

능동형 센서의 깊이 정보를 이용한 3D 객체 생성 (3D object generation based on the depth information of an active sensor)

김상진(Sang-Jin Kim) 유지상(Ji-Sang Yoo) 이승현(Seung-Hyun Lee)¹⁾

요약

본 논문에서는 능동형 센서를 이용하여 실사 객체에 대한 깊이 정보 및 칼라 정보를 획득하고 획득된 데이터를 이용하여 3D 객체를 생성하였다. 깊이 정보를 획득하는 방법은 능동형 센서 모듈을 내장한 ZcamTM 카메라를 이용하였다. <중략>세 번째, 세부 파라미터를 조절하여 깊이 정보의 왜곡을 보정하고 보정된 깊이 정보를 이용하여 3D 메쉬 모델을 생성한 후, 서로 인접한 외곽 점들을 연결하여 완전한 객체 메쉬 모델을 만든다. 최종적으로, 완성된 객체 메쉬 모델에 칼라 영상 데이터의 칼라 값을 적용해 매핑 처리를 수행함으로써 3D 객체를 생성하였다. 실험을 통해 능동형 센서가 장착된 카메라로 획득한 데이터만으로 3D 객체를 생성할 수 있다는 가능성을 제시하였으며, 3차원 전용 스캐너를 이용한 것보다 데이터 획득이 간편하고 용이함을 알 수 있었다.

ABSTRACT

In this paper, 3D objects is created from the real scene that is used by an active sensor, which gets depth and RGB information. To get the depth information, this paper uses the Zcam^T_M camera which has built-in an active sensor module. <중략> Thirdly, calibrate the detailed parameters and create 3D mesh model from the depth information, then connect the neighborhood points for the perfect 3D mesh model. Finally, the value of color image data is applied to the mesh model, then carries out mapping processing to create 3D object. Experimentally, it has shown that creating 3D objects using the data from the camera with active sensors is possible. Also, this method is easier and more useful than the using 3D range scanner.

Keywords: active sensor, 3D modeling, 3D mesh model, depth-Information, Zcam,

논문접수 : 2006. 8. 10.

심사완료 : 2006. 9. 10.

1) 정회원 : 광운대학교 전자공학과

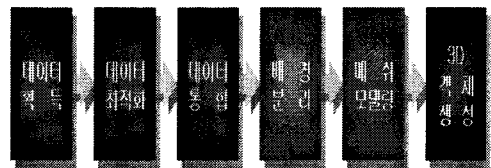
"본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과 (IITA-2006-C1090-0603-0017) 및 2005년도 광운대학교 교내연구비에 의해 연구되었음"

I. 서 론

현재의 3D 객체 생성 방법은 실세계의 객체(Real object)가 아닌 가상의 객체(Computer graphic)가 대부분을 차지하고 있다. 실세계의 객체를 3D 객체로 생성하기 위해서는 고가의 입체 스캐너를 이용해 객체에 대한 방대한 양의 깊이 정보를 획득한 후 메쉬 모델을 생성 [1]하는 방법과 2차원 영상들로부터 3D 정보인 깊이 정보를 복원하여 메쉬 모델을 생성 [2]하는 방법 등이 있다. 이러한 방법들은 모두 3D 메쉬 모델을 생성해야 하는데 Levoy는 다수의 스캐닝 데이터에서 iso-surface를 생성하는 알고리즘을 제안하였다 [3]. Hoppe등은 임의의 점군으로부터 signed-distance를 이용해 메쉬 모델을 생성하였다 [4]. Fua는 노이즈가 있는 점 데이터로부터 클러스터링을 통해 점 데이터를 분리하고 각각을 메쉬로 만드는 알고리즘을 제안하였다 [5]. 또한, 기존의 3D 객체 생성 방법은 컴퓨터 그래픽(CG) 틀을 이용하는 경우가 대부분인데, 이는 많은 시간과 노력을 필요로 하며 디자이너의 숙련된 기술을 요하는 경우가 대부분이다.

본 논문에서는 실세계의 객체에 대한 3D 객체 생성을 목표로 한다. 이를 위해, 3dvsystems사의 Zcam 카메라 [6][7]를 이용해 기초 영상 데이터를 촬영·획득하고, MS사의 VC++ 6.0 컴파일러와 3D 관련 그래픽 처리를 위한 OpenGL 라이브러리를 이용해 화질개선, 데이터의 통합 및 렌더링 관련 처리를 할 수 있는 프로그램을 개발, 이를 이용 3D 객체를 생성하였다. 여기서 사용될 데이터는 능동형 센서인 적외선 카메라 모듈을 내장한 Zcam 카메라를 이용하여 실세계의 객체를 촬영한다. 촬영된 영상 데이터를 이용하여 RGB 영상 데이터에 대해서는 선명화 필터를 이용하여 RGB 영상을 선명하게 개선하고, 깊이정보 영상 데이터에 대해서는 메디안 필터를 이용하여 깊이정보 경계선에 존재하는 잡음을 제거하여 두 영상 각각에 대해 화질 개선과정을 거친다.

화질개선 과정이 끝나면 4시점 각각에서 획득된 영상 데이터의 x좌표와 z좌표의 교환을 통해 하나의 공통된 좌표계를 공유하도록 데이터를 통합하고 배경과 객체를 분리한다. 이렇게 통합된 깊이정보 영상 데이터를 바탕으로 메쉬 모델링을 수행하는데 적용적 샘플링 등의 방법을 이용하여 삼각 폴리곤을 생성하고 RGB 영상 데이터를 이용하여 매핑을 수행하면 3D 객체 모델을 생성할 수 있다. 아래의 그림 1은 능동형 센서를 내장하고 있는 Zcam 카메라 시스템을 이용해 영상 데이터 획득해서 3D 객체 모델을 생성하는 과정 흐름을 그림으로 나타낸 것이다.



[그림 1] 3D 객체 생성 과정

II. 본 론

기존의 RGB 카메라나 적외선 카메라 등과 같은 카메라 장비를 이용할 경우, 우리가 필요로 하는 데이터인 객체 영상에 대한 RGB 데이터와 RGB 영상과 일대일로 대응되는 깊이 정보 데이터를 동시에 획득할 수 없다. 때문에, 이 두 데이터를 동시에 실시간으로 획득할 수 있는 Zcam 카메라 시스템 [6][7]을 이용하여 대상 객체에 대한 기초 영상 데이터를 촬영·획득하였다.

2.1. 데이터 획득

2.1.1 Zcam 카메라 시스템

Zcam 카메라 시스템은 이스라엘 3dvsystems사에서 방송용 ENG 카메라에 영역 감지용 능동형 센서를 부착하여 blue screen 효과를 주기위해 개발한 방송용 카메라인데,

RGB 칼라 영상과 깊이 정보인 깊이정보 영상을 실시간 동시 촬영이 가능한 입체 영상 데이터 획득 카메라로써, 그 구성은 크게 두 부분으로 나누어진다.

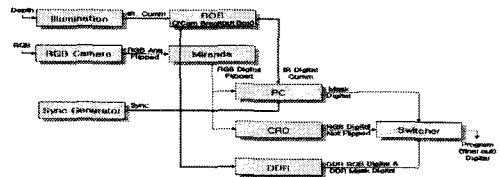
즉, 촬영 대상에 대한 RGB 칼라 영상 데이터 및 깊이정보 영상 데이터 촬영을 위한 카메라 모듈과 촬영된 영상의 A/D변환 및 영상 출력 등을 처리하기 위한 영상 처리 모듈로 구분할 수 있다. 여기서, 카메라 모듈은 다시 대상 객체에 대한 RGB 칼라 영상 데이터 촬영을 위한 RGB 카메라 모듈과 RGB 카메라 모듈에 의해 촬영된 영상에 대해 일대일 픽셀 대응되는 깊이 정보를 나타내는 깊이정보 영상 데이터 촬영을 위한 능동형 센서인 적외선 모듈로 나누어진다.



[그림 2] 3dvsystems 사의 Zcam 카메라

Zcam 카메라 시스템의 동작을 살펴보면, RGB 영상 데이터의 경우, RGB 카메라에서 영상 데이터를 획득해 영상이 아날로그 영상 데이터이면 미란다로 보내주고 디지털 영상 데이터이면 컴퓨터(Person Computer)로 보내준다. 미란다(Miranda)라는 A/D(Analog to Digital) 변환 장비로 보내주면 미란다는 RGB 카메라에서 받은 아날로그 영상 데이터를 디지털 영상으로 변환해 컴퓨터로 보내준다. 깊이정보 영상 데이터의 경우, 능동형 센서(Illumination)에서 깊이정보 영상 데이터를 획득해 BOB(Zcam BrackOut Box)로 보내주면 BOB는 능동형 센서에서 받은 데이터를 8비트 그레이 영상 데이터로 변환해 컴퓨터로 보내준다. 컴퓨터는 RGB 카메라 혹은 미란다에서 받은 디지털 RGB 칼라 영상 데이터와 BOB에서 받은 8비트 그레이 영상 데이터를 사용자가 원하는 형식의 파일로 저장해 준다. 즉, 정지영상의 경우, 32비트 Targa 형식으로 저장이 되는데,

24비트는 RGB 칼라 영상 데이터를 저장하고 8비트는 RGB 칼라 영상의 해당 픽셀에 대응하는 깊이 정보 데이터(0에서 255사이의 값을 갖는 256단계로 그레이 영상 데이터)를 하나의 파일 형태로 저장되어 사용자에게 제공되고, 동영상의 경우, 비압축 AVI 형식으로 RGB 칼라 동영상 데이터와 8비트 그레이 동영상 데이터가 각각 별도로 저장되어 사용자에게 제공된다.



[그림 3] Zcam 카메라 시스템 다이어그램

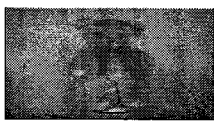
위의 다이어그램을 보면 기본적으로 Zcam 카메라 시스템이 방송용 시스템이기 때문에 CRD(Color Rotate Delay), DDR(Digital Disk Record), 그리고, VR 합성을 위한 스위처(Switcher) 장비까지 연결되어 있다. 하지만, 본 논문은 방송용 영상 데이터 생성을 목적으로 하는 것이 아니므로 컴퓨터에서 저장되는 데이터인 정지 영상 데이터를 이용하게 된다.

2.1.2. 데이터 획득 방법

기초 영상 데이터를 획득하는 방법에는 카메라 수와 촬영 장소, 대상 객체의 상태 등에 따라 여러 가지 방법이 있을 수 있다. 첫째는 대상 객체를 기준으로 정면, 후면, 좌측면, 그리고, 우측면에 90° 간격으로 카메라를 배치하고 4면의 기초 영상 데이터를 동시에 획득하는 방법이다. 둘째는 대상 객체의 정면을 기준으로 90° 간격으로 카메라를 이동시켜 가면서 4면의 기초 영상 데이터를 획득하는 방법이다. 마지막으로 대상 객체의 정면을 기준으로 90° 간격으로 카메라를 이동시키는 것이 아니라 대상 객체를 회전시켜 가면서 4면의 기초 영상 데이터를 획득하는 방법이 있을 수 있다. 첫 번째 방법은 객체가 움직일 경우에는 효과적인

수 있지만 4대의 카메라를 동시에 제어해야 하고 비교적 넓은 촬영 공간이 필요하다는 부담감이 있고, 두 번째와 세 번째 방법은 1대의 카메라를 사용하여 4면의 기초 영상 데이터를 획득할 수 있지만 객체가 움직일 경우에는 사용할 수 없고, 두 번째의 경우에는 비교적 넓은 촬영 공간이 필요하다.

본 논문에서는 움직임이 없는 객체를 대상으로 기초 영상 데이터를 획득하므로, 세 번째 방법인 1대의 카메라를 이용하여 대상 객체를 90° 간격으로 회전시켜 가면서 기초 영상 데이터를 획득하는 방법을 채택하여 데이터를 획득하였다. Zcam 카메라 시스템은 동영상과 정지영상에 대해 파일 형태로 사용자에게 제공할 수 있다. 동영상의 경우, 720 × 486 크기의 AVI 형식으로 획득되는데 RGB 칼라 영상과 깊이 정보인 depth-information 영상이 분리되어 따로따로 저장이 된다. 정지영상의 경우, 720 × 243 크기의 32비트 Targa 형식으로 획득되는데 RGB 칼라 영상과 깊이 정보인 depth-information 영상이 분리되어 저장되는 것이 아니라 하나로 통합되어 저장이 된다. 즉, 24비트의 RGB 칼라 영상 데이터와 8비트 depth-information 영상 데이터가 하나로 통합되어 32비트 RGBA Targa 형식의 파일로 사용자에게 제공된다. 아래의 그림 3.4는 Zcam 카메라 시스템을 이용해 촬영한 기초 영상 데이터의 정면 영상이다.



(a)RGB 칼라 영상 데이터



(b)깊이정보

[그림4]Zcam 카메라 시스템을 이용해 촬영한 기초 영상 데이터(정면 영상)

Zcam 카메라 시스템을 이용하여 기초 영상 데이터를 촬영할 때는 최종 결과 영상의 품질을 위해 Zcam 카메라 시스템의 내부 파라미터에 대한 최적화 작업이 필요하고, 대상 객체에

대한 기초 영상 데이터를 4방향에서 획득한 이유는 잡음 제거 및 데이터 통합시 대상 객체의 범위 및 기타 내부 파라미터 값의 설정과 최종 3D 객체 생성 시 대상 객체에 대한 occlusion (시점이 이동하면서 발생하는 데이터가 존재하지 않는 부분으로 특정 시점에서는 존재하는 데이터가 다른 시점에서는 존재하지 않는 데이터를 의미)을 줄이기 위한 것이다. 각 시점에서 얻어진 기초 영상 데이터는 Zcam 카메라 시스템의 내부 혹은 외부 파라미터는 기초 영상 획득하기 전에 이미 설정되어 있고, 기초 영상 데이터의 촬영·획득이 종료되기 전에는 임의로 변하지 않는다고 가정한다.

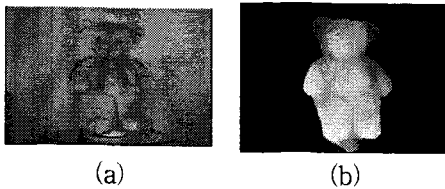
2.2. 데이터 최적화

Zcam 카메라 시스템을 이용하여 기초 영상 데이터의 획득이 완료되면 대용량의 영상 데이터에 대한 고속의 영상 처리와 최종적 결과물인 3D 객체 생성시 더 나은 고품질의 3D 객체와 보다 실제에 가까운 질감 표현 및 재현을 위해 4시점에서 획득한 기초 영상 데이터들 각각에 대해 필요한 경우 크기조절 및 화질개선 과정을 수행하여 기초 영상 데이터에 대해 3D 객체 생성에 적합하도록 최적화한다.

2.2.1 크기조절

Zcam 카메라 시스템은 동영상을 촬영할 경우에는 720 × 486 크기의 영상 데이터를 온전히 획득할 수 있지만, 정지영상을 촬영할 경우에는 720 × 486 크기의 절반 크기에 해당하는 720 × 243 크기로 영상 데이터가 획득된다. 그 이유는, 영상을 획득하기 전에 화면에 출력하는 방식이 인터레이스 방식으로 처리되고 있기 때문이다. 예를 들어, 초당 30프레임을 출력하는 환경에서 인터레이스 방식은 60프레임을 출력하게 되는데, 홀수 프레임에는 영상 데이터의 홀수 라인을 출력하고 짝수 프레임에서는 영상 데이터의 짝수 라인을 출력하게 된다. 이러한 특성 때문에 특정 한 프레임을 저장하게

되는 정지 영상은 원래 영상 데이터의 높이보다 적은 절반(486라인에서 243라인으로 감소)에 해당하는 만큼 줄어들어 상하로 늘리는 영상의 왜곡이 발생하게 된다. 이와같이 획득된 기초 영상 데이터의 왜곡을 바로 잡고 해당 영상 데이터의 연산량을 줄여 대용량의 이미지 연산에 대해 고속 이미지 프로세싱 처리를 위해 기초 영상 데이터의 크기를 조절하게 된다. 즉, 720×243 크기로 왜곡된 기초 영상 데이터를 기존의 4:3 비율인 320×240 크기의 영상으로 재조정하여 3D 객체 생성의 기초 영상 데이터로 사용하게 된다. 아래의 그림 5는 720×243 크기로 획득된 영상 데이터의 크기를 320×240 크기로 재조정된 결과 영상이다.



[그림 5] 크기 조정된 기초 영상 데이터
(a) 칼라 영상 데이터 (b) 깊이 영상 데이터

2.2.2 RGB 칼라 영상 데이터 처리

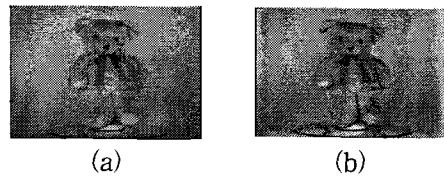
크기조절을 수행한 기초 영상 데이터에 대해 필요한 경우 RGB 칼라 영상 데이터에 대해 컨볼루션 필터를 이용해 객체와 배경의 경계를 또렷하고 영상을 더욱 뚜렷하게 한다. RGB 칼라 영상 데이터의 경우 depth-information 영상 데이터와는 달리 Red, Green, Blue의 세 개의 채널이 존재하게 되는데, 각 채널을 분리한 후, 목표로 하는 결과에 부합하는 채널에 선명화 필터를 적용한 후, 다시 재결합하는 것과 같은 처리가 필요하게 된다. 좀더 정확히 말하면, 먼저 RGB 모델을 이루고 있는 RGB 칼라 영상 데이터를 HSI 모델로 변환한다. HSI 모델은 3가지 채널로 구성되어 있는데, 색상 값에 해당하는 채널인 H(Hue)채널은 0° 에서 360° 의 범위를 가진 각도 값을 갖는다. 채도 값에 해당하는 채널인

S(Saturation)채널은 0과 1사이의 반지름 값을 갖는다. 그리고 명도 값에 해당하는 채널인 I(Intensity)채널은 0과 1사이의 z축으로 표현된다[8]. 이런 HSI 모델은 밝기 성분에 해당하는 I채널이 이미지에서 색 정보와 분리되어 있고, 색상과 채도 성분은 사람이 색을 인지하는 방식과 같은 관계를 가지고 있다[9]. RGB 모델을 HSI 모델로 변환한 후, 명도채널에 해당하는 I채널에 대해서만 선명화 처리 필터를 적용하여 선명화를 수행한다. 적용되는 선명화 처리 필터는 3×3 크기의 마스크로 아래의 그림 6과 같다.

0	-1	0
-1	5	-1
0	-1	0

[그림 6] 선명화 처리 필터

마지막으로 선명화 처리 필터 적용으로 선명화된 명도 채널과 색상 채널, 그리고, 채도 채널을 다시 RGB 모델로 변환하므로써 RGB 칼라 영상 데이터에 대한 선명화 처리를 완료한다. 아래의 그림 7은 선명화 필터 적용 전과 후의 RGB 칼라 영상 데이터의 모습이다. 그림 (b)의 결과 영상을 보면, 객체가 이전 영상(a)에 비해 상당히 선명해졌음을 알 수 있다.



[그림 7] 선명화 필터 적용 전과 후
(a)선명화필터 적용 전 (b)선명화필터 적용 후

2.2.3 깊이 정보 영상 데이터 처리

depth-information 영상 데이터는 촬영 대상 객체에 대한 깊이 정보를 저장하는 것으로 8비트(0에서 255까지 256단계로 저장) 크기의 데이터들의 집합이다. 이 데이터들은 RGB 칼라 영상 데이터와 픽셀 대 픽셀로 대응된다. 즉,

RGB 영상 데이터의 X좌표와 Y좌표에 해당하는 RGB 칼라 데이터의 깊이 정보 혹은 Z좌표 값은 depth-information 영상 데이터의 해당 X좌표와 Y좌표의 8비트 데이터를 읽음으로써 알 수 있다. 그러나 획득한 depth-information 영상 데이터를 보면 대상 객체와 배경 사이의 경계선에 비교적 잡음이 많이 포함되어 있음을 알 수 있다. 이러한 잡음은 3D 객체 생성시 불필요한 데이터일 뿐만 아니라 연산 시간을 높이는 계기가 되므로 제거해야 하는 것들이다. 아래의 그림 8은 대상 객체의 depth-information 영상 데이터에서 일부분을 확대한 그림으로써 객체의 경계선이 매끄럽지 못하고 울퉁불퉁한 모양을 하고 있음을 볼 수 있다.



[그림 8] depth-information 기초 영상 데이터의 부분 확대

위와 같이, depth-information 기초 영상 데이터에서 객체 데이터와 배경 사이의 경계선에 존재하는 잡음 데이터를 제거하기 위해서 3×3 크기의 마스크를 적용한 메디안 필터를 이용한다. 메디안 필터는 비선형 영역처리 기법으로 강한 edge(경계선)를 보존하고 기존의 경계선들을 좀더 상세하게 보존할 수 있다는 장점이 있으므로, 임의 크기의 윈도우를 depth-information 기초 영상 데이터의 좌측 상단에서부터 우측 하단으로 순차적으로 슬라이딩하면서 윈도우 안에 속하는 픽셀들에 대해 오름차순으로 순위를 결정하고 정렬해 중간에 위치하는 값을 해당 윈도우의 중심에 위치한 픽셀의 값과 대치시키는 방식으로 필터링을 수행한다. 이 메디안 필터를 적용할 때, 기초 영상 데이터의 서브 데이터에 해당하는 Red, Green, Blue, depth-information 영상 데이터 각각에 대해 모두 적용하는 것이 아니라 RGB

칼라 영상 데이터에는 적용하지 않고, 메디안 필터의 기능중 뚜렷한 밝기를 가진 픽셀을 그 이웃한 픽셀들과 유사하도록 만들어서 돌출되는 픽셀 값에 대한 제거가 가능하므로 depth-information 영상 데이터의 경계선에 존재하는 불규칙적인 데이터들을 제거하기 위해 depth-information 영상 데이터에 대해서만 메디안 필터를 적용한다. 아래의 그림 9는 메디안 필터를 적용하기 전과 후의 depth-information 영상 데이터이다.



(a)메디안 필터 적용 전 (b)메디안 필터 적용 후

[그림 9] 메디안 필터 적용 전과 후

RGB 칼라 영상 데이터에 대해 메디안 필터를 적용하지 않는 이유는, 각각의 요소(Red, Green, Blue)에 대하여 메디안 필터를 적용한 후, 각각의 필터링 결과를 하나의 픽셀로 결합하므로써 RGB 칼라 영상 데이터에 대한 메디안 필터링의 수행이 가능하나 필터링 수행 후 RGB 칼라 영상 데이터의 각 칼라 요소들 사이의 상관관계에 대한 손실이 발생하고 앞 절에서 논한 칼라 영상에 대한 선명화 필터만으로도 3D 객체 생성 결과에 부족함이 없으므로 RGB 칼라 영상 데이터에 대해서는 메디안 필터를 적용하지 않는다.

2.3. 데이터 통합

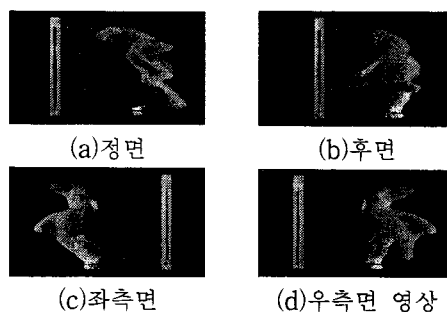
획득된 기초 영상 데이터에 대한 기본적인 영상 처리가 완료되면, 4시점에서 획득된 기초 영상 데이터를 하나의 공통된 좌표계를 공유하는 3차원 큐브상의 공간에 통합 및 재배치를 하게 된다. 이것은 3차원 공간 데이터를 구성하기 위한 X좌표 값, Y좌표 값, 그리고, Z 좌표 값을 완벽하게 구성할 수 있다는 것을 의미

하는 것으로 이렇게 구성된 3차원 좌표(X, Y, Z)값과 RGB 칼라 영상 데이터에서 해당 X좌표 값과 Y좌표 값의 칼라 데이터인 Red값, Green값, Blue값을 참조함으로써 3D 객체 모델링을 위한 기본적인 데이터를 확보할 수 있다.

2.3.1 좌표 재설정

각각의 시점에서 획득한 기초 영상 데이터들을 하나의 3차원 큐브상에 통합 배치할 때에는 각각의 기초 영상 데이터에 대해 좌표를 재설정 해주어야 한다. 왜냐하면, 기초 영상 데이터 획득시 대상 객체의 정면을 기준으로 반시계 방향으로 90° 간격으로 객체를 회전시켜 가면서 촬영했다. 즉, 각각의 시점에서 획득된 기초 영상 데이터의 정확한 X좌표 값과 Z좌표 값을 설정하기 기초 영상 데이터의 X좌표 값과 해당 X, Y좌표의 깊이 정보 값을 참조하여 획득시 반시계 방향으로 회전한 만큼을 다시 역으로 회전시켜야 한다. 이는 삼각함수를 이용하여 각각의 시점 영상 데이터에서 각각의 X좌표와 Z좌표를 계산하여 정확한 좌표를 설정할 수 있다. 하지만, 각각의 픽셀에 대해 삼각함수 연산을 수행하는 것은 상당한 시간이 소요되기 때문에 여기서는 좌표 간의 교환을 통해 같은 효과를 얻을 수 있다. 이때, Y좌표는 객체의 수평적 회전과는 관련성이 없으므로 Y좌표를 뺀 X좌표와 Z좌표 간에 대해서만 교환·재설정을 한다. 객체의 정면 기초 영상 데이터를 기준으로 할 때, 정면 기초 영상 데이터는 X, Y, Z좌표 값을 교환 없이 그대로 사용한다. 즉, 우측면 기초 영상 데이터의 경우 시계 방향으로 90° 회전한 영상 데이터이기 때문에, 새로이 설정되는 X좌표 값은 이전 기초 영상 데이터의 Z좌표 값으로 Z좌표 값은 이전 기초 영상 데이터의 X좌표 값으로 재설정하면 된다. 후면 기초 영상 데이터의 경우 시계 방향으로 180° 회전한 영상 데이터이기 때문에, 새로이 설정되는 X좌표 값은 이전 기초 영상 데이터에서 X좌표 값의 역순으로 Z좌표 값은

이전 기초 영상 데이터에서 Z좌표 값의 역순으로 재설정한다. 마지막으로 좌측면 기초 영상 데이터의 경우 시계 방향으로 270° 회전한 영상 데이터이기 때문에 X좌표 값의 경우 이전 기초 영상 데이터의 Z좌표 값의 역순으로 Z좌표 값의 경우 이전 기초 영상 데이터의 X좌표 값의 역순으로 재설정한다. 그림 3.11은 각각의 시점에서 획득한 기초 영상 데이터를 기초로 각각의 기초 영상 데이터를 3차원 큐브 상에 단순히 배치한 것이다. 배치는 단순히 RGB 칼라 영상 데이터에서는 X좌표, Y좌표, 그리고, 그 좌표에 해당하는 칼라 값인 Red값, Green값, Blue값을 참조하고, 깊이정보 영상 데이터에서는 RGB 칼라 영상 데이터의 X좌표, Y좌표와 동일한 좌표에 위치한 값인 객체의 깊이 정보를 Z좌표로 참조한다. 그래서 X좌표, Y좌표, Z좌표로 해당 포인트의 위치 값을 결정하고, Red값, Green값, Blue값으로 해당 포인트 위치의 칼라 값을 결정할 수 있다. 아래의 그림 10은 각각의 기초 영상 데이터를 3차원 큐브에 출력한 후 반시계 방향으로 90° 회전시켜 출력한 결과이다.



[그림 10] 각각의 기초 영상 데이터를 3차원 큐브에 출력한 후 90° 회전

그림 10에서 보는 것과 같이 획득된 기초 영상 데이터에서 깊이 정보를 나타내는 깊이정보 영상 데이터의 깊이 정보 자체에도 상당한 데이터 왜곡이 존재함을 볼 수 있다. 또한, 배경에 대한 데이터가 같이 존재한다. 이것을 배경 데이터를 제거하고 세부 파라미터는 조정하여 가상의 3D 큐브 공간에 통합·배치하면 아

래의 그림 11과 같은 결과가 나타난다.

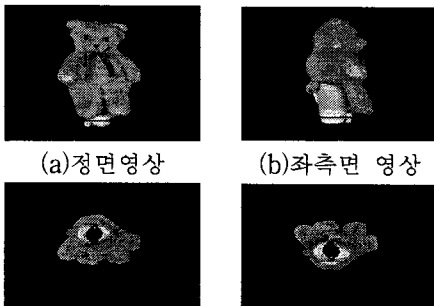


(a)정면에서 본 영상 (b)상단에서 본 영상

[그림 11] 세부 파라미터 조정 전의 기초 영상 데이터 통합

2.3.2 파라미터 조절

가상의 3차원 큐브에 각각의 시점에서 획득된 기초 영상 데이터들에 대해 좌표 재설정으로 통합한 후, 대상 객체 데이터와는 관련이 없는 배경 데이터를 제거한 후, 세부 파라미터를 조정하면 대상 객체와 흡사한 외형을 복원할 수 있다. 세부 파라미터의 조정으로 대상 객체와 흡사한 외형을 복원할 수 있다고 했는데, 원래 객체의 복원에 필요한 파라미터는 데이터 획득 시 대상 객체를 회전시켜 기초 데이터를 획득하므로 발생하는 좌우 위치 조절이 가능해야 하고, 대상 객체의 깊이 정보의 외곽을 보정할 수 있는 깊이값 조절이 가능해야 하며, 대상 객체의 깊이 정보의 출력에 있어 최대 값 혹은 최소 값의 조절이 가능해야 하고, 깊이값 조절은 모든 픽셀 데이터에 대해 동일한 상수 값의 적용이 아닌 동일한 비율 값이 적용된다. 아래의 그림 12는 세부 파라미터 조정 후 기초 영상 데이터들을 통합한 결과이다.



(a)정면영상

(b)좌측면 영상

(c)상단 영상

(d)하단 영상

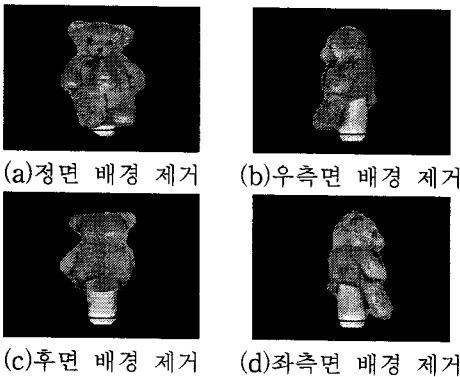
[그림 12] 세부 파라미터 조정 후의 기초 영상 데이터 통합

그림 12에서 보이는 바와 같이 객체에 대한 데이터가 존재하지 않는 영역(occlusion)과 불필요한 쓰레기 데이터(garbage data)가 발생하는 것을 볼 수 있다. 데이터가 존재하지 않는 영역의 경우, 기초 영상 데이터 획득시 정면을 기준으로 대상 객체가 수평으로 회전하면서 획득된 것이라 정면과 양측면, 후면보다는 상단면과 하단 면에서 많이 발생함을 알 수 있고, 쓰레기 데이터의 경우, 전체적으로 발생하는 것을 볼 수 있다.

2.4. 배경분리

배경분리 처리의 경우 데이터 통합 과정에서 동시에 수행된다. RGB 칼라 영상 데이터에서 배경 데이터와 객체 데이터를 분리할 때는 depth-information 영상 데이터를 이용하게 된다. RGB 칼라 영상 데이터에서 배경 데이터로부터 대상 객체 데이터를 분리할 때는 분리할 객체 데이터와 배경 데이터간의 칼라 밀도 차이를 이용하는 방법, 스테레오 영상에서 폭주각에 의한 좌우 영상의 미세한 차이를 이용하는 방법, 그리고 앞의 방법들을 결합한 방법 등이 있다. 이러한 방법들은 나름대로의 장·단점이 있다. 예를 들면, 칼라 밀도 차이를 이용하는 방법의 경우, 주변환경이나 카메라의 특성으로 인해 분리할 객체와 배경 영상의 컬러 밀도 차이에 오류가 발생한 경우에 배경의 분리가 어렵다. 또, 영상의 폭주각에 의한 좌우 영상 차이를 이용하는 경우, 기초 영상 데이터 획득 후 방대한 양의 후처리 작업과정이 필요하다[10]. 본 논문에서 사용된 기초 영상 데이터는 32비트 영상 데이터이다. 즉, 일반 카메라에 의해 획득된 24비트 크기의 RGB 칼라 영상 데이터와 능동형 센서에 의해 획득되고 칼라 영상 데이터와 일대일 픽셀 대응의 깊이 정보를 나타내는 8비트 크기의 깊이 정보 영상 데이터가 공존하고 있다. 깊이 정보 영상 데이터는 대상 객체의 깊이 정보 혹은 멀고 가까움을 256단계로 저장하는 것으로, 능동형 센서의

감지 범위를 대상 객체보다 넓게 설정하고 감지하면, 대상 객체에 대한 깊이 정보를 실시간으로 획득이 가능하다. 즉, $D=0$ (D 는 해당 픽셀에 대응하는 깊이 정보값)이면, 대상 객체에 대한 깊이 정보 데이터를 갖지 않는 것이고, $D \neq 0$ (D 는 해당 픽셀에 대응하는 깊이 정보 값)이면, 대상 객체에 대한 깊이 정보 데이터를 갖는 것이다. 그러면, 대상 객체 데이터를 제외한 나머지 배경에 대한 데이터에 대한 깊이 정보 값이 0이 된다. 그러므로 각각의 기초 영상 데이터를 통합할 때 조건 연산시 간단한 조건식(깊이 정보 값이 임계값보다 작은가?) 하나만으로 대상 객체 데이터와 배경 영상 데이터 사이의 분리가 가능하고, 경우에 따라선 임계값의 조절에 의한 깊이 정보의 한정화로 불필요한 쓰레기 데이터를 어느 정도 제거할 수 있다. 아래의 그림 13은 깊이정보 영상 데이터를 이용해 RGB 칼라 영상 데이터에서 배경 데이터를 제거한 결과이다.



[그림 13] 배경 데이터 제거

2.5. 메쉬 모델링

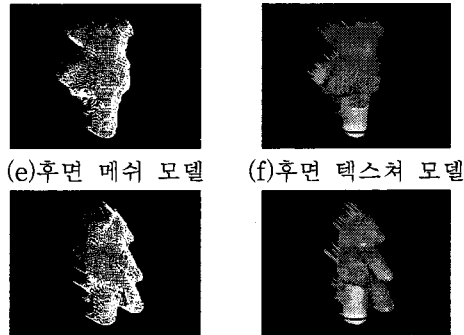
메쉬 모델은 통합된 4시점의 영역 데이터에 대해 한번에 통합된 메쉬 모델을 생성하지 않고, 4시점 각각에 대해 메쉬 모델을 생성하고, 외곽의 점들을 상호 연결함으로써 완전한 3D 객체 메쉬 모델을 생성한다. 텍스처 매핑 시에는 이미 저장되어 있는 칼라 데이터(R, G, B)

쌍이 사용되지 않고 점의 위치 데이터(X, Y, Z)쌍을 이용해 RGB 칼라 기초 영상의 좌표와 대응시켜 해당 위치의 텍스처 소스를 얻어와 메쉬 모델에 텍스처 매핑을 처리한다. 텍스처 매핑시 좌표값은 0에서 1사이의 실수 값을 갖는데 점 위치 데이터(X, Y, Z)쌍과 RGB 칼라 기초 영상의 좌표간의 대응은 RGB 칼라 기초 영상 데이터와 depth-information 기초 영상 데이터가 같은 크기의 이미지이고 각각의 이미지 내의 대상 객체의 위치가 동일하므로 단순히 점 위치 데이터(X, Y, Z)쌍의 요소중 X좌표와 Y좌표를 사용하고 각각에 대해 RGB 칼라 기초 영상 데이터 이미지의 X(320), Y(240) 사이즈를 나누어 줌으로써 대응을 완료할 수 있다.

2.5.1 삼각 메쉬 모델 생성

이전 단계의 결과로 얻어진 4시점의 깊이 정보 영상 데이터에 대해 삼각 메쉬 모델을 만들기 위해 삼각화를 수행한다. 이때, 4시점이 통합된 전체 깊이 정보 영상 데이터에 대해 삼각화를 수행하지 않고, 각 시점에서 얻어진 깊이 정보 영상 데이터에 대해 삼각화를 수행해 각 시점에 대한 삼각 메쉬 모델을 생성하고, 이후 각 시점의 삼각 메쉬 모델의 경계선에 대한 연결을 통해 하나로 통합한다. 삼각화를 위한 방법으로는 첫째, 사용자가 원하는 메쉬 모델을 얻을 수 있는 Delaunay 삼각화[11][12] 방법과 둘째, Delaunay 삼각화 방법보다는 상대적으로 단순한 Regular 삼각화 방법이 있다. Delaunay 삼각화는 객체의 전체 역영을 포함하는 삼각형에 샘플링을 통해 선택된 점 데이터를 이용해 삼각형을 늘려가는 방식으로 각 시점의 깊이 정보 영상 데이터에 분포하는 점들의 convex hull을 생성하기 때문에 원하는 형태의 원하는 형태의 메쉬 모델을 얻기 위해서는 불필요한 외부의 삼각형을 제거해야 한다. 이러한, 불필요한 외부의 삼각형을 제거하기 위해서 깊이 정보 영상 데이터에서 객체의 외곽선에 존재하는 점들을 찾고, Delaunay 삼

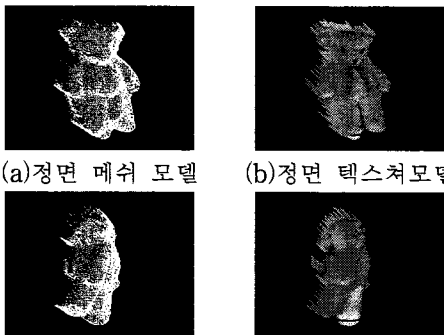
각화의 결과로 생성된 삼각형들 중 외곽선에 분포하는 점들로만 이루어진 삼각형을 제거해야 원하는 결과의 메쉬 모델을 얻을 수 있다. 이 방법은 사용자가 원하는 메쉬 모델을 생성할 수 있다는 것과 깊이 정보에서 메쉬 모델을 생성할 때 점 데이터의 수를 현격히 감소시킬 수 있다는 장점이 있다. 하지만, 이 방법은 기존 삼각화 된 영역 내에 새로운 점 데이터가 추가될 경우, 인접한 점의 위치를 파악하고 인접한 점과의 관계를 고려하여 기존 삼각화 영역내에 새로운 삼각화를 수행해야 하며, 삼각막이 완료된 이후 객체의 외곽선에 존재하는 불필요한 삼각형을 제거하기 위해 객체의 외곽선에 존재하는 점 데이터를 별도로 관리해야 하는 부담이 있다. 본 논문에서는 특정 한 시점에 대한 메쉬 모델을 생성하는 것이 아니기에 삼각화를 위해 객체의 외곽 점 데이터를 별도로 관리할 필요가 없고 위와 같은 부담이 존재하는 Delauny 삼각화 방식을 사용하지 않고, 상대적으로 간단한 Regular 삼각화 방식을 사용한다. 이 방식의 샘플링 방식은 시작 점 데이터부터 마지막 점 데이터까지 일정한 간격마다 점 데이터를 선택하고 인접한 두 점과 삼각형을 이루게 되는데 샘플링 순서는 Z자 모양 순서로 점 데이터를 선택해 나간다. 아래의 그림 14는 각각의 시점별로 Regular 삼각화를 통해 생성한 메쉬 모델과 텍스처 매핑 처리 결과이다.



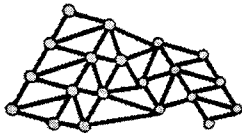
(e)후면 메쉬 모델 (f)후면 텍스처 모델
(g)좌측면 메쉬 모델 (h)좌측면 텍스처 모델
[그림 14] 부분 메쉬 모델과 텍스처 모델

2.5.2 삼각 메쉬 모델 통합

위의 그림 3.15에 부분 메쉬 모델 결과 영상을 보면 객체의 중심에서 배경과의 경계선 쪽으로 가까워지면 객체의 깊이 정보의 왜곡으로 인해 메쉬 삼각형이나 매핑 처리된 결과가 늘어지는 것을 확인할 수 있다. 이러한 왜곡에 대한 처리는 각각의 삼각 메쉬 모델의 외곽선에 존재하는 점들을 서로 삼각형으로 연결하여 하나로 처리하는 통합과정에서 처리한다. 각 시점별로 생성된 메쉬 모델은 서로 같은 공간 좌표를 공유하면서 존재하지만 서로 독립적으로 생성되어 있기 때문에 서로 인접한 두 메쉬 모델을 서로 연결해 주어야 한다. 즉, 정면 메쉬 모델의 오른쪽 외곽 점들은 우측면 메쉬 모델의 왼쪽 외곽 점들과 연결하고, 우측면 메쉬 모델의 오른쪽 외곽 점들은 후면 메쉬 모델의 왼쪽 외곽 점들과 연결하고, 후면 메쉬 모델의 오른쪽 외곽 점들은 좌측면 메쉬 모델의 왼쪽 외곽 점들과 연결하고, 마지막으로, 좌측면 메쉬 모델의 오른쪽 외곽 점들은 정면 메쉬 모델의 왼쪽 외곽 점들과 연결한다. 외곽 점들의 연결 방법은 아래의 그림 3.16과 같이 삼각화하여 연결해 각각 따로 존재하던 4개의 메쉬 모델을 하나로 통합하여 3D 객체를 생성한다.

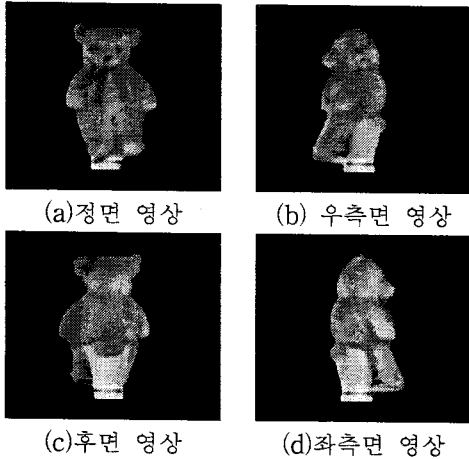


(a)정면 메쉬 모델 (b)정면 텍스처 모델
(c)우측면 메쉬 모델 (d)우측면 텍스처 모델



[그림 15] 메쉬 모델의 통합 방법

각 시점에서 획득된 기초 영상 데이터를 가공·처리하여 각각의 기초적인 메쉬 모델을 생성한 후 위의 그림 15와 같은 외곽 점 데이터 연결 방법을 이용하여 하나로 통합하면, 아래의 그림 16과 같이 각 시점에서 생성된 3D 객체 영상을 볼 수 있다.

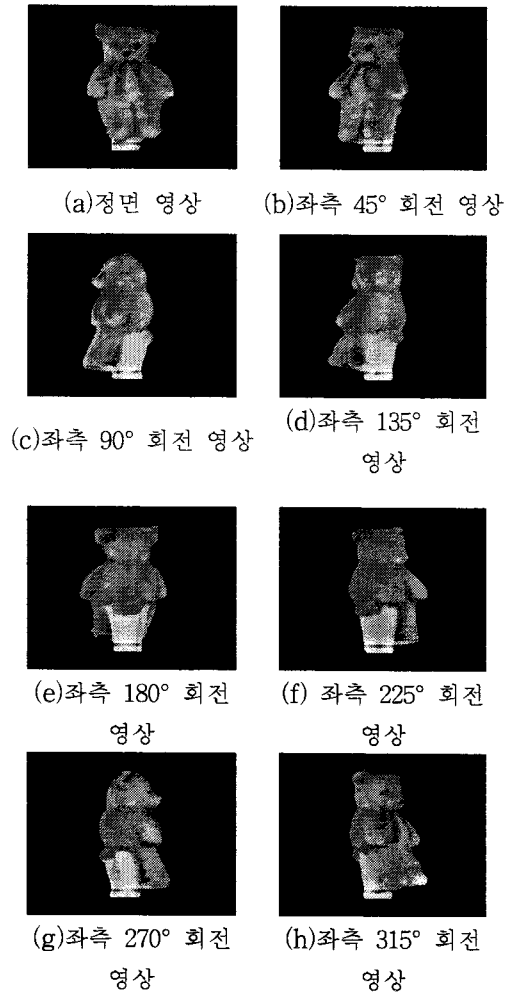


[그림 16] 메쉬 모델의 통합

III. 실험 결과

실험을 위한 데이터는 3dvsystems사의 Zcam 카메라 시스템을 사용하여 720 × 243 크기의 해상도를 갖는 정지 이미지를 320 × 240 크기의 해상도로 재조정하여 사용되었다. 제안된 3D 객체 생성 방법의 입력 데이터는 실제 객체의 외형적 데이터라고 할 수 있는 3D 점 데이터와 색상 데이터라고 할 수 있는 각 점의 컬러 값을 사용하였다. 전면과 후면, 좌측면과 우측면 이렇게 각각의 다른 시점 4곳에서 획득된 기초 영상 데이터 4장을 이용해 3D 객체를

생성해 시각자가 원하는 위치 어느 곳이나 볼 수 있도록 객체의 회전 및 확대·축소가 가능하다. 아래의 그림 17은 생성된 3D 객체를 왼쪽 45° 간격으로 회전시킨 영상들이다. 결과 영상을 보면 외곽의 점들을 연결한 절단면에서 색감의 차이가 발생함을 알 수 있다. 이것은 초기 영상 데이터 획득시 카메라 한대를 이용하여 객체를 회전시켜 가면서 초기 영상 데이터를 획득했고, 데이터를 획득하는 과정에서 카메라 특성 및 주변 환경의 변화로 인해 발생한 것이다.



[그림 17] 3D 렌더링 영상

IV. 결 론

본 논문에서는 능동형 센서의 깊이 정보를 이용한 3D 객체 생성에 관한 방법을 제안하였다. 제안된 방법은 3dvsystems사의 능동형 센서를 장착한 Zcam 카메라를 이용하여 실세계의 객체를 정면을 기준으로 90° 간격으로 객체를 회전시켜 4시점 각각에서 기초 영상 데이터를 획득한다. 즉, 획득된 기초 영상 데이터에는 대상 객체의 RGB 칼라 데이터와 RGB 칼라와 일대일의 대응관계이면서 해당 픽셀의 깊이 정보를 저장한 깊이 정보 영상 데이터를 동시에 실시간으로 획득이 가능하다. 이와같이 획득된 데이터는 각각의 영상에 대해 크기조절 및 화질개선의 단계를 거쳐 하나의 공통된 좌표계에 표시하기 위해 하나의 공통된 3차원 큐브에 좌표 변환 과정을 거쳐 통합하고, 카메라의 특정 및 주변의 외부환경 변환에 따른 세부 조절을 거쳐 각각 메쉬 모델을 만들고 각각의 메쉬 모델에 대해 외곽 점들을 삼각형 구조로 연결하여 통합하고 RGB 칼라 영상 데이터의 정보를 이용하여 텍스처 매핑 처리를 완료함으로써 하나의 3D 객체를 생성할 수 있다. 일반 카메라와 능동형 센서만으로 획득된 데이터를 이용해 3D 객체를 생성할 수 있다는 가능성을 제시하였고, 능동형 센서에 의해 깊이 정보를 획득하기 때문에 전용 스캐너를 이용한 것보다 데이터 획득이 간편하고 용이하며, 능동형 센서에 의해 획득된 깊이 정보를 기반으로 메쉬 모델을 생성하고 같이 획득된 RGB 칼라 영상을 기반으로 매핑 처리를 완료하므로 실세계 객체에 대한 3D 객체 생성이 용이해졌다.

참고문헌

[1] V.Sequeira, K.C. Ng, E.Wolfart, J.G.M Goncalves, and D.C. Hogg., "Automated 3D reconstruction of interiors with multiple scan-views", Proceedings of SPIE, Electronic Imaging, 1999.

[2] Johnson, S. Kang, "Registration and Integration of Textured 3-D Data" Tech. report CRL96/4, Digital Equipment Corporation, Cambridge Research Lab, 1996.

[3] Curless, B., and Levoy, M., "A volumetric method for building complex models from range images," Proceedings of SIGGRAPH'96, pp.303-312, 1996

[4] H. Hoppe, "Surface Reconstruction from Unorganized Points", PhD Thesis, Department of Computer Science and Engineering, University of Washington, 1994.

[5] P.Fua. From Multiple Stereo Views to Multiple 3D Surfaces. International Journal of Computer Vision, pp. 19-35, August 1997.

[6] <http://www.3dvsystems.com/technology/technology.html>

[7] <http://ww.3dkorea.co.kr/kr/page6.html>

[8] http://www.cs.kookmin.ac.kr/project/project_2003/project2003_04/project2003_04.htm

[9] 이창수, 지정규, "멀티미디어 서비스를 위한 동영상 이미지의 특징정보 분석 시스템에 관한 연구", 한국학술진흥재단, 제9권 제3호

[10] 홍동표, 우운택, "컬러의 채널 특성을 고려한 확률기반 배경분리 기법", 광주과학기술원, 논문자료

[11] H. Edelsbrunner. "Algorithms in Computational Geometry", Springer-Verlag, New York, 1987.

[12] S. M. Kim, J. C. Park, Kwan H. Lee, "Depth Video-based Human Actor Modeling for Multi-modal Interactions", 광주과학기술원, 논문자료