 광주과학기술원 정보통신공학과 (석사)
- 2003년 8월 : 광주과학기술원 정보통신공학과 (박사)
- 2007년 12월 ~ 2008년 8월 : 한국전자통신연구원 실감방송시스템연구팀 위촉연구원
- 2007년 9월 ~ 현재 : 광주과학기술원 정보기전공학부 박사후과정
- 주관심분야 : 깊이영상기반 모델링, 3차원 메쉬 표현 및 압축, 3차원 TV, 실감방송



호 요 성

- 1981년 2월 : 서울대학교 전자공학과 (학사)
- 1983년 2월 : 서울대학교 전자공학과 (석사)
- 1983년 3월 ~ 1995년 9월 : 한국전자통신연구소 선임연구원
- 1989년 12월 : University of California, Santa Barbara Department of Electrical and Computer Engineering (박사)
- 1990년 1월 ~ 1993년 5월 : 미국 Philips 연구소 Senior Research Member
- 1995년 9월 ~ 현재 : 광주과학기술원 정보통신공학과 교수
- 주관심분야 : 디지털 신호처리, 영상신호처리 및 압축, 멀티미디어 시스템, 디지털 TV와 고선명 TV, MPEG 표준, 3차원 TV, 실감방송