

형상 스무딩과 Z-buffer 렌더링을 이용한 깊이 영상의 노이즈 필터링

김승만¹, 박정철², 조지호³, 이관행⁴
광주과학기술원, 지능설계 및 그래픽스 연구실^{1 2 3 4}
sman2000@gmail.com¹, {jucpark², jhcho³}@gist.ac.kr, lee@kyebek.gist.ac.kr⁴

Noise filtering for Depth Images using Shape Smoothing and Z-buffer Rendering

Seung-Man Kim¹, Jeung-Chul Park², Ji-ho Cho³, Kwan H. Lee⁴
IDEG lab., Gwangju Institute of Science and Technology^{1 2 3 4}

요약

본 논문에서는 동적 객체의 3 차원 정보를 표현하는 깊이 영상의 노이즈 필터링 방법을 제안한다. 실제 객체의 동적인 3 차원 정보는 적외선 깊이 센서가 장착된 깊이 비디오 카메라를 이용하여 실시간으로 획득되며, 일련의 깊이 영상, 즉 깊이 비디오(depth video)로 표현될 수 있다. 하지만 측정환경의 조명조건, 객체의 반사속성, 카메라의 시스템 오차 등으로 인해 깊이 영상에는 고주파 성분의 노이즈가 발생하게 된다. 이를 효과적으로 제거하기 위해 깊이 영상기반의 모델링 기법(depth image-based modeling)을 이용한 3 차원 메쉬 모델링을 수행한다. 생성된 3 차원 메쉬 모델은 깊이 영상의 노이즈로 인해 경계 영역과 형상 내부 영역에 심각한 형상 오차를 가진다. 경계 영역의 오차를 제거하기 위해 깊이 영상으로부터 경계 영역을 추출하고, 가까운 순서로 정렬한 후 angular deviation 을 이용하여 불필요하게 중복된 점들을 제거한다. 그리고 나서 2 차원 가우시안 스무딩 기법을 적용하여 부드러운 경계영역을 생성한다. 형상 내부에 대해서는 경계영역에 제약조건을 주고 3 차원 가우시안 스무딩 기법을 적용하여 전체적으로 부드러운 형상을 생성한다. 최종적으로 스무딩된 3 차원 메쉬모델을 렌더링할 때, 깊이 버퍼에 있는 정규화된 깊이 값들을 추출하여 원래 깊이 영상과 동일한 깊이 영역을 가지도록 저장함으로써 전역적으로 연속적이면서 부드러운 깊이 영상을 생성할 수 있다. 제안된 방법에 의해 노이즈가 제거된 깊이 영상을 이용하여 고품질의 영상기반 렌더링이나 깊이 비디오 기반의 햅틱 렌더링에 적용할 수 있다.

Keyword : Depth image, Noise filtering, Shape modeling,, Z-buffer rendering

1. 서론

최근 디지털 방송 기술의 발달과 함께 3 차원 방송에 대한 관심이 증가하고 있다. 3 차원 방송은 방송 콘텐츠에서 입체 영상을 생성하고 재현함으로써 시청자에게 3 차원 모니터를 이용하여 입체감을 제공하는 기술이다. 입체 영상을 생성하는 기술로는 영상 기반 모델링 및 렌더링 기술(IBM: Image-based modeling and rendering)이 있다 [1, 2, 3]. IBM은 실세계 영상을

이용하여 입체 영상을 생성하거나, 임의 시점에서의 영상을 합성해 내는 기술이다. IBMR에서는 장면의 3 차원 정보가 영상의 각 화소당 깊이 값으로 정의될 수 있다. 이는 깊이영상 기반의 3 차원 정보 표현방법 (DIBR : Depth image-based representation)이라고 불린다 [4]. 특히 각 프레임마다 색상 영상에 동기화되어 있는 깊이 영상을 가지는 비디오를 깊이 비디오(depth video)라고 한다.

동적인 객체로부터 깊이 비디오를 획득하기 위해서는 능동 깊이 센서나 수동적 스테레오 알고리즘이 사용될 수 있다. 스테레오 알고리즘은 두 장의 영상에 대해서 카메라 보정, 대응점 탐색 등의 과정을 거쳐 장면에 대한 변이 맵을 계산하는 것이다. 이에 대해서 많은 연구가 진행되었지만 폐색이나 대응점 불일치 등의 근본적인 문제로 인해 정밀한 3 차원 정보를 획득하는 데는 한계가 있다 [5, 6, 7, 8]. 이에 반해 능동적 깊이 센서를 사용하는 깊이 카메라는 가격이 비싼 단점이 있지만, 수동적 방법에 비해서 상대적으로 정밀한 깊이 값을 픽셀 단위로 실시간 획득하는데 사용되고 있다 [9]. 본 연구에서는 적외선 센서가 장착된 깊이 비디오 카메라를 사용하여 움직이는 객체에 대한 3 차원 정보를 획득한다.

깊이 비디오는 입체 방송 시스템에서 전송이나 압축에 효율적으로 사용되고 있는 효율적인 3 차원 표현방법이다. 하지만 실제 동적인 객체를 실시간으로 촬영하는 과정에서 깊이 비디오의 각 깊이 영상마다 고주파 성분의 노이즈가 포함된다. 만약 노이즈로 인해 왜곡된 3 차원 정보를 포함하고 있는 깊이 비디오가 그대로 전송되고 입체 영상을 생성하는데 사용되면 렌더링되는 3 차원 장면이 심각하게 왜곡된다. 이를 해결하기 위해 미디언 필터링이나 저주파 통과 필터를 사용할 수도 있지만 국부적인 필터링 연산자가 적용되기 때문에 전역적으로 부드러운 입체 영상을 획득하기 어렵다 [10]. 본 논문에서는 깊이 영상을 기반으로 명시적인 3 차원 형상 모델링 기술과 Z-buffer 렌더링 기술을 이용하여 깊이 영상내의 관심 객체에 대해서 전역적으로 부드러운 깊이 비디오를 생성하는 기법을 제안한다.

2. 깊이 영상기반 3차원 형상 모델링

2-1. 실사 깊이 영상 획득 및 전처리

깊이 영상은 객체의 3 차원 위치 정보를 8 비트로 양자화하여 회색 영상으로 표현한 것으로 동적인 객체의 3 차원 정보를 실시간으로

획득하여 저장하는데 사용된다. 깊이 비디오를 획득하기 위해 적외선 센서의 TOF(Time of Flight)방식을 이용하는 깊이 카메라인 ZCam™ 을 이용한다 [11]. 깊이 비디오는 획득 과정에서 객체의 반사속성, 조명조건, 깊이 카메라의 속성 등으로 인해 많은 노이즈를 포함한다. 우선 깊이 비디오의 형상을 왜곡시키는 고주파 성분을 제거하기 위해 2 차원 영상 도메인에서 미디언 필터링을 적용하여 일부 고주파 성분을 제거하였다. 하지만 관심 객체 내에서 국부적인 형상 오차들이 그대로 남아있기 때문에 전역적으로 부드러운 깊이 영상을 생성하기 위해 다음의 제안된 알고리즘을 사용한다.

2-2. 깊이 영상 기반 3차원 형상 모델링

깊이 영상으로부터 추출된 관심 객체의 3 차원 형상을 비정형 메쉬형태로 생성하기 위해 깊이 영상기반 모델링 기법(DIBM: Depth image-based modeling)을 사용하였다 [12]. 생성된 3 차원 모델은 깊이 영상의 전역적 평활도를 향상시키는 데 사용된다.

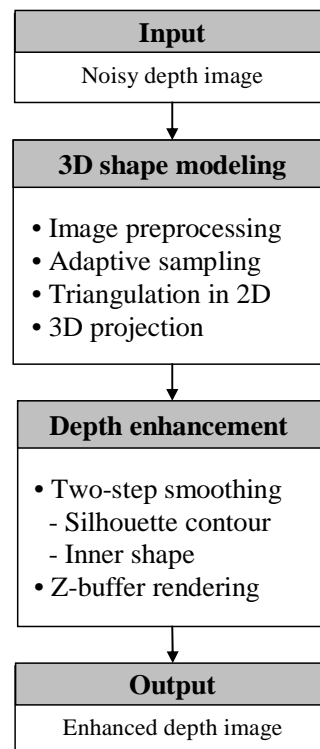


Figure 1. Flowchart of the proposed algorithm

DIBM 기법은 노이즈가 포함된 깊이 영상으로부터 부드러운 형상을 생성하기 위해 카메라 시점에 의존적인 일련의 모델링 파이프라인이다. 그림 1은 3차원 형상 모델링 과정과 깊이 영상의 품질 향상을 위한 과정을 보여주고 있다. DIBM에서 제안된 것과 같이[12], 우선 깊이 영상으로부터 관심객체를 추출하고, 깊이 영상의 1차 그래디언트 해석을 통해 깊이의 변화가 큰 영역에서 특징점을 적응적으로 추출한다 [13]. 추출된 특징점 중에서 경계 라인을 제약조건으로 하는 Delaunay 삼각화 기법을 이용하여 2차원 메쉬를 생성한다 [14]. 2차원 메쉬 모델은 pinhole 카메라 모델을 이용하여 3차원으로 투영되며, 그림 2는 깊이 영상으로부터 3차원 메쉬 모델을 생성하는 일련의 과정에 대한 결과를 보여주고 있다.

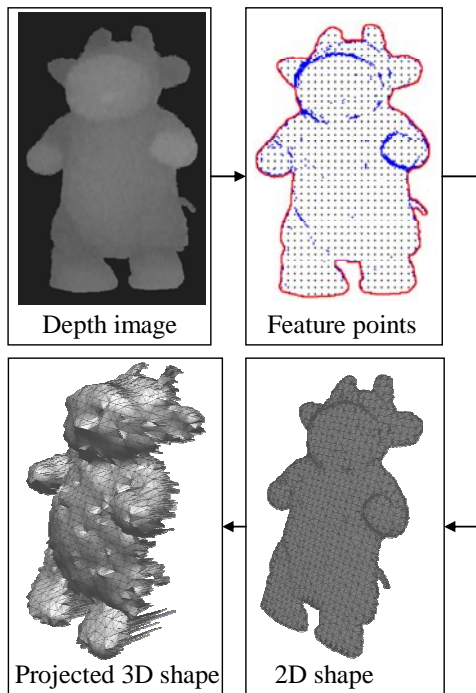


Figure 2. 3D shape modeling from a depth image

3. 경계 라인 및 내부 형상의 스무딩

3-1. 경계 라인 스무딩

깊이 영상으로부터 추출된 특징점 중에서 경계 라인은 형상 경계의 정밀도를 표현하는 중요한 역할을 한다. 하지만 깊이 영상의 오차로 인해 경계를 구성하는 점들은 그림 3(a)와 같이

중복되거나 심각한 형상 왜곡을 포함하고 있다. 이러한 경계 라인을 부드럽게 스무딩하기 위해서 우선 chain code 알고리즘을 사용하여 경계 라인에 존재하는 점들을 가까운 순서대로 정렬한다 [15]. 정렬된 경계 라인 상의 각 점에서 이웃한 점들을 이용하여 각도 크기(angular deviation)를 계산하고, 이를 기반으로 각도 크기가 큰 값을 중복된 점으로 간주하여 그림 3(b)에서 보는 것과 같이 중복된 점들을 제거한다. 그리고 2차원 가우시안 스무딩(Gaussian smoothing) 기법을 경계 라인에 적용하여 재귀적으로 스무딩함으로써 그림 3(c)에서 보는 것과 같이 부드러운 경계 라인을 생성한다. 결국 전체적인 경계 형상은 유지하면서 고주파 성분의 노이즈가 제거된 것을 볼 수 있다.

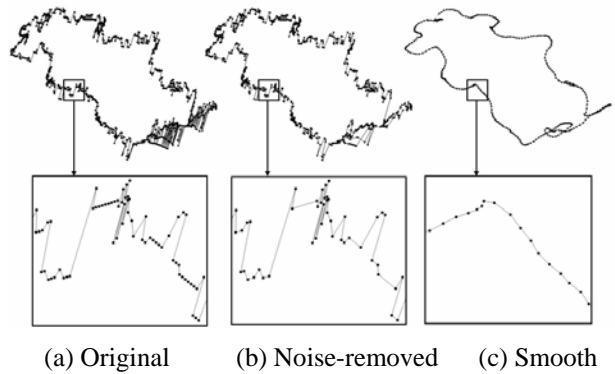


Figure 3. Smoothing of silhouette contour

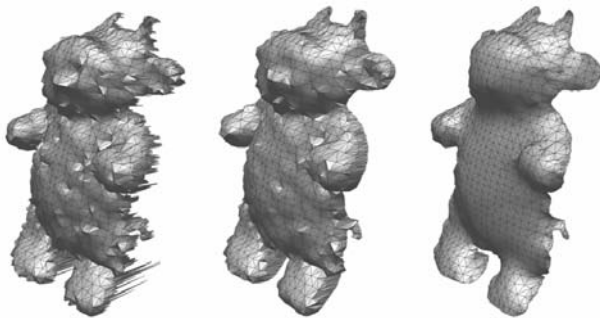
3-2. 내부 형상 스무딩

2차원 영상 도메인에서 미디언 필터링을 적용하고 경계 라인을 스무딩 하더라도, 그림 4(b)와 같이 3차원으로 투영된 3차원 모델에서 보면 모델 내부에 부드럽지 못한 형상 오차가 존재하는 것을 볼 수 있다. 이러한 형상 오차를 줄이기 위해 메쉬 형태로 표현되어 있는 내부 형상에 3차원 가우시안 스무딩을 적용한다. 라플라시안 연산자 $L(p_i)$ 는 식 1과 같이 각 정점 p_i 에서 정의된다 [16].

$$L(p_i) = \frac{1}{n_i} \sum_{j=0}^{n_i-1} (p_j - p_i), \quad (1)$$

여기서 p_j 는 정점 p_i 의 j 번째 이웃점이며, n_i 는 이웃점의 개수이다. 그림 4(c)는 경계 라인과

내부 형상이 스무딩된 결과를 보여주고 있다. 노이즈가 심한 깊이 영상으로부터 저용량의 부드러운 3 차원 메쉬 모델이 생성된 것을 볼 수 있으며, 이를 이용하여 깊이 값을 추출함으로써 전역적으로 부드러운 깊이 영상을 재생성할 수 있다.



(a) Original (b) Silhouette (c) Silhouette + inner

Figure 4. Inner shape smoothing

4. Z-buffer 렌더링을 이용한 깊이 영상의 복원

DIBM 기법과 형상 스무딩 방법에 의해 생성된 3 차원 모델은 전역적으로 부드러우면서 연속적인 곡면을 가지고 있다. OpenGL API 를 이용하여 3 차원 모델을 렌더링할 때 한 삼각형의 내부는 각 정점의 값에 의해 선형 보간된다. 즉 Z 버퍼에 있는 깊이 값도 마찬가지로 보간되기 때문에, 이를 이용하여 3 차원 모델로부터 깊이 영상을 복원하게 되면 전역적으로 부드러운 깊이 영상을 복원할 수 있다.

Z 버퍼에 있는 깊이 값은 다음의 OpenGL 함수로부터 추출될 수 있다.

```
void glReadPixels(x, y, width, height,
format, type, *pixels)
```

여기서 x 와 y 는 프레임 버퍼로부터 읽혀지는 첫 번째 화소의 윈도우 좌표이며, $width$ 와 $height$ 는 영상의 가로, 세로 크기를 의미한다. 화소 데이터의 포맷은 $format$ 으로 표현되고, 만약 $format$ 값이 $GL_DEPTH_COMPONENT$ 이면 추출되는 화소 값은 0 부터 1 까지 정규화된 깊이

값을 의미한다. 복원되는 깊이 영상의 깊이 값은 $[0,1]$ 로 표현되기 때문에 이를 원래 영상의 깊이 범위와 일치하도록 해야 한다. 이를 위해 원래 영상의 최대 및 최소 깊이 값을 $Original_{max}$ 와 $Original_{min}$ 라 하고, 정규화된 깊이 값을 $Norm_i$ 라 한다. 정규화된 깊이 값 중에서 최대, 최소 값을 $Norm_{max}$, $Norm_{min}$ 라 하고, 복원되는 깊이 값을 $Enhanced_i$ 라 한다. 그러면 복원되는 깊이 값은 식 2 와 같이 계산된다. 모든 픽셀에 대해서 제안된 식을 적용하여 전역적으로 부드러운 깊이 영상을 생성하게 된다.

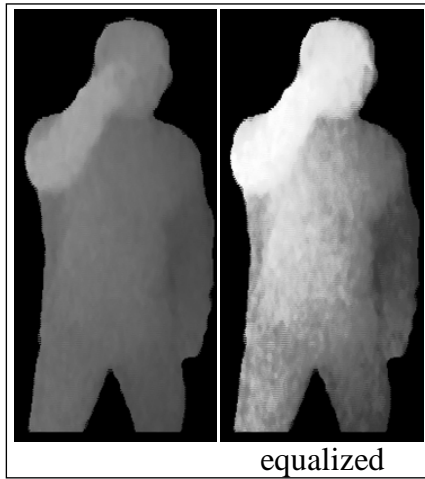
$$Enhanced_i = Original_{max} - NF_i(Original_{max} - Original_{min})$$

$$where, NF_i = \frac{Norm_{max} - Norm_i}{Norm_{max} - Norm_{min}} \quad (2)$$

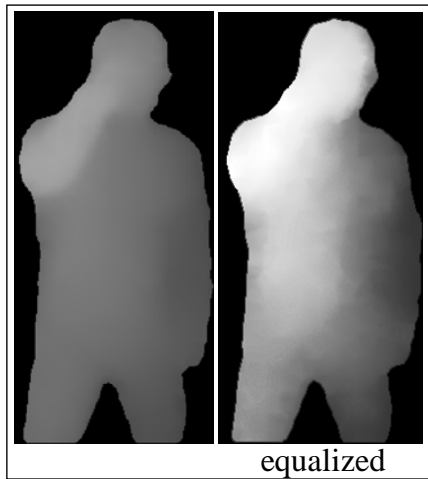
5. 실험 결과 및 평가

5-1. 형상 평활도 향상

제안된 알고리즘을 적용하기 위해 ZCamTM 을 이용하여 287 장으로 구성된 10 초 정도의 깊이 비디오를 획득하였다. 그림 5 는 발췌된 영상 한 장에 대한 노이즈 필터링 결과를 보여주고 있다. 그림 5(a)의 왼쪽 영상은 원래 영상을, 오른쪽 영상은 가시화를 위해 평활화된 영상을 보여 주고 있으며, 형상 내부에 고주파 성분의 노이즈가 많이 포함된 것을 볼 수 있다. 즉 불규칙한 형상 노이즈로 인해 렌더링이나 촉감 제시 장치를 통한 햅틱 상호작용에서 매끄럽지 못한 결과를 초래한다. 하지만 형상 모델링을 이용한 필터링 기법을 적용하면, 그림 5(b)에서와 같이 대부분의 고주파 노이즈가 제거된 부드러운 깊이 영상을 획득할 수 있다. 세부적인 형상이 어느 정도 사라졌지만 전체적인 평활도(smoothness)는 크게 향상된 것을 볼 수 있다.



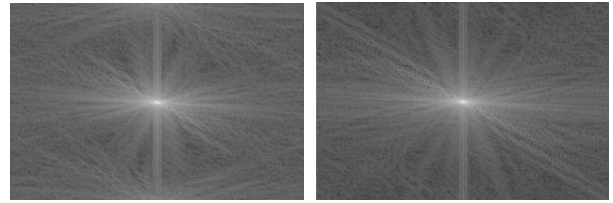
(a) Original depth image



(b) Enhanced depth image

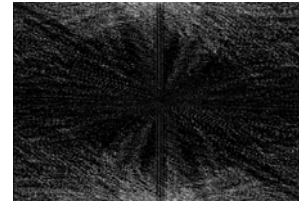
Figure 5. An experimental result of depth enhancement

그림 6 는 주파수 해석을 통해 형상의 평활도 향상에 대한 검증 결과를 보여준다. 그림 6(a)는 원래 영상을 푸리에 변환한 것이며, 그림 6(b)는 복원된 영상을 변환한 것으로 고주파 성분이 많이 사라진 것을 볼 수 있다. 분명한 비교를 위해 그림 6(c)는 변환된 두 영상의 차영상을 보여주고 있다. 차영상에서와 같이 대부분의 고주파 성분이 제거되었기 때문에 가장자리 부분에서 밝은 값을 나타내고 있고, 저주파 성분은 유지되었기 때문에 중간 부분에서 어두운 값을 가지는 것을 볼 수 있다. 즉 전체 형상은 유지하면서 고주파로 표현되는 형상 노이즈가 효과적으로 제거되었음을 알 수 있다.



(a) Original depth image

(b) Enhanced depth image



(c) Difference image that subtracts (b) from (a)

Figure 6. Evaluation of shape smoothness using frequency analysis

5-2. 압축 효율 검증

제안된 알고리즘으로 생성된 깊이 비디오의 압축 효율에 대한 영향을 평가하기 위해 H.264 표준을 이용하여 압축실험을 수행하였다. 비교 실험은 원래 깊이 비디오와 미디언 필터링만 적용된 비디오, 그리고 제안된 기법에 의해 노이즈가 제거된 깊이 비디오로부터 각각 100 장의 영상을 추출하여 사용하였다. 표 1 은 압축 실험의 결과를 보여주고 있다.

Table 1. The compression efficiency

Depth sequence (100 images)	Original	Median-filtered	Enhanced
Compressed size	154 KB	102 KB	82 KB
Absolute compression ratio	0.449%	0.297%	0.239%
Relative compression ratio	100%	151%	185%

Original depth images (BMP file format) : 34,300KB

압축되지 않은 BMP 파일로 표현되었을 때 34,300 KB 였던 100 장의 영상 시퀀스를 압축하였을 때, 원래 비디오, 미디언 필터링된 비디오, 제안된 기법에 의해 필터링된 비디오가 각각 154 KB, 102 KB, 82 KB 로 압축되는 것을 알 수 있었다. 결국 원래 비디오에 비해 제안된

방법에 의한 비디오의 압축률이 상대적으로 185% 정도가 향상된 것을 볼 수 있으며, 이는 형상의 부드러움이 크게 향상되었기 때문이다.

6. 결론

본 논문에서는 3 차원 형상 모델링 기법과 Z-buffer 렌더링 기법을 이용하여 동적인 실제 객체의 깊이 비디오에 포함된 고주파 성분의 노이즈를 제거하고, 전역적으로 부드러운 깊이 비디오를 재 생성하였다. 이는 방송 시스템에서 고품질의 입체 영상을 생성하거나 깊이 비디오 기반 햅틱 렌더링에 활용될 수 있다.

후기

본 연구는 광주과학기술원 실감방송 연구센터(RBRC)를 통한 정보통신부 대학 IT 연구센터(ITRC) 사업과 실감콘텐츠 연구센터(ICRC)를 통한 과학기술부 특정연구개발 사업의 지원에 의한 것입니다.

참고문헌

[1] M.M. Oliveira, "Image-Based Modeling and Rendering Techniques: A Survey", RITA - Revista de Informática Teórica e Aplicada, Vol. IX, No. 2, October 2002, pp. 37-66.

[2] M. Levoy, T. Whitted, "The Use of Points as a Display Primitive", Technical Report 85-022, Computer Science Department, UNC at Chapel Hill, 1985.

[3] C. Zhang and T. Chen, "A Survey on Image-Based Rendering - Representation, Sampling and Compression", EURASIP Signal Processing: Image Communication, Vol. 19, No. 1, 2004, pp. 1-28.

[4] A. Ignatenko, A. Konushin, "A Framework for Depth Image-Based Modeling and Rendering," 13th International Conference on Computer Graphics and Vision GraphiCon-2003 Moscow, September 5 -10, 2003, pp.169-172.

[5] M. Lemmens, "A survey on stereo matching techniques", Proc. of ISPRS, vol. 27, 1988, pp. 11-23.

[6] L. Falkenhagen, "Depth estimation from stereoscopic

image pairs assuming piecewise continuous surfaces", in Image Processing for Broadcast and Video Production, November 1994, pp. 115-127.

[7] M. Ziegler, L. Falkenhagen, R. ter Horst, D. Kalivas, "Evolution of stereoscopic and three-dimensional video", Image Communication, vol. 14, 1998, pp. 173-194.

[8] J. Ohm, K. Grüneberg, E. Hendriks, E.Izquierdo, D. Kalivas, M. Karl, D. Papadimitos, A. Redert, "A realtime hardware system for stereoscopic videoconferencing with viewpoint adaptation", Signal Processing: Image Communication, vol. 14, 1998, pp. 173-363.

[9] G.J. Iddan and G. Yahav, "3D Imaging in the studio and elsewhere", Proceedings of SPIE, Vol. 4298, 2001, pp.48-55.

[10] A. K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Inc., 1989.

[11] ZCamTM add-on, <http://www.3dvsystems.com>, 2005.

[12] Seung-Man Kim, Jeung-Chul Park, Kwan H. Lee, "A Single Depth Image-Based Complete Mesh Generation", Proceedings of the 6th Israel-Korea Bi-National Conference, Israel, November 2005, pp. 67-72.

[13] Seung-Man Kim, Jeung-Chul Park, Kwan H. Lee, "Depth Map-based 3D Mesh Modeling Using an Adaptive Sampling Method," Proceedings of the 2005 Korean Signal Processing Conference, Vol. 18, No. 1, 2005, pp.157.

[14] V. Domiter, "Constrained Delaunay triangulation using plane subdivision," Proceedings of the 8th Central european seminar on computer graphics, Budmerice, April 19-21, 2004, pp. 105-110.

[15] J.A. Saghri, H.Freeman, "Computer Processing of Line Drawing Images", IEEE Trans. PAMI, 1981, pp. 533-539.

[16] M. Eck, T. DeRose, T. Duchamp, H. Hoppe, M. Lounsbery, W. Stuetzle, "Multiresolution analysis of arbitrary meshes", SIGGRAPH proceedings, 1995, pp.173-182.