

특집논문-04-09-3-01

# 깊이정보 카메라 및 다시점 영상으로부터의 다중깊이맵 융합기법

엄기문<sup>a)\*</sup>, 안충현<sup>a)</sup>, 이수인<sup>a)</sup>, 김강연<sup>b)</sup>, 이관행<sup>b)</sup>

## Multi-Depth Map Fusion Technique from Depth Camera and Multi-View Images

Gi-Mun Um<sup>a)\*</sup>, Chunghyun Ahn<sup>a)</sup>, Soo In Lee<sup>a)</sup>, Kang Yeon Kim<sup>b)</sup> and Kwan H. Lee<sup>b)</sup>

### 요 약

본 논문에서는 정확한 3차원 장면복원을 위한 다중깊이맵 융합기법을 제안한다. 제안한 기법은 수동적 3차원 정보획득 방법인 스테레오 정합기법과 능동적 3차원 정보획득 방법인 깊이정보 카메라로부터 얻어진 다중깊이맵을 융합한다. 전통적인 두 개의 스테레오 영상 간에 변이정보를 추정하는 전통적 스테레오 정합기법은 차폐 영역과 텍스처가 적은 영역에서 변이 오차를 많이 발생한다. 또한 깊이정보 카메라를 이용한 깊이맵은 비교적 정확한 깊이정보를 얻을 수 있으나, 잡음이 많이 포함되며, 측정 가능한 깊이의 범위가 제한되어 있다. 따라서 본 논문에서는 이러한 두 기법의 단점을 극복하고, 상호 보완하기 위하여 이 두 기법에 의해 얻어진 다중깊이맵의 변이 또는 깊이값을 적절하게 선택하기 위한 깊이맵 융합기법을 제안한다. 3-시점 영상으로부터 가운데 시점을 기준으로 좌우 영상에 대해 두 개의 변이맵들을 각각 얻으며, 가운데 시점 카메라에 설치된 깊이정보 카메라로부터 얻어진 깊이맵들 간에 위치와 깊이값을 일치시키기 위한 전처리를 행한 다음, 각 화소 위치의 텍스처 정보, 깊이맵 분포 등에 기반하여 적절한 깊이값을 선택한다. 제안한 기법의 컴퓨터 모의실험 결과, 일부 배경 영역에서 깊이맵의 정확도가 개선됨을 볼 수 있었다.

### Abstract

This paper presents a multi-depth map fusion method for the 3D scene reconstruction. It fuses depth maps obtained from the stereo matching technique and the depth camera. Traditional stereo matching techniques that estimate disparities between two images often produce inaccurate depth map because of occlusion and homogeneous area. Depth map obtained from the depth camera is globally accurate but noisy and provide a limited depth range. In order to get better depth estimates than these two conventional techniques, we propose a depth map fusion method that fuses the multi-depth maps from stereo matching and the depth camera. We first obtain two depth maps generated from the stereo matching of 3-view images. Moreover, a depth map is obtained from the depth camera for the center-view image. After preprocessing each depth map, we select a depth value for each pixel among them. Simulation results showed a few improvements in some background regions by proposed fusion technique.

Keywords : 3D scene reconstruction, depth map

a) 한국전자통신연구원 디지털방송연구단 방송시스템연구그룹  
ETRI Digital Broadcasting Research Division, Broadcasting System Research Group

b) 광주과학기술원 기전공학과 GIST, Department of Mechatronics

\* 본 연구는 정보통신부 출연 "지능형 통합정보 방송(SmarTV) 기술 개발" 사업의 일환으로 수행되었습니다.

## I. 서 론

최근 3차원 장면복원(3D scene reconstruction) 분야의 연구에 있어 3차원 깊이 센서(3D depth sensors) 개발에 대한 연구가 특히 많이 이뤄지고 있다. 이 3차원 깊이 센서

는 능동 방식과 수동 방식의 크게 두 가지로 나눌 수 있다<sup>[1][2]</sup>. 첫 번째 방식인 능동 3차원 깊이 센서(active 3D depth sensors)는 능동 조사(active illumination)기법을 사용하며, 신뢰성 높은 깊이정보를 제공한다. 레이저(laser)나 구조화된 광선(structured light)기법을 이용한 제품들에는 미국의 Cyberware, 영국의 Wicks & Wilson사의 제품들이 있는데, 이들 시스템들은 모두 주로 정지된 객체나 장면의 3차원 정보 획득에 사용되며, 일반적 비디오 프레임 비율(video frame rate)로 움직이는 3차원 장면을 획득하는 것은 불가능하다. 그러나 최근 이스라엘의 3DV systems사에서는 빛의 비행시간(time-of-flight of light)에 기초한 깊이 측정기법을 사용하여 비디오 프레임 비율로 깊이정보를 취득할 수 있는 방송용 카메라 모듈인 Z-cam<sup>TM</sup>을 발표하였다<sup>[3]</sup>. 그러나 이러한 능동 센서를 이용한 제품들은 가격이 비싸며, 취득된 데이터 량이 너무 크고, 제한된 환경 하에서만 이용가능하다는 단점이 있다.

반면에 수동 3차원 깊이 센서(passive 3D depth sensors)는 카메라로부터 취득된 영상들로부터 3차원 정보를 취득하며, 값이 저렴하고 깊이정보와 함께 텍스처 정보를 같이 얻을 수 있다는 장점이 있는 반면, 차폐(occlusion), 광량 변화(illumination changes)가 발생하는 영역이나 텍스처가 적은(low texture) 영역에서는 오차를 많이 발생하며, 또한 화소 단위로 연산을 수행하므로 수행시간이 오래 걸린다는 단점이 있다. 수동 3차원 깊이 센서의 대표적인 기법으로는 동일 시간대의 서로 다른 시점을 가진 두 카메라로부터 얻어진 영상을 이용하는 스테레오(Shape from Stereo)기법, 실루엣 기반 기법(Shape from Silhouette), 카메라 이동에 의해 촬영된 여러 시점의 정적 객체에 대한 3차원 정보를 계산하는 움직임기반 모양 추정기법(Shape from Motion (SFM)) 등이 있다. 표 1은 지금까지 설명한 능동 3차원 깊이 센서와 수동 3차원 깊이 센서들의 장단점을 비교하여 보여주고 있다.



그림 1. 3시점 및 깊이정보 카메라 시스템  
Fig. 1. 3-view and depth camera system

본 논문에서는 그림 1과 같이 구성된 3-시점 SD급 카메라 및 깊이정보 카메라(depth camera)로부터 얻어지는 2개의 변이맵(disparity map) 및 깊이맵(depth map)을 융합(fusion)하여 보다 향상된 가운데 카메라 시점에 대한 깊이맵을 획득하는 다중깊이맵 융합기법을 제안한다.

본 논문의 구성은 먼저 2장에서 기존의 복합적 깊이맵 획득기법들에 대해 간략하게 소개하고, 3장에서 제안한 깊이맵 융합기법에 대해 기술한다. 또한 4장에서는 제안한 기법의 실험 결과를 보이며, 5장에서 결론을 맺는다.

## II. 기존의 복합적 깊이맵 획득기법

기존의 능동 또는 수동 3차원 깊이 센서를 이용한 깊이맵 획득기법들은 앞장에서 살펴본 바와 같이 서로 장단점을 가지고 있다. 이러한 두 가지 방식의 단점을 극복하기 위한 기법으로서 최근에는 여러 가지 3차원 깊이 센서로 구해진 데이터들을 융합하거나, 여러 가지 기법을 동시에

표 1. 능동식 3차원 깊이 센서와 수동식 3차원 깊이 센서의 비교  
Table 1. Comparison between active 3D depth sensor and passive 3D depth sensor

항 목	능동식 3차원 깊이 센서	수동식 3차원 깊이 센서
대표적기법들	Laser Scanning, Structured Light Pattern, 깊이정보 카메라(depth camera)	Shape from Stereo(SFS) or Shape from Motion(SFM)
3차원 정보 측정 방법	직접 측정	영상이나 비디오로부터의 간접 측정
공간 해상도	제한적 해상도	고해상도
구입/제작비용	고가	비교적 저가
정확도	정밀함	텍스처가 적은 영역이나 차폐영역에서 오차 발생

같이 이용함으로써 단독으로 사용한 기법에 비해 개선된 결과를 얻어내려는 연구가 많이 시도되고 있다<sup>[4][5]</sup>. 이들 기법들은 사용되는 기법의 종류에 따라 서로 다른 수동적 깊이맵 획득 기법들을 복합적으로 사용하는 기법과, 능동적 기법과 수동적 기법을 복합적으로 사용하는 기법으로 나눌 수 있다. 또한 서로 다른 기법에 의해 얻어진 깊이맵 데이터들을 어떻게 융합하느냐에 따라 부피 격자(volume grid)를 이용하여 융합하는 기법<sup>[6][7]</sup>, 영상 도메인(image domain) 상에서 융합하는 기법<sup>[8]</sup>, 그리고 변형 모델 틀 구조(deformation model framework)를 이용한 기법<sup>[9]</sup> 등이 있다. 본 논문에서 제안하는 기법은 능동적 기법인 깊이정보 카메라와, 수동적 기법인 스테레오 정합기법(Stereo Matching)을 이용하는 복합적 기법에 속하며, 이 두 기법에 의해 얻어진 영상 형태의 깊이맵 데이터들을 융합하는 기법이므로, 영상 도메인 상에서 융합하는 기법의 부류에 속한다. 이러한 깊이맵을 융합하는 연구들의 예로는 통계적 근사화(stochastic approximation)를 이용한 기법<sup>[10]</sup>, 스테레오 정합과 관찰 외피(visual hull)를 이용한 기법<sup>[11][5]</sup>, 3차원 스캐너로부터 얻어진 3차원 데이터와 스테레오 정합 기법을 이용하여 구해진 3차원 데이터를 융합하는 기법<sup>[11]</sup> 등이 있다. 이 기법들은 객체의 3차원 모델생성을 목적으로 한 기법도 있고, 전체 장면에 대한 모델생성을 목적으로 한 기법도 있는 등 3차원 정보를 구하기 위한 대상에 따라 다양하게 적용되고 있다. 따라서 구해지는 깊이맵 데이터의 특성에 맞는 융합기법이 필요하다. 본 논문에서는 능동 깊이 센서인 Z-cam<sup>TM</sup> 깊이정보 카메라로부터 얻어진 깊이맵 데이터와, 가운데 시점을 기준으로 좌측과 우측 영상과의 변이 추정기법에 의해 구해진 변이맵 데이터들로부터 이들의 특성을 분석하여 영상 도메인 상에서 융합함으로써 서로의 장단점을 보완하는 기법을 제안한다.

### III. 제안한 깊이맵 융합기법

#### 1. 깊이맵 취득(depth map acquisition)

제안한 깊이맵 융합기법의 입력 데이터인 깊이맵을 취득하기 위하여 그림 1과 같은 3-시점 카메라 및 가운데 시점 카메라에 설치된 Z-cam<sup>TM</sup> 깊이정보 카메라를 사용하였다. 이 카메라 시스템으로부터 가운데 시점 영상을 기준으로 좌우 시점 영상과의 스테레오 정합을 통하여 변이맵을 계산하

는 수동적 기법과, 가운데 시점의 깊이정보 카메라로부터 깊이맵을 얻는 능동적 기법으로 두가지 타입의 변이/깊이맵을 획득 하였다.

먼저 수동적 기법인 스테레오 정합기법은 그림 2와 같이 가운데 시점과 좌측 시점 영상, 가운데 시점과 우측 시점 영상 쌍 간에 각각 스테레오 정합기법을 적용하여 변이맵을 구하였다. 일반적으로 스테레오 정합을 이용한 변이맵 생성과정은 영역기반 정합의 경우 기준 영상의 모든 점에 대하여 대상 영상의 탐색 범위 내 화소들을 탐색하여 가장 유사도가 높은 점을 정합점으로 선택하고 이로부터 변이값을 구하는 과정이다. 따라서 고속의 변이맵 생성을 위해서는 탐색 범위의 축소가 필수적이다. 본 논문에서는 이러한 탐색 범위를 축소하기 위하여 에피폴라 조건(epipolar constraint)을 적용하였다. 에피폴라 조건을 적용하면, 탐색 범위를 2차원 탐색에서 1차원 탐색으로 줄일 수 있다. 이 조건의 적용을 위하여 변이 추정 전에 두 영상의 에피폴라 선(epipolar line)을 영상의 행 방향(scan line)과 일치시키는 영상 보정(image rectification) 과정<sup>[12]</sup>을 각 스테레오 영상 쌍에 대해 적용하였다. 이 과정은 교차축(toed-in) 스테레오 카메라로 촬영된 스테레오 영상을 평행축(parallel) 스테레오 카메라로 촬영된 영상으로 보정하는 과정과 동일하다. 보정된 두 스테레오 영상 쌍에 대하여 박 성찬 등이 제안한 스테레오 정합기법<sup>[13]</sup>을 적용하여 가운데 시점 영상에 대한 2개의 변이맵을 계산하였다.

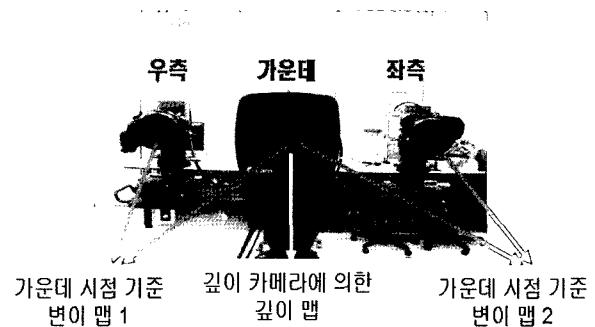


그림 2. 깊이맵 취득 방법  
Fig. 2. Depth map acquisition method

다음으로 능동적 기법으로서 가운데 시점의 깊이정보 카메라로부터 깊이맵을 얻는 기법을 사용하였는데, 깊이정보 카메라로서 3DV systems 사에 의해 개발된 Z-cam<sup>TM[3]</sup>을 사용하였다.

Z-cam™은 2 개의 IR 센서를 이용하여 카메라로부터 카메라 시야 내 각 점들까지의 거리를 실시간으로 계산하여 미리 정의된 거리 범위 내에서 각 거리를 256 단계 그레이 영상(gray image) 형태로 표현하여 영상 형태로 대응되는 칼라 영상과 함께 기록해주는 기능을 가지고 있다. 따라서 기록된 깊이맵의 최소/최대 거리를 사전에 알고 있을 경우, 깊이맵 내 각 화소들의 밝기를 실제 거리로 간단한 수식에 의해 변환할 수 있다. 표 2는 Z-cam™의 사양(specification)을 나타내고 있다.

표 2. Z-cam™ 깊이정보 카메라 사양<sup>[12]</sup>  
Table 2. Z-cam™ depth camera specifications <sup>[13]</sup>

감지 거리 범위	0.5 to 7 m(1.5- 23 feet)
거리 해상도	1cm(최대/최소 거리 설정에 따라 다름)
FOV(Field Of View)	40도
출력 영상 포맷	NTSC or PAL(4:3 ratio, Standard Definition)
출력 신호	SMPTE 259 M D1

## 2. 깊이맵 융합기법

일반적으로 스테레오 정합에 의한 변이 추출기법의 단점은 영상 내 다른 물체 또는 카메라 각도 차이에 의해 한쪽에서는 보이나, 다른 쪽 영상에서는 보이지 않은 영역이 존재할 경우 또는 정합하려는 화소 주변의 칼라가 거의 유사한 영역이 존재할 경우 오정합을 발생한다는 점이다. 따라서 이러한 영역에서의 변이맵 신뢰도는 매우 낮다고 볼 수 있다. 한편, Z-cam™으로부터 얻어지는 깊이맵은 스테레오 정합에 의해 얻어진 변이맵에 비해 전반적으로 장면 내 객체와 객체 간의 깊이 차이를 명확하게 나타내주며, 특히 객체의 경계부분에서도 변이맵에 비해 명확한 깊이 변화를 나타낸다. 그러나 영상 내에 잡음이 많이 포함되어 있으며, 2개의 IR 발신기(emitter)와 수신기(sensor)를 깊이 추출에 이용함으로써 인해 객체에 검은 색을 포함하거나 반사되는 부분에서는 정확한 깊이를 얻지 못하는 단점이 있다. 또한 깊이맵 취득 시 최대/최소 깊이 범위에 따라 깊이 정보의 정밀도가 달라지므로, 깊이맵을 취득할 대상에 따라 적절한 깊이 범위를 선택하여야 한다.

본 논문은 다시점 영상과 깊이맵으로부터 고화질의 임의 시점 영상을 생성하는데 목적이 있으므로, 객체보다는 장면 전체의 깊이맵을 얻는 것이 필요하며 이럴 경우 깊이정보 카메라의 최대/최소 깊이 범위가 넓어져서 개별 객체의 깊

이 정밀도는 저하되게 된다.

이러한 두 깊이 정보 추출기법의 장단점을 보완하고, 보다 정확한 깊이맵을 획득하기 위해 본 논문에서는 이 두 기법에 의해 얻어진 깊이맵들을 전처리하여 깊이값의 분포와 위치를 일치시킨 후, 가운데 시점의 영상 각 화소에 해당하는 깊이값을 구하기 위해 각 기법에 의해 얻어진 깊이값의 신뢰도를 통계적 기준에 의해 조사하여 가장 신뢰도가 높은 깊이값을 선택하는 깊이맵 융합기법을 제안한다.

제안한 기법의 특징은 서로 다른 기법에 의해 얻어진 여러 개의 깊이맵을 각 기법이 가지는 특성을 정량화하여 적절하게 깊이값을 선택함으로써 정확도를 향상시키는데 있다. 제안한 깊이맵 융합기법은 그림 3과 같이 구성되며, 상세한 내용은 다음과 같다.

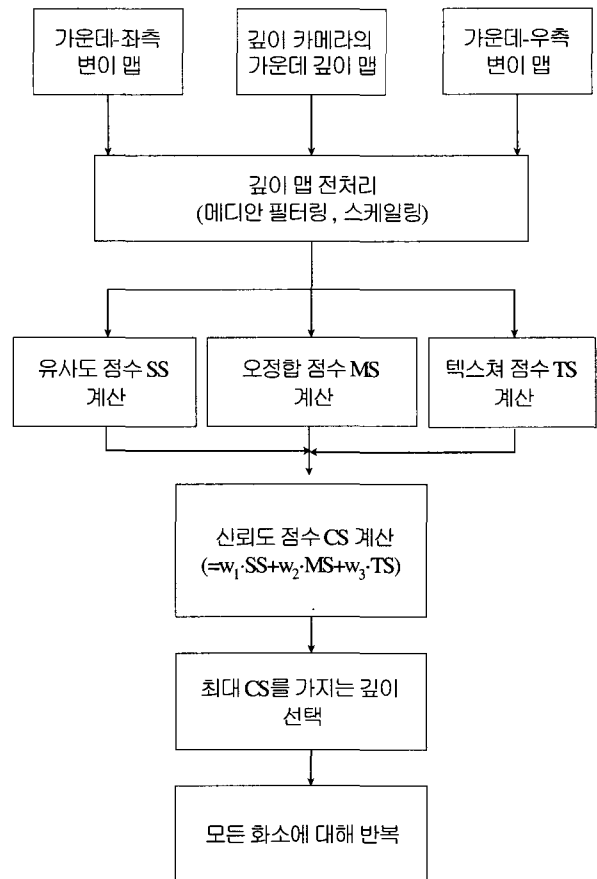


그림 3. 제안한 깊이맵 융합기법의 순서도  
Fig. 3. Flowchart of proposed depth map fusion method

먼저, 스테레오 정합 및 Z-cam<sup>TM</sup>에 의해 얻어진 변이맵들과 깊이맵의 전처리를 수행한다. 전처리 과정에는 오정합된 변이값의 제거, 잡음 제거 필터링, 두 기법에 의해 얻어진 데이터의 범위(scale)와 위치(position)를 일치시키기 위한 정합(registration) 및 스케일링(scaling)이 포함된다. 그림 1과 같은 카메라 시스템을 사용할 경우, 2 쌍의 스테레오 영상으로부터 2 개의 변이맵을 얻을 수 있으며, 깊이정보 카메라에 의해 1 개의 깊이맵을 얻을 수 있다. 이때, 변이맵은 가운데 시점의 깊이맵을 기준으로 생성하고, 영상보정에 사용되는 기반 행렬(Fundamental Matrix)을 이용하여 가운데 시점이 영상으로 역변환 한다. 깊이맵 간의 정합(registration)은 모두 가운데 시점의 깊이맵이므로, 따로 정합할 필요가 없으며, 두 데이터의 스케일(scale)을 맞추기 위하여 다음 식 (1)-(4)를 이용한다.

$$disp(i, j) = disp_{min} + (int) \left( \frac{gray_{disp} \cdot (disp_{max} - disp_{min})}{256} \right) \quad (1)$$

$$depth(i, j) = depth_{max} - (int) \left( \frac{gray_{depth} \cdot (depth_{max} - depth_{min})}{256} \right) \quad (2)$$

For the rectified stereo camera,  $depth(i, j)$

$$= \frac{B \cdot f}{disparity(i, j)} \quad (3)$$

$$disparity(i, j) = disp(i, j) \cdot dx \quad (4)$$

식 (1), (2)에서  $disp(i, j)$ 는 점  $(i, j)$ 에 대한 화소 단위 변이(pixel disparity)값,  $disp_{max}$ ,  $disp_{min}$ 은 각각 최대/최소 변이값,  $gray_{disp}$ 는 256 레벨(level)로 저장된 변이맵의 밝기,  $depth(i, j)$ 는 점  $(i, j)$ 에 대한 실제 깊이(depth) 값,  $depth_{max}$ ,  $depth_{min}$ 은 각각 최대/최소 깊이값,  $gray_{depth}$ 는 256 레벨(level)로 저장된 깊이맵의 밝기를 그리고  $dx$ 는 CCD의 가로방향 1 화소에 해당하는 크기를 나타낸다. 또한  $int$ 는 정수 부분만을 취하는 연산을 나타낸다. 식 (3)은 평행 스테레오 비전(Stereo Vision)의 깊이값과 변이값의 상호 변환 식을 나타내고 있는데, 여기서  $B$ 는 두 카메라 간의 간격,  $f$ 는 두 영상의 보정된 초점거리이다. 식 (1)과 (2)의 결

과를 식 (3)에 대입함으로써 깊이값을 변이값으로 또는 변이값을 깊이값으로 상호 변환 가능하다.

이러한 과정을 거쳐 변환된 변이값을 이용하여 가운데 시점 영상 기준의 각 화소에 대하여 깊이맵 융합을 수행한다. 먼저 각 변이맵을 조사하여, 해당 위치의 깊이값이 유효한 깊이값을 가지는지 여부를 조사한다. 만약 세 개의 깊이맵 중 하나의 깊이맵에서만 유효한 깊이 정보를 가지고 있고, 나머지 깊이맵에서는 유효한 깊이 정보를 가지지 않는 경우 유효한 깊이 정보를 가지는 깊이값만을 융합된 깊이맵의 깊이로 선택한다.

다음으로, 세 개의 깊이맵으로부터 얻어진 해당 화소에 대한 깊이값이 모두 유효한 값을 가지면, 이 세 깊이 값 중에서 규칙에 의해 가장 정확한 깊이값으로 판단된 깊이맵의 깊이값을 융합된 깊이맵의 해당 화소에 대한 깊이로 선택한다.

모든 영상 화소에 대해 두 번째와 세 번째 과정을 반복하여 최종적으로 융합된 깊이맵을 얻게 된다.

한편, 세 번째 과정의 각 깊이맵에 대한 신뢰도를 조사하기 위하여, 다음 식 (5)에 의해 신뢰도 점수 CS를 계산하게 된다.

$$Confidence\ Score\ CS = w_1 \cdot SS + w_2 \cdot MS + w_3 \cdot TS \quad (5)$$

여기서, SS는 유사도 점수(Similarity Score), MS는 오정합 점수(Mismatch Score), TS는 텍스처 정도 점수(Texture Score)이며,  $w_1$ ,  $w_2$ ,  $w_3$ 는 각 점수에 대한 가중치이다. 가중치는 실험적으로 결정되는 값으로서 보통 정수로 주어지나, 실수로 줄 수도 있다.

식 (5)에 나타난 각 점수들의 계산 기법은 다음과 같다.

먼저 유사도 점수 SS는 각 깊이맵의 깊이값을 위의 식 (3), (4)에 의해 변이값으로 변환하고, 이로부터 구해지는 정합점 간의 색(color) 차이를 식 (6)에 의해 각각 계산한 다음, 이 중 차이가 가장 작은 깊이맵의 깊이값에 대하여 1을 할당하고, 나머지 깊이맵들의 깊이값에 대하여는 0을 식 (7)과 같이 할당하게 된다.

$$Color\_diff_k(i, j) = |L(or R)(i, j + disp_k(i, j)) - C(i, j)| \quad (6)$$

여기서,  $i, j$ 는 영상 좌표이고,  $k$ 는 깊이맵 종류를 나타내는 인자(index)이고,  $L(i, j)$ ,  $R(i, j)$ , 그리고  $C(i, j)$ 는 각각

$$Similarity\ Score\ SS = \begin{cases} SS(k) = 1 & \text{if } \arg \min_k (Color\_Diff_k(i, j)) = k \ (k = 1, 2, 3) \\ SS(k) = 0 & \text{Otherwise} \end{cases} \quad (7)$$

좌우, 가운데 시점의 칼라영상이다. 또한  $L(i, j+Disp_k(i, j))$ 는 기준 영상(C, 가운데 시점 영상)에 대한 변이값에 의해 구해진 좌측 영상의 대응점을 의미한다.

다음으로, 오정합 점수 MS는 다음 식 (8)을 이용하여 계산하는데, 이는 장면 내 배경에 대한 깊이값이 깊이정보 카메라로부터 얻어진 깊이맵에서는 거의 동일한 값으로 표현됨으로 인해 오차가 많이 발생하므로, 보다 정밀한 깊이값을 얻기 위하여 스테레오 정합에 의한 변이맵으로부터 얻어진 깊이값을 우선 선택하기 위한 것이다.

$$\begin{cases} \text{If } depth_{z-cam} > BTh & MS = 1 \\ \text{Otherwise} & MS = 0 \end{cases} \quad (8)$$

여기서, BTh는 깊이정보 카메라로부터 얻어진 깊이맵에서 멀리 있는 배경으로 판단되어야 할 최소 깊이 임계값이며,  $depth_{z-cam}$ 은 깊이정보 카메라의 깊이 값이다.

한편, 텍스처 점수 TS는 다음 식 (9)을 이용하여 계산하였다.

$$\begin{cases} \text{If } ((T_w < Th1) \text{ and } (M_w < Th2)) & TS = 0 \\ \text{Otherwise} & TS = 1 \end{cases} \quad (9)$$

여기서, Th1과 Th2는 실험적으로 결정되는 임계값이며,  $T_w$  및  $M_w$ 는 식 (9)와 (10)에 의하여 계산된다.

$$T_w = \frac{1}{W_y} \times \sum_{r=1}^{W_x} \left( 1 - \frac{m_r}{\sqrt{\left(\frac{W_x-1}{W_x}\right)\sigma_r^2 + m_r}} \right) \quad (10)$$

$$M_w = \sum_{r=1}^{W_x} \frac{m_r}{W_y} \quad (11)$$

여기서,  $W_x$ 와  $W_y$ 는 각각 창들의 열과 행 크기이며,  $m_r$ 과  $\sigma_r$ 은  $W_x \times W_y$ 창들 내에서 각 행마다 계산되는 평균 밝기 및 밝기 값의 표준 편차로서 기준 영상에서 계산된다.

이렇게 계산된 SS, MS, TS와 가중치에 의해 깊이맵 데

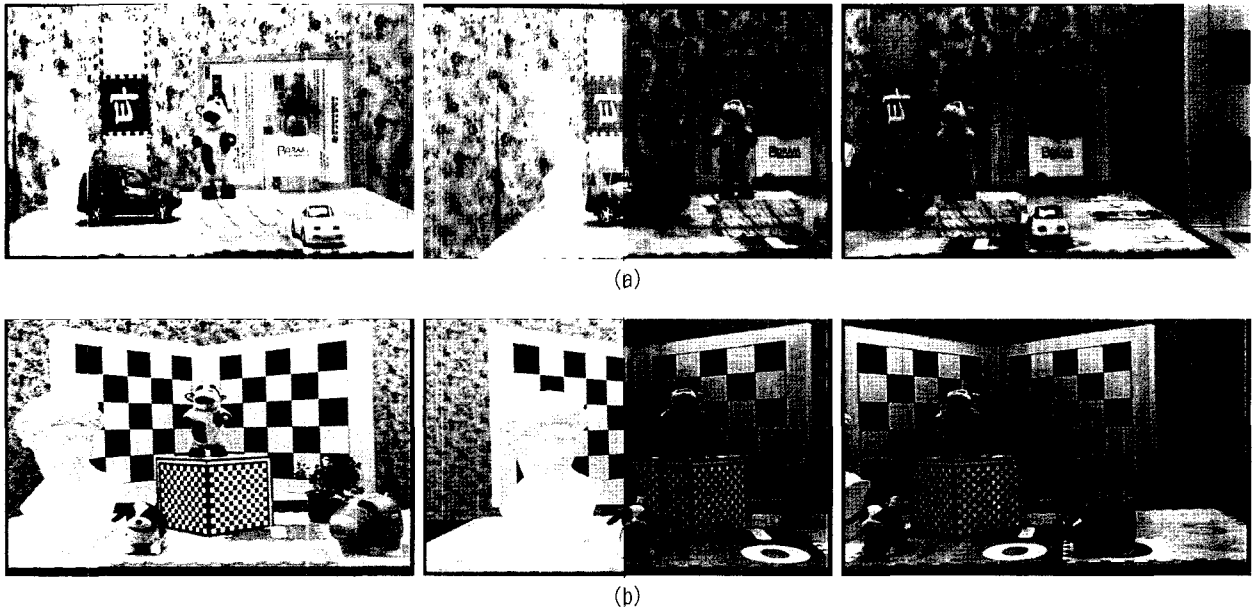


그림 4. 3-시점 실험 영상 (a)실험 영상 1 (b)실험영상 2  
Fig. 4. 3-view test images. (a)Test image 1 (b)Test image 2

이터의 CS를 계산하고, 각 화소마다 최대 CS를 가지는 깊이맵의 데이터를 가장 신뢰도가 높은 깊이값으로 판단하여, 이를 최종 융합된 깊이맵의 해당 화소에 대한 깊이값으로 결정하여 저장한다.

#### IV. 실험 결과 및 토의

##### 1. 실험 영상

본 논문에서 제안한 깊이맵 융합기법의 성능을 검증하기 위하여 그림 1의 카메라 시스템을 이용하여 3-시점의 영상을 촬영하였다.

그림 4는 그림 1의 카메라 시스템에 의해 촬영된 실험영상 1과 2를 나타내고 있으며, 표 3은 각 영상에 대한 정보를 나타내고 있다.

3장 1절에서 언급한 바와 같이 가운데 시점에 대한 깊이맵은 깊이정보 카메라인 3DV systems의 Z-cam<sup>TM</sup>으로부터 획득하였고, 가운데 시점과 좌측 시점, 가운데 시점과 우측 시점 스테레오 영상 쌍에 대해 영상 보정기법<sup>[12]</sup>을 각각 적용, 영상 보정을 수행하였다. 보정된 영상 쌍에 대하여 실

험영상 1에 대하여 박 성찬 등이 제안한 스테레오 정합기법<sup>[13]</sup>을, 실험영상 2에 대하여 Zitnick 등이 제안한 Cooperative기법<sup>[14]</sup>을 적용하여 변이맵을 각각 획득하였다. 최대최소 변이구간이 실험 1의 영상의 경우 ±(115 ~200 화소), 실험 2의 영상의 경우(좌영상-가운데영상 :+55~+65, 가운데영상 - 좌영상 :-90~+80화소)이다.

또한 각 변이맵 및 깊이맵은 제안한 깊이맵 융합기법의 적용을 위하여 3장에서 기술한 기법에 의해 정합 및 스케일링 과정을 수행하였다.

그림 5는 각 기법에 의해 구해진 실험 영상 1의 깊이맵 데이터들을 나타내고 있으며, 그림 6은 실험 영상 2의 깊이맵 데이터를 나타내고 있다.

먼저 그림 5의 각 깊이맵을 비교해 보면, 그림 5(a)의 깊이정보 카메라로부터 획득된 깊이맵의 오차가 스테레오 정합에 의해 구해진 두 변이맵들에서 나타나는 오차보다 전반적으로 적으며, 특히 객체와 배경의 깊이값의 차이가 잘 나타남을 볼 수 있다. 그러나 그림 5(a)의 경우에도 각 객체 내부 또는 배경에서의 미세한 깊이 변화는 잘 나타나지 않음을 볼 수 있으며, 비정규화 된 잡음이 대량 발생됨을 볼 수 있다.

그림 6의 경우는 그림 6(a)에서 알 수 있듯이, 영상에서 객체에 해당하는 공간에 깊이정보 카메라의 측정 깊이로

표 3. 각 실험 영상 1, 2에 대한 정보  
Table 3. Information about Test Images 1, 2

	실험영상 1	실험영상 2
영상 크기(pixel×pixel)	720×486	720×486
변이 범위(화소)	좌-중앙 영상: -200~-115 중앙-우 영상: 115~200	좌-중앙 영상: -55~-65 중앙-우 영상: 80~90
깊이정보 카메라의 깊이 범위(cm)	261 ~ 437	103~300
f(초점거리)·B(카메라사이거리)	좌-중앙 영상: 32,015cm 중앙-우 영상: 32,015cm	좌-중앙 영상: 7,355cm 중앙-우 영상: 10,040cm

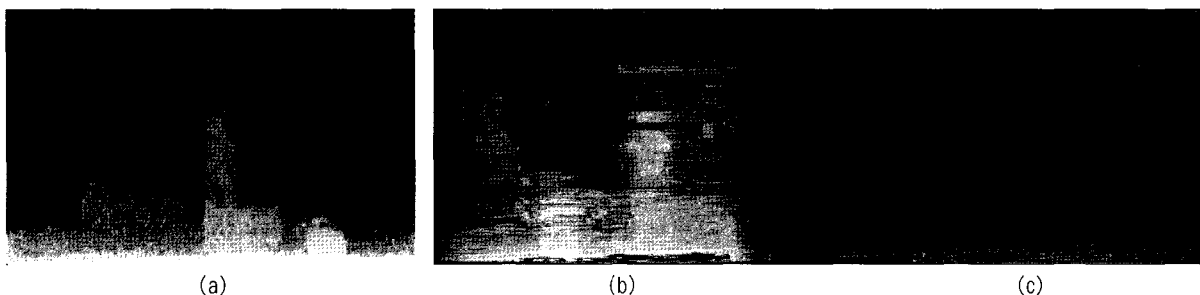


그림 5. 실험 영상 1에 대한 각 기법의 깊이맵. (a)깊이정보 카메라 (b)C-L 스테레오 정합 (c)C-R 스테레오 정합  
Fig. 5. Depth maps for the Test Image 1. (a)Depth camera (b)C-L stereo matching (c)C-R stereo matching

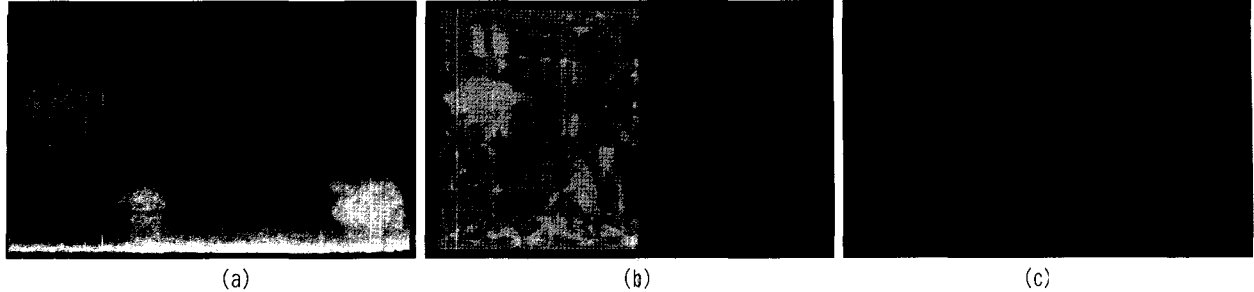


그림 6. 실험 영상 2에 대한 각 기법의 깊이맵. (a)깊이정보 카메라 (b)C-L 스테레오 정합 (c)C-R 스테레오 정합  
 Fig. 6. Depth maps for the Test Image 2. (a)Depth camera (b) C-L stereo matching (c)C-R stereo matching

설정하여 객체 부분에서 상당히 정확한 깊이값을 가지나 배경에 대한 정보가 전혀 나타나지 않는 것을 알 수 있다. 배경에 대한 정보를 보상해 주기위해, 표 3에 나타난 바와 같이 6(b), (c)의 변이맵은 변이 범위를 10화소 이내로 주어 가능한 한 배경에 해당하는 부분에서 정합이 정확하게 변이가 구해지도록 하였다.

그림 7은 제안된 깊이맵 융합기법을 적용한 결과를 나타내고 있는데, 그림 7(a)은 실험영상 1의 깊이정보 카메라로부터 얻어진 깊이맵을 사용하지 않고 좌우 변이맵만을 이용하여 제안한 깊이맵 융합기법을 적용한 결과이며, 그림 7(b), (c)는 각각 실험영상 1, 2의 모든 깊이맵을 사용하여 제안한 기법을 적용한 결과이다. 기존의 깊이맵 융합기법들은 본 논문에서 사용한 입력 깊이맵들과 그 특성이 다르기 때문에 객관적인 성능 비교가 어려워 비교 실험에서 제외하였다.

그림 7(a)의 결과를 보면, 영상 전체적으로 평가하였을 때, 개별 변이맵의 결과보다는 제안한 깊이 융합에 의해 정확도가 약간 향상되었음을 볼 수 있다. 다만, 깊이정보 카메라로부터의 깊이맵을 사용하지 않았으므로, 기본적인 스테레오 정합과정에서의 오차는 해결되지 못하였다. 특히 두 변이맵 간에 깊이값의 차이가 큰 영역이나, 한쪽 변이맵에서만 데이터가 존재하는 영역과 두 변이맵에서 모두 데이터가 존재하는 영역 경계에서 깊이의 불연속이 나타나고 있다. 이러한 융합을 거친 후 후처리를 통한 잡음 제거가 필요할 것으로 판단된다.

또한, 그림 7(b)의 결과를 보면 제안한 깊이맵 융합기법에 의해 배경부분의 깊이값에서 스테레오 정합에 의한 변이맵의 결과가 대부분 선택됨으로써 정확도가 향상되었음을 볼 수 있다. 그러나 전경에 해당하는 여러 객체 내부에서의 미세한 변화는 대부분 깊이정보 카메라로부터 얻어진

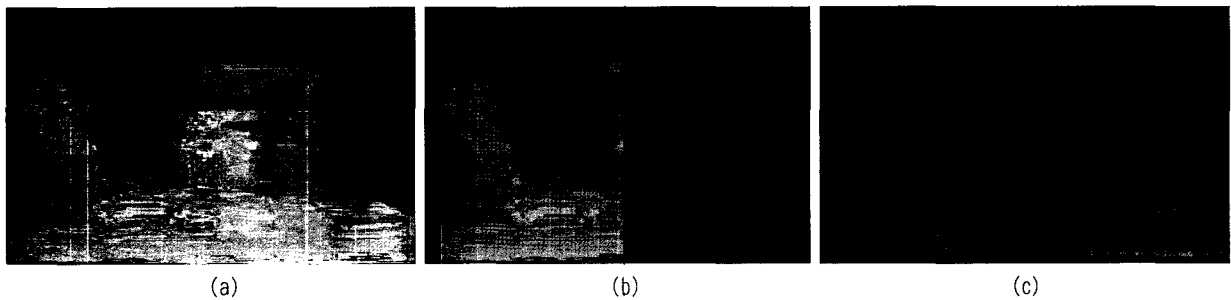


그림 7. 제안한 깊이맵 융합기법에 의해 융합된 깊이/변이맵 (a)실험 영상 1의 깊이정보 카메라 깊이맵을 사용하지 않은, 변이맵만의 융합 변이맵 (b) 실험 1영상의 모든 깊이맵 데이터를 사용한 융합 변이맵( $w_1=w_3=1, w_2=2, BTh=60, Th1=10, Th2=40$ ) (c) 실험 2영상의 모든 깊이맵 데이터를 사용한 융합 변이맵( $w_1=w_3=1, w_2=2, BTh=60, Th1=10, Th2=40$ )

Figure 7. Fused disparity/depth maps by proposed depth map fusion method. (a)Fused disparity map of test image 1 without using depth map from depth camera (b)Fused depth map of test image 1 using all three depth maps( $w_1=w_3=1, w_2=2, BTh=60, Th1=10, Th2=40$ ) (c)Fused depth map of test image 2 using all three depth maps( $w_1=w_3=1, w_2=2, BTh=60, Th1=10, Th2=40$ )



깊이가 선택됨으로 인해 미세한 깊이 변화는 반영되지 못하였다. 이를 해결하기 위해서는 스테레오 정합에 의한 변이맵을 좀 더 향상시키기 위한 연구가 더 필요할 것으로 판단된다.

다음으로, 실험 영상 2와 같이, 영상 내 존재하는 정합점 전체 변이범위가 60~ 250 화소의 영상의 스테레오 정합을 통한 정밀한 변이맵 생성은 상당히 어려운 작업임에도 그림 7(c) 결과를 보면 깊이정보 카메라에서 나타나지 않았던 배경에 해당하는 정보가 융합된 것을 알 수 있다. 그러나 깊이맵과 변이맵 사이의 스케일링 오차에 의해 그림 7(c)의 줄리앙 왼쪽 어깨 부분의 불연속 영역이 발생하였다. 그림 8(a)와 그림 8(b)는 실험 영상 1에 대해 각각 깊이정보 카메라로부터 획득된 깊이맵과 제안된 깊이맵 융합 기법에 얻어진 깊이맵을 이용하여 3차원 텍스처 투영(texture mapping)을 수행한 결과이다. 이 두 결과를 비교하였을 때, 객체 내부나 유사한 깊이를 가지는 부분의 깊이 연속성이 향상된 것을 볼 수 있다.

## VI. 결론 및 추후 과제

본 논문에서는 능동 3차원 깊이 센서인 깊이정보 카메라

와, 수동 3차원 깊이 센서인 스테레오 정합으로부터 얻어진 깊이/변이맵의 장단점을 분석하여 개별적인 깊이맵의 정확도보다 향상된 깊이맵을 얻기 위한 깊이맵 융합 기법을 제안하였다. 제안한 기법은 각각의 기법에서 오차를 발생시키는 인자들을 각각 정량화하고, 이로부터 구해지는 깊이맵 데이터의 신뢰도로부터 정확도가 높을 가능성이 높은 깊이값을 선택하게 된다. 본 논문의 특징은 최근 3차원 모델링 기법 중의 하나인 융합기법을 이용하여 3차원 깊이정보의 정확도를 향상시키는데 있으며, 제안한 깊이맵 융합기법을 3-시점 카메라 및 깊이정보 카메라로부터 얻어진 깊이맵과 변이맵에 적용한 결과, 전반적으로 깊이정보 카메라에 의해 깊이 정보가 취득되지 않거나 부정확한 배경 영역에서의 정확도를 스테레오 정합기법과의 융합에 의해 개선할 수 있었으며, 깊이맵의 연속성이 향상됨을 볼 수 있었다.

그러나 실험 영상 1에서 사용한 실시간 스테레오 정합기법 정확도의 한계로 인하여 스테레오 정합에 의해 얻을 수 있는 미세한 깊이 변화 정보가 반영되지 못하였고, 변이맵과 깊이맵 사이의 스케일링 오차에 의해 깊이 불연속구간이 발생하였다. 따라서 스테레오 정합에 의한 변이맵의 정확도를 향상시키기 위한 연구와 적절한 깊이맵과 변이맵의 스케일링 기법에 관한 연구가 필요할 것으로 판단된다.



그림 8. 실험 영상 1에 대한 깊이정보 카메라 및 제안한 융합기법에 의한 깊이맵의 3차원 텍스처 투영 결과. (a)깊이정보 카메라 (b)제안한 융합기법  
 Fig. 8. 3D texture mapping results for depth maps for test image 1 by depth camera and proposed fusion method. (a)Depth camera (b)Proposed fusion method

## 참고문헌

- [1] C. H. Esteban and F. Schmitt, "Silhouette and Stereo Fusion for 3D Object Modeling," Proceedings of 3DIM'03, pp. 46-53, Oct. 2003.
- [2] P. Dias, V. Sequeira, F. Vaz, and J. G. M. Gonçalves, "Registration and Fusion of Intensity and Range Data for 3D Modeling of Real World Scenes," Proceedings of 3DIM'03, pp. 418-426, Oct. 2003.
- [3] <http://www.3dvsystems.com>
- [4] S. F. El-Hakim, C. Brenner, G. Roth, "A Multi-sensor Approach to Create Accurate Virtual Environments," ISPRS Journal for Photogrammetry and Remote Sensing, vol. 54, no.6, pp. 379-391, Dec. 1998.
- [5] S. Y. Park and M. Subbarao, "A Range Image Refinement Technique for Multi-view 3D Model Reconstruction," Proceedings of 3DIM'03, pp. 147-154, Oct. 2003.
- [6] G. Cross, and A. Zisserman, "Surface Reconstruction from Multiple Views Using Apparent Contours and Surface Texture," Proceedings of NATO Advanced Research Workshop on Confluence of Computer Vision and Computer IGraphics, pp. 25-47, 2000.
- [7] Y. Matusumoto, K. Fujimura, and T. Kitamura, "Shape-from-silhouette/stereo and its Application to 3-d Digitizer," Proceedings of Discrete Geometry for Computing Imagery, pp. 177-190, 1999.
- [8] M. Li, H. Schirmacher, M. Magnor, and H. P. Seidel, "Combining Stereo and Visual Hull Information for On-line Reconstruction and Rendering of Dynamic Scenes," IEEE Workshop on MMSP, pp. 9-12, 2002.
- [9] P. Fua, and Y. G. Leclerc, "Object-centered Surface Reconstruction: Combining Multi-image Stereo and Shading," Int. J. of Computer Vision, vol. 16, pp.35-56, 1995.
- [10] A. R. Chowdhury and R. Chellappa, "A Robust Algorithm for Fusing Noisy Depth Estimates Using Stochastic Approximation," Proceedings of IEEE, ICASSP-01, 2001.
- [11] P. Dias, V. Sequeira, F. Vaz, and J. G. M. Gonçalves, "Registration and Fusion of Intensity and Range Data for 3D modeling of Real World Scenes," Proceedings of 3DIM'03, pp. 418-426, Oct. 2003.
- [12] J. Gluckman and S. K. Nayar, "Rectifying Transformations that Minimize Resampling Effects," Proceedings of Computer Vision and Pattern Recognition 2001 (CVPR'01), pp.111-117, vol. 1, Dec. 2001.
- [13] Hong Jeong, Sung-Chan Park, "Trellis-based Systolic Multi-layer Stereo Matching," Proceedings of IEEE Workshop on Signal Processing Systems 2003, 257-262, Aug., 2003.
- [14] C. L. Zitnick and T. Kanade, "Cooperative Algorithm for Stereo Matching and Occlusion Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 7, July 2000.

## 저 자 소 개



## 엄 기 문

- 1991년 2월 : 서강대학교 전자공학과 졸업
- 1993년 2월 : 서강대학교 전자공학과 석사학위 취득
- 1998년 2월 : 서강대학교 전자공학과 박사학위 취득
- 1998년 6월 ~ 2000년 3월: ETRI 컴퓨터 소프트웨어 기술 연구소 Post-Doc. 연구원
- 2000년 4월 ~ 현재 : ETRI 디지털방송연구단 방송시스템연구그룹 3DTV시스템연구팀 선임연구원
- 주관심분야 : 3DTV, 휴먼 팩터, 다시점 영상처리, Computer Vision 등



## 안 충 현

- 1985년 2월 : 인하대학교 해양학과 졸업
- 1989년 2월 : 인하대학교 해양학과 석사학위 취득
- 1995년 3월 : 지바대학교 자연과학연구과 공학박사
- 1995년 12월 : 지바대학교 공과대학 정보공학과 조수
- 1996년 1월~현재 : ETRI 디지털방송연구단 방송시스템연구그룹 3DTV시스템연구팀 책임연구원

저 자 소 개



이 수 인

- 1985년 2월 : 경북대학교 전자공학과 졸업
- 1989년 2월 : 경북대학교 전자공학과 석사학위 취득
- 1996년 2월 : 경북대학교 전자공학과 박사학위 취득
- 1990년~현재 : ETRI 디지털방송연구단 방송시스템연구그룹 책임연구원
- 주관심분야 : 지상파 DTV, DMB 시스템



김 강 연

- 2000년 2월 : 인하대학교 기계항공자동화공학부 졸업
- 2002년 8월 : 광주과학기술원 기전공학과 석사학위 취득
- 2002년 9월 ~ 현재 : 광주과학기술원 기전공학과 박사과정
- 주관심분야 : 다시점영상처리, 실감재질표현기술 등



이 관 행

- 1976년 2월 : 서울대학교 섬유공학과 졸업
- 1982년 2월 : 서울대학교 산업공학과 석사학위 취득
- 1985년 8월 : North Carolina State Univ. 산업공학과 석사학위 취득
- 1988년 8월 : North Carolina State Univ. 산업공학과 박사학위 취득
- 1988년 8월 ~ 1994년 : Northern Illinois Univ., Dekalb, IL, 조교수
- 1995년 1월~현재 : 광주과학기술원 기전공학과 교수
- 주관심분야 : 실감모델링 기술, 깊이기반 다시점 영상 생성기술, 광원모델링기술 등