

손실된 영역의 복원을 이용한 깊이 영상 개선 기법

Depth Image Improvement using Estimation of Lost Region

조지호, Ji-Ho Cho, 박정욱, Joung Wook Park, 장인엽, In-Yoep Chang, 이관행, Kwan H. Lee
광주과학기술원 기전공학과, 지능설계 및 그래픽스 연구실

요약 본 논문에서는 깊이 영상을 개선하는 방법으로 깊이 영상 획득 시 손실된 영역을 복원하는 기법을 제안한다. 대상 객체의 동적인 3차원 정보는 적외선 깊이 센서가 장착된 깊이 비디오 카메라를 통하여 실시간으로 획득한다. 이때, 깊이 비디오뿐만 아니라 각 프레임마다 컬러영상이 동시에 획득된다. 그러나 대상 객체의 일부 또는 전체가 반짝이는 검은 재질로 되어있을 경우, 획득된 깊이 영상에 손실이 발생한다. 특히 방송용 콘텐츠로서 연기자의 3차원 정보를 획득할 때 머리카락 영역이 손실되는 심각한 문제가 발생한다. 이를 해결하기 위해 먼저 컬러 영상을 이용하여 손실된 영역의 위치 정보를 알아낸다. 손실된 영역 내 경계부분의 깊이 정보를 복원한 후 2차 베지어 커브로 보간하여 내부의 깊이 정보를 복원한다. 개선된 깊이 영상을 기반으로 일련의 모델링 과정을 수행하면 보다 자연스러운 3차원 모델을 생성할 수 있다. 생성된 3차원 모델은 실감방송용 콘텐츠로 사용될 수 있으며, 시청자에게 시각상호작용과 촉각상호작용 등 다차원 감각의 상호작용을 제공할 수 있다.

핵심어: 깊이 비디오 카메라, 3차원 객체 모델링, 깊이 영상 개선 기법

1. 서론

최근 3차원 방송 또는 실감 방송에 대한 관심이 높아짐에 따라 3차원 콘텐츠의 생성이 더욱 중요해지고 있다. 시청자에게 사실감을 부여하기 위해 실세계의 동적인 객체나 배우를 모델링 하는 연구가 활발히 진행되고 있다 [1].

동적인 객체로부터 3차원 정보를 획득하는 방법으로는 객체 주위에 여러 대의 카메라를 설치하여 완전한 3차원 모델을 생성하는 방법과 능동 깊이 센서 또는 스테레오 카메라를 이용하여 2.5차원의 깊이 정보를 획득하는 방법이 있다.

여러 대의 카메라를 이용하여 동적인 객체의 3차원 정보를 획득하는 방법으로 shape-from-silhouettes 기법이 많이 이용되고 있다 [2-3]. 이 방법을 이용하기 위해서는 카메라 보정을 통해 각 카메라의 내외부 파라미터를 획득해야 한다. 각 카메라로부터 획득된 영상의 실루엣 정보를 이용하여 convex hull 형태의 형상이 생성된다. 이 기법은 비교적 빠르고 정확하게 객체의 3차원 정보를 획득할 수 있지만, 제한된 스튜디오 환경에서만 사용이 가능하다. 또한 획득된 3차원 모델의 품질이 설치된 카메라의 대수와 카메라 보정결과에 크게 좌우된다.

능동 깊이 센서를 사용하는 깊이 카메라는 가격이 비싼 단점이 있지만, 수동적 방법인 스테레오 카메라를 이용하는 것보다 상대적으로 정밀한 깊이 값을 실시간으로 획득할 수 있다. 또한 촬영장소에 구애받지 않고 자유로운 환경에서 측정할 수 있는 장점을 가지고 있다 [4-5].

근래에 이러한 깊이 비디오를 기반으로 한 모델링 기술이 개발되었다 [6]. 이 기법을 이용하면 각 깊이 영상이 해당하는 3차원 형상으로 변환되어, 깊이 비디오가 3차원 메쉬 스트림의 형태로 변환된다. 변환된 3차원 모델은 비교적 좋은 품질이면서 동시에 촉각 상호작용을 할 수 있도록 충분히 부드러운 형상을 가진다.

그러나 깊이 비디오를 이용하여 획득된 깊이 영상에는 반짝이면서 검은색을 가지는 영역이 포함하지 않는다. 따라서 획득된 깊이 영상에 3차원 모델링 기법을 그대로 적용하면 검은 머리카락을 가진 사람을 제대로 표현할 수 없다. 보다 자연스러운 3차원 모델을 생성하기 위해서는 손실된 영역의 깊이 정보를 복원해야 한다.

본 논문에서는 깊이 카메라 시스템으로부터 획득한 두장의 영상 즉, 컬러 영상과 깊이 영상을 이용하여 손실된 영역의 위치 정보를 파악하고, 그 영역의 깊이 정보를 복원하는 기법을 제안한다. 개선된 깊이 영상은 깊이 영상 기반 모델링 과정을 통해 3차원 메쉬형태로 변환되고 이

는 본 연구실에서 수행중인 실감방송연구의 콘텐츠로 활용된다.

본 논문의 구성은 다음과 같다. 2장에서는 깊이 영상의 획득 방법과 획득한 깊이 영상의 문제점에 대하여 기술한다. 3장에서는 손실된 영역의 깊이 정보를 복원하는 방법에 대하여 자세히 기술한다. 4장에서는 깊이 영상 기반의 3차원 모델링 기법을 설명한다. 5장에서 실험결과와 응용 분야에 대하여 논하고 6장에서 결론을 맺는다.

2. 깊이 영상의 획득

깊이 비디오를 획득하기 위해 적외선 센서의 TOF(Time of Flight)방식을 이용하는 깊이 카메라인 ZCam™을 이용한다. ZCam™시스템을 이용하여 획득한 영상은 대상객체의 기하정보를 표현하는 깊이 영상과 객체의 외관을 나타내는 컬러 영상으로 구성된다. 그림 1은 ZCam™으로부터 획득한 컬러비디오 및 깊이 비디오 중 한 프레임만을 나타내고 있다.

그러나 깊이 비디오 카메라를 이용하여 획득한 깊이 영상에는 반짝이면서 검은색을 가지는 영역을 포함하지 않기 때문에 모델링에 그대로 적용할 수 없다. 그림 1(b)의 머리카락 영역이 그 예를 보여주고 있다. 이처럼 검은 머리카락을 가진 사람을 제대로 표현할 수 없다. 이를 해결하기 위해서 이전연구에서는 배우가 머리에 모자를 쓰고 촬영에 임해야 했다. 본 논문에서는 보다 자연스러운 3차원 모델을 생성하기 위해 손실된 영역의 깊이 정보를 복원하는 방법을 제시한다.



(a) 컬러 영상 (b) 깊이 영상

그림 1 ZCam™으로 획득한 데이터

3. 손실된 영역의 복원과정

그림 2는 깊이 영상에서 손실된 영역의 깊이 값을 복원하는 과정을 나타낸다. 손실된 영역의 깊이 정보를 복원하기 위해서 먼저 손실된 영역의 위치를 파악해야 한다. 대응되는 컬러 영상에는 깊이 영상에서 손실된 머리카락 영역의 위치 정보를 포함하고 있다. 따라서 컬러영상의 실루엣 정보를 추출하면 손실된 영역의 위치를 찾을 수 있다.

손실된 영역의 위치를 탐색한 후에 그 깊이 값을 복원하기 위해서 경계의 깊이 값이 복원되어야 한다. 앞에서

언어진 실루엣 이미지에 체인코드 알고리즘을 적용하여 객체의 경계정보를 획득한다. 획득된 경계정보에 기존 깊이 영상의 깊이 값을 할당하더라도 손실된 영역의 깊이 값을 알 수 없다. 선형 보간법을 이용하여 손실된 경계의 깊이 값을 복원하였다.

경계의 깊이 값 만을 복원했기 때문에 깊이 영상에는 커다란 홀(hole)이 존재한다. 이 홀을 효과적으로 채우기 위해서 2차 베지어 곡선이 이용되었다. 사람의 머리는 타원형의 부드러운 형상을 가지기 때문에 3개의 컨트롤 포인트를 가지는 2차 베지어 곡선을 이용하여 손실된 영역의 내부부를 복원하였다. 그 후에도 여러 개의 작은 홀이 생기는데 이는 closing operator를 이용하여 복원하였다.

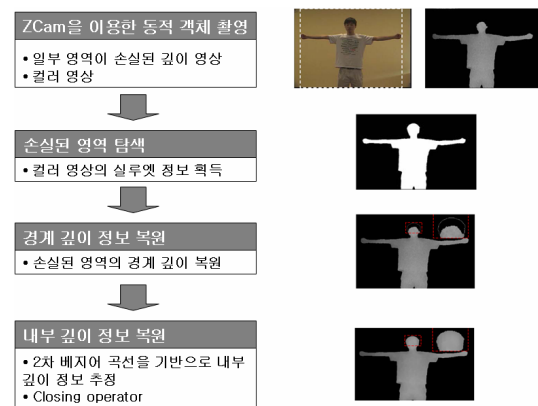


그림 2 손실된 영역의 복원과정

3.1 손실된 영역 탐색

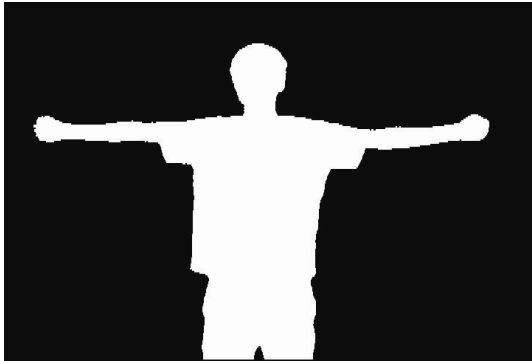
원본 컬러 영상에는 손실된 부분이 없고 깊이 영상에만 손실된 영역이 존재하므로, 손실된 영역을 탐색하기 위하여 컬러영상에서 대상 객체의 실루엣 정보를 추출한다.

실루엣 정보를 추출하기 위해 그림 1(a)에서와 같이 단일한 색을 가지는 배경에서 촬영하였다. 깊이 영상의 실루엣은 영상의 이진화를 통해 쉽게 구할 수 있다. 그 결과는 그림 3(a)가 보여준다. 컬러영상의 실루엣과 깊이 영상의 실루엣 정보는 머리카락영역을 제외하면 거의 일치한다. 따라서 깊이 영상의 실루엣 정보를 이용하여 오브젝트의 바운딩박스(Bounding Box)를 계산할 수 있다.

계산된 바운딩박스를 제외한 영역을 배경으로 간주하고 배경영역의 평균값 μ 와 표준편차 σ 를 R, G, B 각 채널 별로 계산한다. 이렇게 계산된 배경의 컬러정보를 바탕으로 대상객체를 분리한다. 배경과 객체를 분리하는 문턱값 (Threshold value)로는 $|\mu| \pm 3\sigma$ 를 사용하였다. 이를 이용하여 그림 3(b)와 같이 컬러영상의 실루엣 영상을 생성한다.



(a) 깊이 영상의 실루엣



(b) 컬러 영상의 실루엣

그림 3 실루엣 생성결과

3.2 경계 깊이 정보 복원

손실된 영역의 위치를 탐색한 후에 그 깊이 값을 복원하기 위해서 먼저 경계의 깊이 값이 복원되어야 한다. 앞절에서 얻어진 컬러영상의 실루엣 이미지에 체인코드 알고리즘을 적용하여 대상 객체의 경계 리스트 l_c 를 획득한다. 그림 4는 리스트 l_c 를 보여준다.

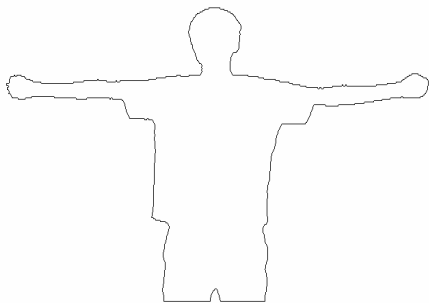


그림 4 컬러 영상의 경계 리스트(l_c)

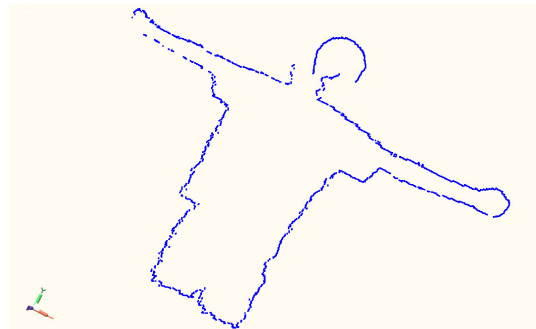
l_c 는 시계방향으로 정렬되어 있다. 획득된 경계정보에 기존 깊이 영상의 깊이 값을 할당하면 리스트 l_c 의 각 원

소(element)는 위치정보에 해당하는 x, y 값과 깊이 정보에 해당하는 z 값을 가진다. 그러나 원본 깊이 영상의 손실된 영역의 깊이 값이 존재하지 않기 때문에 그림 5(a)에서 보는 것과 같이 l_c 중 머리카락 영역에 해당하는 원소의 z 값은 0이다.

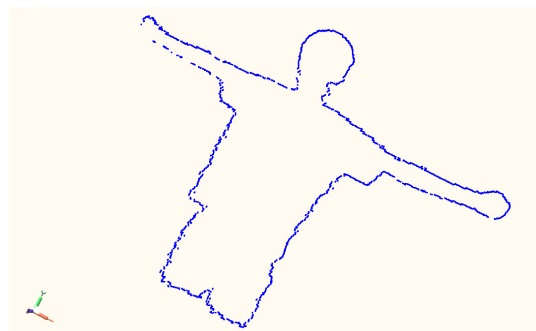
이러한 문제점을 해결하기 위해 z 값이 0인 원소들의 깊이 값을 복원한다. 리스트 l_c 중 처음으로 그 깊이 값이 0이 아닌 원소를 p_i 맨 마지막에 나타나는 0이 아닌 원소를 p_c 로 지정한다. 일반적으로 p_i 와 p_c 의 깊이 값이 동일하지 않기 때문에, p_i 와 p_c 의 x 좌표 상의 중간에 위치한 원소 p_c 가 사용된다. 그리고 나서 p_i 에서부터 p_c 사이 존재하는 원소들의 깊이 값은 다음과 같은 선형 보간법에 의해서 결정된다.

$$p_i.z = (1 - a)p_i.z + a p_c.z, i = 0, \dots, n-1 \quad (1)$$

n 은 포인트 p_i 와 p_c 사이에 존재하는 원소들의 개수를 나타내고 $a = i/n$ 이다. p_i 에서 p_c 사이의 결과는 식(1)과 동일하게 수행하면 얻을 수 있다. 그림 5(b)는 복원된 실루엣 리스트를 3차원으로 보여준다.



(a) 컬러 실루엣의 경계 리스트 (3차원)



(b) 복원된 경계 리스트 (3차원)

그림 5 경계 깊이 정보 복원 결과

3.3 내부 깊이 정보 복원

그림 3(b)인 깊이 영상의 실루엣영상에 체인코드 알고리즘을 적용하면 또 하나의 경계 리스트인 l_2 가 생성된다. 그림 6은 깊이 영상의 경계 리스트인 l_2 를 보여준다.

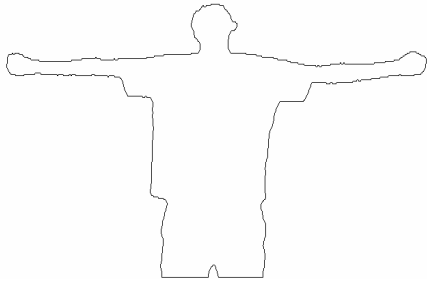


그림 6 깊이 영상의 경계 리스트(l_2)

앞 절에서 구했던 p_i 과 q_i 는 컬러영상의 경계리스트인 l_1 뿐만 아니라 깊이 영상의 경계리스트인 l_2 에도 존재하기 때문에 이 두 포인트 p_i 과 q_i 를 이용하면 머리카락영역을 의미하는 두 개의 리스트 즉, l_1 과 l_2 를 생성할 수 있다. l_1 은 l_2 에서 추출한 리스트이므로 머리카락영역 중 윗 부분에 해당하고 l_2 는 아랫부분을 의미한다.

일반적으로 l_1 에 포함된 원소의 개수가 l_2 에 포함된 원소의 개수보다 많고, 각 원소의 x 값만을 고려하면 다음과 같은 $l_1^x \supseteq l_2^x$ 의 성질은 만족한다.

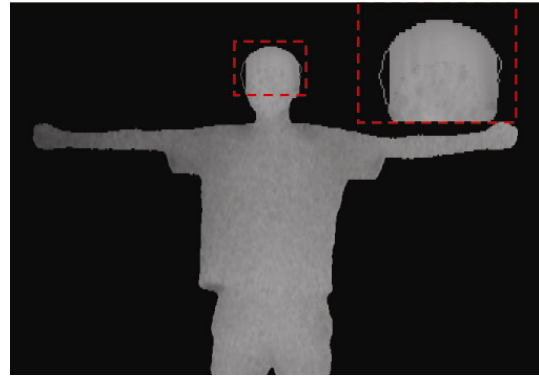
각 리스트의 원소 중 $l_1^x \cap l_2^x$ 를 만족하는 원소들만을 남기고 모두 제거하면 l_2 가 l_1 의 부분집합이므로 l_2 는 변화가 없고, l_1 은 교집합이 성립하지 않는 원소들이 제거되어 l_1 과 l_2 모두 같은 개수의 원소를 가지게 된다.

이렇게 생성된 두 리스트 l_1 과 l_2 사이의 깊이 정보를 보간 할 수 있다. 사람의 머리는 타원형의 부드러운 형상을 가지기 때문에 선형 보간 기법을 이용하면 부자연스러운 형상을 보인다. 그러므로, 보다 자연스러운 형상을 얻기 위해서 3개의 컨트롤 포인트를 가지는 2차 베지어 곡선을 이용하여 손실된 영역의 내부를 복원하였다[7]. 두 리스트의 모든 원소는 같은 x 값을 가지기 때문에, y 값과 z 값만을 가진 2차원 공간으로 간소화 하여 생각 할 수 있다. 각 y 값을 베지어 커브의 파라미터로 사용하면 다음과 같은 공식을 이용하여 손실된 영역의 내부를 복원할 수 있다.

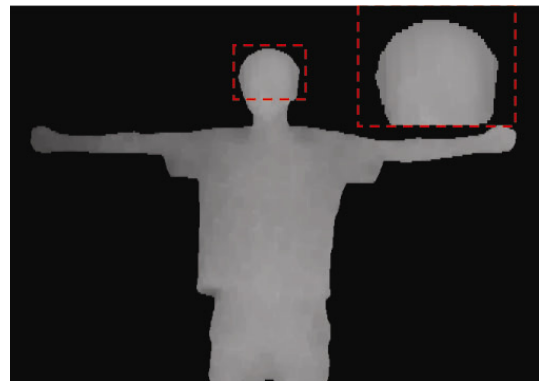
$$z_{ij} = (1-t)^2 p_i.z + 2t(1-t)m_i + t^2 q_i.z, j=0, \dots, n-1 \quad (2)$$

여기서 p_i 와 q_i 는 각각 l_1 과 l_2 의 원소를 의미한다. n 은 p_i 와 q_i 사이의 포인트 개수, 즉 $p_i.y - q_i.y$ 를 의미한다. m_i

$= \max(p_i.z, q_i.z)$, $t = j/n$ 이다. 2차 베지어 커브로 보간한 결과는 그림 7(a)와 같다. l_1 에서 교집합이 성립하지 않는 원소들을 모두 제거 했으므로, 제거되었던 원소에서 여전히 홀이 발생한다. 이 결과에 closing 연산자를 적용하면 그림 7(b)와 같은 홀이 모두 채워진 깊이 영상이 생성된다.



(a) 2차 베지어 커브 보간 결과



(b) closing 연산 결과

그림 7 내부 깊이 정보 복원 결과

4. 깊이 영상 기반 3차원 모델링에 응용

본 장에서는 깊이 영상 기반 3차원 모델링 (DIBM: Depth Image Based Modeling)기법을 소개한다[8-9].

DIBM의 입력으로는 깊이 비디오로 획득한 깊이 비디오가 사용된다. 본 논문에서는 그림 7(b)와 같이 머리카락이 복원된 깊이 비디오가 사용되었다. 획득된 원본 깊이 영상은 노이즈가 매우 심하기 때문에, 2D 영상의 노이즈 제거 기법 중 미디언 필터링이 사용된다.

샘플링 과정에서는 적응적 샘플링 기법을 사용한다. 깊이 영상의 모든 픽셀들을 삼각메쉬 생성에 이용하면 필요 이상의 중복된 삼각메쉬들이 생성된다. 따라서 최종 메쉬의 품질과 제한된 컴퓨팅 자원을 고려한다면 적은 수의 의미 있는 픽셀만을 선별하여 이용해야 한다. 깊이 정보의 변화량이 큰 의미 있는 픽셀을 특징점으로 이용한다. 특징점 추출은 입력영상 1차 그래디언트 (1st gradient)

해석을 통해 깊이의 변화가 큰 영역에서 특징점을 적응적으로 추출한다.

추출된 특징점들에 경계라인을 제약조건으로 하는 Delaunay 삼각화 기법을 적용하여 2차원 메쉬를 생성한다[10]. 2차원 메쉬 모델은 핀홀(pinhole) 카메라 모델을 이용하여 3차원으로 투영되면 3차원 메쉬가 생성된다. 그러나, 생성된 3차원 메쉬 모델도 여전히 형상 노이즈를 포함하고 있기 때문에 가우시안 스무딩 (Gaussian smoothing) 기법을 이용하여 형상의 노이즈를 제거한다.

최종적으로 부드러운 3차원 메쉬모델이 생성되면, 입력 컬러영상이 텍스처로 사용되어 실감나는 3차원 모델을 출력한다. 그림 8은 DIBM의 개괄적인 순서도를 보인다.

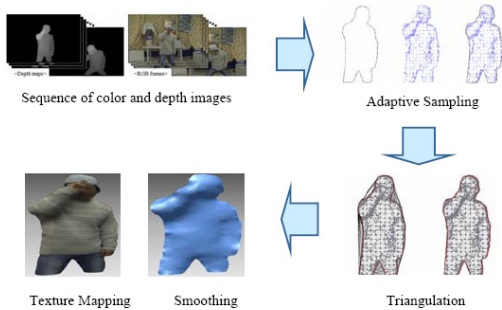


그림 8 DIBM 순서도

5. 실험 결과 및 응용

5.1 결과 비교

ZcamTM으로 측정된 모든 영상은 720 X 486의 해상도를 가지는 Standard Definition (SD)급의 영상이다. 깊이 영상 기반 모델링 기법[9]을 이용하여 이 영상 기반 모델링 3차원 모델을 생성하였다. 그림 9(a)는 제안된 방법으로 개선된 깊이 영상을 기반으로 모델링 한 결과이고 그림 9(b)는 원본 깊이 영상을 기반으로 모델링 한 결과이다. 제안된 복원기법을 거쳐 손실된 영역이 복원된 영상에 대한 결과가 훨씬 더 자연스럽다.



(a) 개선된 깊이 영상을 기반으로 3차원 모델링 한 결과



(b) 개선된 깊이 영상을 기반으로 3차원 모델링 한 결과

그림 9 깊이 영상 기반 모델링 결과

5.2 응용분야

생성된 3차원 모델과 가상의 3차원 배경을 합성하면 방송용 콘텐츠를 제작할 수 있다. 생성된 모델을 이용하여 여러 시점의 가상카메라를 설정한 후 렌더링 하면 다 시점의 장면을 생성하는 것이 가능하다. 이러한 기술을 이용하면 미래의 방송에서는 시청자가 다양한 시점에서 배우를 바라보는 것이 가능해진다. 그림 10은 다 시점 생성의 결과를 보여준다. 또한 햅틱장비를 이용하면 시청자가 화면 속의 배우나 CG 객체를 만져볼 수 있다[9,11].

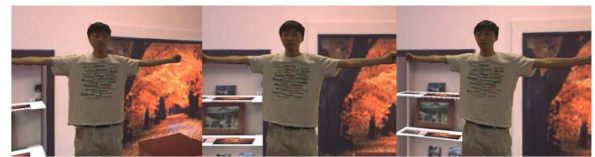


그림 10 다 시점 생성 결과

6. 결론

본 논문에서는 ZCamTM을 이용하여 동적 객체의 3차원 정보를 얻을 시 생길 수 있는 데이터 손실 문제를 다루었다. 손실된 정보는 해당 컬러 정보와 원본 깊이 정보의 깊이 정보를 기반으로 복원된다.

먼저 컬러 영상의 실루엣 정보를 이용하여 손실된 영역을 탐색한다. 그리고 나서 경계부분의 깊이 정보 복원을 수행한다. 마지막으로 2차 베지어 곡선과 closing 연산자를 통해 손실된 영역 내부를 복원한다. 이렇게 손실된 영역이 복원된 깊이 영상은 3차원 모델링을 통해 방송 시스템에서 3차원 콘텐츠로 이용될 수 있다.

본 연구에서는 검은 머리카락을 가진 사람을 측정할 때 생기는 문제점만을 다루었다. 또한 장발과 같은 특별한 헤어스타일을 가진 사람일 경우에는 제대로 적용되지 않는다.

향후 연구로 배우가 가진 허리띠나 가방 등의 소품으로 인한 깊이 영상 손실이나 다양한 헤어스타일에도 적용

할 수 있는 손실 복원 기법을 개발할 예정이다.

감사의 글

본 연구는 정보통신부 및 정보통신연구진흥원의 대학 IT연구센터(RBRC) 지원사업(IITA-2006-C1090-0603-0017)과 실감 콘텐츠 연구센터(ICRC)의 연구결과로 수행되었음.

참고문헌

- [1] A. Smolic et al., "Free Viewpoint Video Extraction, Representation, Coding, and Rendering," ICIP 2004, IEEE International Conference on Image Processing, Singapore
- [2] W.Niem, "Robust and Fast Modeling of 3D Natural Objects from Multiple Views," SPIE Proceedings, vol. 2182
- [3] J. Carranza, C. Theobalt, M. Magnor, H. P. Seidel, "Free-viewpoint video of human actors," ACM Transaction on Graphics, vol. 22, no. 3
- [4] G.J. Iddan and G. Yahav, "3D Imaging in the studio and elsewhere", Proceedings of SPIE, Vol. 4298
- [5] ZCam™ add-on, <http://www.3dvsystems.com>
- [6] 김승만, 박정철, 이관행, "적응적 샘플링을 이용한 깊이 맵 기반 3 차원 매쉬 모델링," 제 18 회 신호처리합동 학술대회, Vol. 18, No. 1, 2005.10. pp.157.
- [7] D. Marsh, "Applied geometry for computer graphics and Cad," 1999.
- [8] 김승만, 박정철, 조지호, 이관행, "형상 스무딩과 Z-buffer 렌더링을 이용한 깊이 영상의 노이즈 필터링," HCI 2006 학술대회(2006. 2. 13~16), Vol. 1, pp. 1188-1193
- [9] Seung-man Kim, Jeung Chul Park, and Kwan H. Lee, "Natural-Textured Mesh Stream Modeling from Depth Image-Based Representation," In Proc. International Conference on Computational Science and its Applications, UK (2006) 480-489
- [10] 박정철, 김승만, 이관행, "경계라인 제약조건을 이용한 깊이 맵 기반 매쉬 모델링," 한국정보과학회, 가을 학술대회, Vol. 32, No. 2, 2005.11. pp. 682-684.
- [11] Jongeun Cha, Seung-man Kim, Sung-yeol Kim, Sehwan Kim, Seung-uk Yoon, Ian Oakley, Jeha Ryu, Kwan H. Lee, Yo-sung Ho, Woontack Woo, "Client System for Realistic Broadcasting: A First

Prototype" Lecture Notes in Computer Science / PCM 2005, 6th Pacific Rim Conference on Multimedia, Jeju, Korea, 2005