

웨이블릿 변환을 이용한 빠른 3D modeling

고병철*, 노윤향*, 이해성*, 변혜란*, 유지상**

*연세대학교 컴퓨터과학과

**광운대학교 전자공학과

e-mail : soccer1,yhrho,geneel,hrbyun @aipiri.yonsei.ac.kr

jsyoo@daisy.kwangwoon.ac.kr

Fast 3D reconstruction using wavelet transform

ByoungChul Ko*, Yoon-Hyang Rho*, Hae-Sung Lee*, Hyeran Byun*, Jisang Yoo**

*Dept. of Computer Science, Yonsei University

**Dept. of Electronics Engineering, Kwangwoon University

요 약

본 논문에서는 웨이블릿 변환을 이용하여 추정된 변위 벡터와 이를 이용한 물체의 분할을 통해 특징 점을 추출하고 3 차원 와이어 프레임(wire-frame)을 생성하는 알고리즘을 제안한다. 우선, 웨이블릿 변환을 이용하여 빠른 시간 안에 변위를 측정하고, 이를 통해 배경과 물체를 분리해 내었다. 그런 뒤에, 변위 벡터를 이용하여, 깊이 정보를 추정해 내고, 동시에 물체로부터 두드러진 특징 값들을 추출하여 3 차원 와이어 프레임 생성을 위한 거리 값으로 사용하였다. 마지막으로, 일반적인 delaunay triangulation 에서 생길 수 있는 오 정합을 본 논문에서 제안하는 전경/배경 분할 알고리즘을 이용하여 제거 하여 정확한 3 차원 모델을 생성하였다. 아울러, 본 논문에서 제안하는 웨이블릿을 이용한 빠른 3D 모델링 방법을 원 영상을 이용한 방법과 비교하여, 더 좋은 결과를 보여줌으로써, 계산 시간 뿐만 아니라 정확성에서도 만족할 만한 결과를 얻을 수 있었다.

1. 서론

최근 들어, 실세계와 가상 세계를 혼합한 AR(Augmented Reality), 3 차원 화상회의 등과 같은 멀티미디어 응용 분야들이 새로운 관심분야로 부각되고 있다. 하지만, 실제로 사용되고 있는 3 차원 모델들은 여러 가지 모델링 툴들을 이용하여 전체적으로 사람의 손에 의해 모델링 되므로 상당한 노력과 시간 및 장비를 필요로 한다. 이러한 문제점들을 해결하기 위해서, 대안적인 방법으로 영상을 이용한 모델링 방법들이 연구되고 있다. 이 방법은 사람의 양안 시차를 이용하여, 카메라의 위치를 이동시키고, 위치에 따라 조금씩 다르게 인식되는 사물의 모양을 통해서, 3 차원 정보를 얻는 방법이다. 이와 같은 2 차원 스테레오 영상 혹은 다 시점 영상을 통해서 사물의 3 차원 정보를 얻기 위해서는 다음과 같은 단계가 필요하다 [1]. 첫째, 영상으로부터 배경과 사물을 구분해 내야 한다. 이 작업을 위해서 우선, 2 차원 스테레오 영상으로부터 변위(disparity)를 측정한다. 그런 뒤에, 일정한 변위가 발생한 영역만을 사물로 간주하고 이

변위를 이용하여 3 차원 깊이 정보를 계산한다. 하지만, 이 과정은 정합 과정이 높은 정밀도의 요구와 많은 연산시간으로 인하여 고속 연산을 요구하는 실시간 구현에 장애가 되는 단계이다 [2]. 따라서, 효율적인 3 차원 모델링을 위해서는 정확하고 빠른 변위 측정이 필수적이다. 두 번째로, 카메라 캘리브레이션 작업이 필요하다. 이 작업을 통해서, 작업에 사용된 카메라들의 외부 파라메타를 비교해 봄으로써 동일 직선상에 평행한 광축을 갖도록 카메라가 배치 되었음을 확인 할 수 있다. 세 번째로, 분할된 사물로부터 특징 점을 추출하는 작업이 필요하다. 이렇게 추출된 특징 점들은 마지막 네 번째 단계에서, 모델링(modeling)과 렌더링(rendering)작업의 성능을 좌우하므로, 적절한 수의 특징 점을 추출해 내도록 해야 한다.

본 논문에서 제안하는 주요 알고리즘은 다음과 같다.

- 웨이블릿을 이용한 빠른 변위 벡터 측정,
- SUSAN edge detector 를 응용한 정확하고 의미

있는 특징 점 추출,

- Delaunay triangulation 을 이용한 와이어 프레임 (Wire-frame) 생성과 오 정합에 의한 와이어 프레임 제거를 위한 전경/배경 분할 알고리즘의 제안.

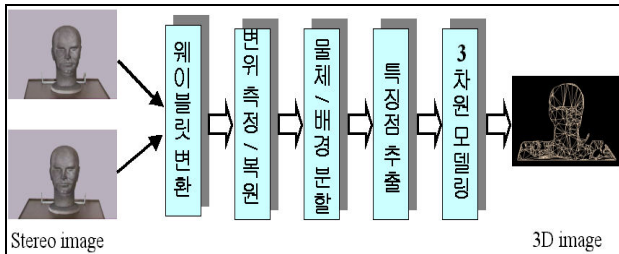


그림 1. 본 논문의 3D 모델링 순서도

본 논문은 다음과 같이 구성된다. 2 장에서는 본 논문에서 제안하는 웨이블릿 기반 변위 추출 알고리즘을 소개하고, 3 장에서는 모델링을 위한 특징점 추출을 설명한다. 4 장에서는 추출된 특징점을 이용한 와이어 프레임(wire-frame) 생성 과정을 설명하고 실험 결과를 제시한다. 마지막으로, 5 장에서는 본 논문에서 제안하는 방법에 대한 결론 및 향후 개선 방향을 제시한다.

2. 웨이블릿 기반 변위 추출 알고리즘

영상으로부터 3 차원 정보를 추출하기 위해서는 두 영상간의 대응되는 특징 점을 선별해 내야 한다. 이 두 점들간의 거리가 변위가 되고, 이 변위 정보로부터 반비례 관계에 있는 깊이 정보를 얻을 수 있다. 변위 측정을 위한 방법에는 영역 기반 방법(Area-based stereo), 특징 기반 방법(Feature-based stereo), 화소 기반 방법(Pixel-based stereo) 등이 있다 [5]. 이 중에서 영역 기반 방법은 기준 화면의 모든 화소에 대해 대응점을 구함으로 연산 시간이 많이 걸리고, 폐색(occlusion)영역이나 화소 값의 차이가 작은 영역에서는 그 정확도가 떨어지지만, 단순하고 조밀한 변위를 구할 수 있는 장점이 있다. 본 논문에서는 영역기반 방법 중 SSD(Sum of Squared Distance) 방법을 웨이블릿 변환된 저주파 영역에 적용시켜 변위를 측정하였다. 웨이블릿 변환으로 얻어지는 저주파 영역은 원 영상을 그대로 몇 배 작게 줄여 놓는 것으로, 모든 웨이블릿 부 영역 중에서 가장 중요한 정보를 가지고 있다 [3]. 따라서, 웨이블릿 변환의 저주파 영역을 이용하여, 영역기반 방법의 최대 단점인 속도 문제를 해결하고, 잡음이나 작은 범위의 폐색 영역에 강건한 변위를 측정 할 수 있다.

본 논문에서는 영역 기반 방법 중 SAD(Sum of Absolute Distance), SSD, NCC(Normalized Correlation Coefficient) 방법과 웨이블릿 기반 SSD, NCC 방법의 성능을 각각 비교하여 보았다(표-1).

$$SAD = \sum_w |R_{x,y} - L_{x+d,y}| \quad (1)$$

$$SSD = \sum_w [R_{x,y} - L_{x+d,y}]^2 \quad (2)$$

$$NCC = \frac{\sum_w [R_{x,y} - \overline{R_{x,y}}] \times [L_{x+d,y} - \overline{L_{x+d,y}}]}{\sqrt{\sum_w [R_{x,y} - \overline{R_{x,y}}]^2} \times \sqrt{\sum_w [L_{x+d,y} - \overline{L_{x+d,y}}]^2}} \quad (3)$$

$R_{x,y}$: 우 영상의 (x,y)위치에서의 화소 밝기 값

$L_{x+d,y}$: 우 영상에서 d 만큼 이동된 좌 영상 위치에서의 화소 밝기 값

$\overline{R_{x,y}}$: 우 영상에서 상관윈도우 내의 화소 평균값

$\overline{L_{x+d,y}}$: 우 영상에서 d 만큼 이동된 좌 영상 상관윈도우 내의 화소 평균값

	계산시간(단위:초)		추출된 특징점 수	
	1번영상	2번영상	1번영상	2번영상
SAD	22	19	784	398
SSD	97	89	758	408
NCC	273	285	708	391
WT_SSD	5	5	926	599
WT_NCC	9	8	943	626

표 1. 영역기반 변위 측정 테스트 결과

이때, 탐색 윈도우의 크기는 32x16, 상관 윈도우의 크기는 16x16 으로 정의하였다. 또한 웨이블릿 변환을 위해서 length-2 의 Daubechies 웨이블릿 필터를 사용하여, 1/4 크기로 줄여든 1 단계의 저주파 영역으로부터, 탐색 윈도우 크기를 16x8, 상관윈도우를 8x8 로 정의하여 실험하였다. 또한, 표에 나타난 웨이블릿 변환을 이용한 계산 시간은 웨이블릿 변환을 포함한 시간이다. 이때, 탐색윈도우와 상관윈도우의 y 축 길이가 동일 한 것은 stereo 영상 자체가 x 축의 위치 이동만 존재하기 때문이다. 본 논문에서는 실험을 위해 웹(www-syntim.inria.fr/syntim/analyse/paires-eng.html)상으로부터 얻은 2 개의 영상을 사용하였다.



(a) 1번 영상(좌, 우) (b) 2번 영상 (좌, 우)

그림 2. 실험에 사용된 stereo 영상 (320x210)

2.1. 변위 히스토그램을 이용한 배경 분할

변위를 측정된 후에 변위 영상에 대하여 메디안 필터를 적용한다. 이렇게 함으로써 잡음 등에 의해 잘못 찾아진 변위를 제거해 준다.

다음으로 웨이블릿 변환된 저주파 영역으로부터 구

해진 변위 값을 원래 영상의 변위 값으로 복원 시켜주는 작업이 필요하다. 본 논문에서 사용한 저주파 영역은 원 영상에 대해 1/4 크기로 줄어든 영상이므로 그림 3 과 같이, 웨이블릿 변환과 SSD 방법으로 구해진 변위 벡터에 2를 곱함으로써 원 영상의 변위 벡터로 복원 할 수 있다.

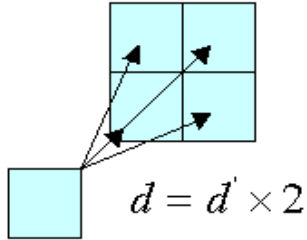


그림 3. 변위벡터 보정

지금까지의 과정을 통해 얻어진 변위 영상은 그림 4 와 같다.

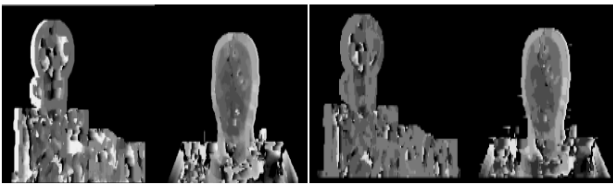


그림 4. 변위 영상(좌:SSD, 우:WT_SSD)

다음으로, 변위 벡터를 이용하여 변위 히스토그램을 구성하여 물체와 배경을 구분해 낸다. 이것은 일반적으로 스테레오 영상에서 변위 정보가 전체 탐색 영역에 고르게 분포되어 있지 않고, 배경과 물체의 변위에 해당하는 값만을 갖는 특성을 이용한 것이다 [2].

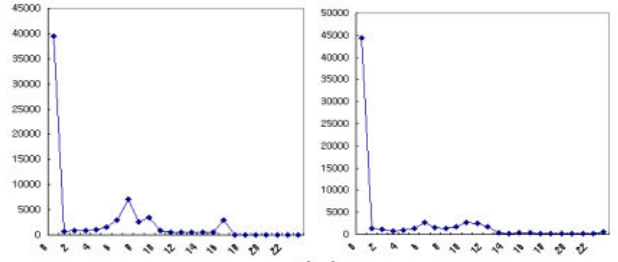
즉, Stereo 영상에서 물체는 일반적으로 카메라의 앞부분에 위치함으로 변위 벡터가 크게 나타나고, 배경의 경우는 뒷부분에 위치함으로 0 에 가까운 값을 가지게 된다. 따라서 이러한 특성을 이용하여, 계산된 변위 히스토그램으로부터 물체와 배경을 구분해 낼 수 있다.

그림 3 에 의해 원 영상의 변위벡터를 보정하고, 그림 5 로부터 배경과 물체영역을 구분해 내고 나면, 변위 벡터와 카메라간의 거리(b), 카메라와 평면사이의 거리(f)를 이용하여 물체의 Z 좌표를 식(4)를 이용하여 추정해 낼 수 있다.

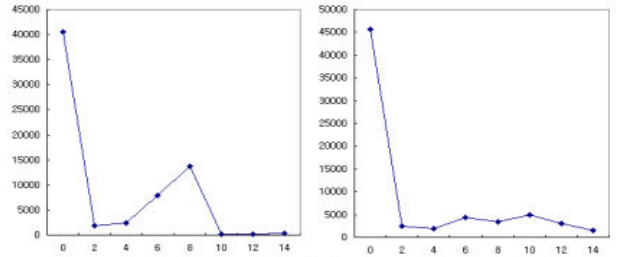
$$Z = \frac{bf}{X_L - X_R} = \frac{bf}{d} \quad (4)$$

3. SUSAN 에지 검출기를 이용한 특징점 추출

특징점이란 영상안의 물체가 가지고 있는 두드러지는 에지(edge)나 모서리(corner)에 해당하는 점으로, 이



(a)



(b)

그림 5. (a) 원 영상(1-2 번)에서의 변위 히스토그램 (b) 웨이블릿 변환된 영상에서의 변위 히스토그램 (가로축 - 변위, 세로축 - 변위 빈도수)

특징점 들을 이용하여, 움직임 추정, 3 차원 인식, 모델링과 같은 분야에 적용시킬 수 있다.

본 논문에서는 특징점 추출을 위한 여러 방법 중 SUSAN edge detector 방법을 사용하였다. SUSAN 의 방법은 edge-map 이나 단순히 gray-level 의 밝기 값을 이용한 방법보다 잡음에 덜 민감하고, 비교적 정확한 특징 값을 추출해 낼 수 있다 [4].



그림 6. 물체의 다른 부분에 위치한 4 개의 원형 마스크

그림 6 에서 보듯이 SUSAN 방법에서는 원형 마스크를 사용하고 있다. 이때 영상 안의 한 점 p(x, y)가 마스크의 중심점이 되고, 이 중심점의 화소 값과 마스크 안의 다른점 q 와의 차이 값을 식(5)를 통해 계산한다.

$$C(p, q) = e^{-\frac{(I(p)-I(q))^6}{t}} \quad (5)$$

이렇게 계산된 C(p, q)의 합이 마스크 크기의 반(W/2)보다 작을 때 p(x, y)점은 특징 점이 된다. 본 논문에서는 특징 점을 위한 임계 값 t 를 25 로 정의하고, 특징 점이 물체의 지역적(local)인 부분에서 과도하게 많이 추출되는 것을 막기 위해 원 논문에서 사용한 마

스크(5x5)보다 큰 11x11 크기의 마스크를 사용하였다.

본 논문에 의해 추출된 특징 점은 그림 7 과 같다. 이때, 특징 점이 되기 위한 조건에서 W/2 를 W/3, W/4 로 조정함으로써, 특징 점의 수를 조절 할 수 있다.



(a)



(b)

그림 7. SUSAN 방법에 의한 특징 점 추출 결과
(a) 원 영상(1-2 번)에서 추출된 특징 점 (b) 웨이블릿 변환된 영상에서 추출된 특징 점

4. 3D 모델링

추출된 특징 값과 이 특징 값들의 x, y, z 좌표 값들을 이용하여 실제적으로 3 차원으로 모델링 하는 과정이 필요하다. 본 논문에서는 3 차원 모델 생성을 위하여 Delaunay triangulation 방법을 이용하였다.

우선, 영상으로부터 추출된 최상위 특징 점을 초기 와이어 그리드(wire-grid) 생성을 위한 시작 노드로 사용하고, 2.1 장에서 계산된 초기 2 차원 와이어 그리드 노드들의 깊이 정보를 이용하여 초기 3 차원 와이어 프레임을 생성한다. 초기 와이어 프레임이 생성된 후에, 추가적인 노드들은 triangular planar patch 와 물체의 다른 특징 점들과의 거리 계산에 의해 결정된다. 즉, 거리는 triangular patch 가 놓여 있는 평면(plane)에 수직 방향으로 계산되고, 물체의 특징 점과 triangular patch 와의 거리가 미리 정의된 임계 값을 초과한다면, 이 점을 새로운 노드로 선언하고, 와이어 프레임에 추가시킨다. 새로운 노드를 갖게 된 triangular patch 는 다시 새로운 노드에 vertex 를 갖는 여러 개의 sub-triangle 들로 나뉘어 지게 된다. 하지만, 일반적으로 사용되는 이 방법은 몇 가지의 문제점을 가지고 있다. 그림 8 에서 보는 것과 같이 대부분의 경우, 와이어 프레임이 정확히 생성되지만, 일부분의 경우, triangular patch 와 특징 점들간의 거리 값을 이용하므로, 오정합된 와이어 프레임을 생성할 수 있다.

하지만, 정확한 텍스처 매핑을 위해서는 필수적으로 정확한 와이어 프레임이 생성되어야 한다. 따라서, 이러한 오 정합된 와이어 프레임을 제거해 주기 위해 본 논문에서는 전경/배경 분할(Foreground/background segmentation) 알고리즘을 제안한다.

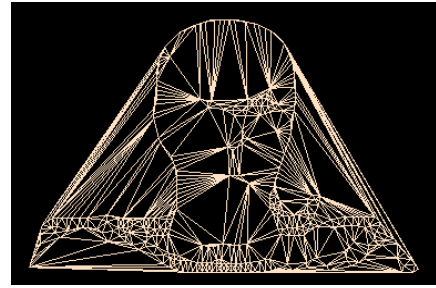
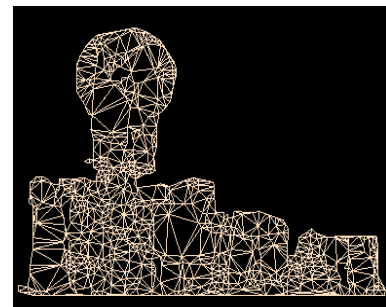


그림 8. Delaunay triangulation 결과

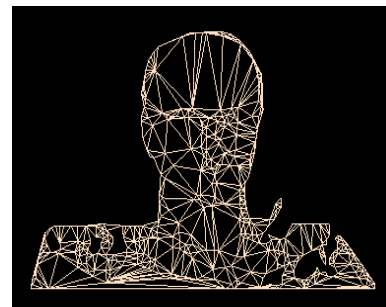
4.1. 오 정합 와이어 프레임 제거를 위한 전경/배경 분할 알고리즘

우선 변위 히스토그램으로부터 얻어진 배경 영역을 버퍼에 저장 시켜둔다. 그런 뒤에 delaunay triangulation 을 이용하여 얻어진 각각의 와이어 프레임으로부터 중심 좌표 값을 구한다. 이렇게 구해진 각 와이어 프레임의 중심 좌표 값을 버퍼에 저장된 배경영역과 매칭 시키는 작업을 실시한다. 그리고 만약, 중심 좌표 값이 배경 영역에 포함된다면, 이 와이어 프레임은 제거되거나 또는 새로운 와이어 프레임으로 선언되지 않게 된다. 이 작업은 delaunay triangulation 작업에 몇 개의 코드만을 추가하여 이루어짐으로 전체적인 와이어 프레임 생성시간에는 거의 영향을 주지 않는다.

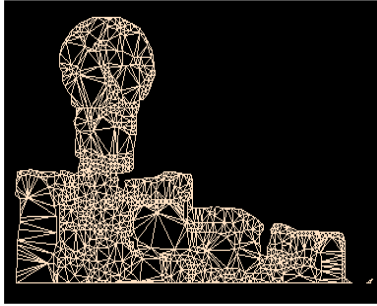
이 과정을 거쳐 최종적으로 생성된 3D 모델은 그림 9 와 같다.



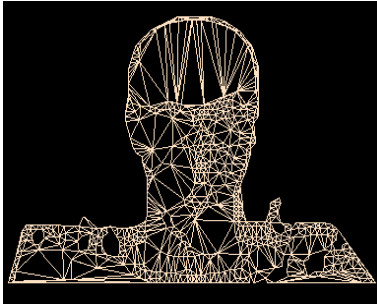
(a)



(b)



(c)



(d)

그림 9. 전경/배경 분할 알고리즘 적용 결과
(a)(b): 원 영상(1-2 번)을 이용하여 생성된 와이어 프레임, (c)(d): 웨이블릿 변환을 이용하여 생성된 와이어 프레임

5. 결론 및 향후 연구과제

본 논문에서는 웨이블릿 변환을 이용하여 빠른 시간 안에 변위를 측정하고, 이를 이용하여 배경과 물체를 분리해 내었다. 그런 뒤에, 이 변위를 통해, 깊이 정보를 추정해 내고, 동시에 물체로부터 두드러진 특징 값들을 추출하여 3 차원 와이어 프레임 생성을 위한 거리 값으로 사용하였다. 그리고 마지막으로, 일반적인 delaunay triangulation 에서 생길 수 있는 오 정합을, 본 논문에서 제안하는 전경/배경 분할 알고리즘을 이용하여 제거할 수 있었다. 아울러, 본 논문에서 제안하는 웨이블릿을 이용한 빠른 3D 모델링 방법을 원 영상을 이용한 방법과 비교하여, 더 좋은 결과를 보여 줌으로써, 계산 시간 뿐만 아니라 정확성에서도 만족할 만한 결과를 얻을 수 있었다. 하지만 본 논문에서는 보다 좋은 결과를 위해서 몇 가지 사항이 추가적으로 보완되어야 한다.

1. 보다 정확한 변위 벡터가 측정 되어야 한다. 현재 본 논문에서는 영역기반 방법을 웨이블릿에 적용했지만, 영역기반 방법만을 사용할 경우 패색 구간이나, 빛의 변화가 심한 부분에서는 잘못된 결과를 나타낼 수 있으므로 이를 보완할 수 있는 추가적인 알고리즘이 적용되어야 한다.
2. 정확한 카메라 캘리브레이션 작업이 이루어져야 한다. 현재 본 논문에서는 웹상에서 공개된 데이터를 사용하고 있으므로, 영상이 작아지거나, 변형될 경우 카메라 캘리브레이션 수치들도 변하게 된다. 따라서, 정확한 깊이 정보를 추정

해 내기 위해서는 정확한 카메라 파라메타들을 얻을 수 있어야 한다.

3. 텍스처 매핑 작업과 렌더링 작업이 이루어져야 한다. 사용자에게 보다 자연스럽고 실제적인 느낌을 주기 위해서는 실제 영상의 텍스처 매핑 작업과 렌더링 작업은 필수적이다. 따라서, 최종 단계에는 이 작업에 대한 연구가 진행되어야 한다.

참고문헌

- [1] Ebroul Izquierdo M. and Silko Kruse, "Image analysis for 3D modeling, rendering, and virtual view generation," Computer vision and Image understanding, Vol. 71, No. 2, 1998.
- [2] 김남규, 이광도, 김형근, 차균현, "저 해상도 변위 히스토그램을 이용한 고성능 변위정보 추출 알고리즘", 전자공학회 논문지 제 35 권 제 3 호, pp. 131-143, 1998.
- [3] 고병철, 이해성, 변혜란, "다중 특징을 이용한 영상 및 비디오 내용기반 검색 시스템," 정보과학회 논문지(B), 제 26 권 12 호, pp. 1519-1530, 1999, 12 월.
- [4] Stephen M. Smith, "SUSAN-A new approach to low level image processing," Int. Journal of Computer Vision 23(1), 45-78, 1997
- [5] Guo-Qing Wei, Wilfried Brauer, Gerd Hirzinger, "Intensity and gradient-based Stereo matching using Hierarchical gaussian Basis Function," IEEE transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, 11, 1998