

2단계 하이브리드 방법을 이용한 2D 스테레오 영상의 3D 모델링

(3D Modeling from 2D Stereo Images using 2-Step Hybrid Method)

노윤향[†] 고병철^{**} 변혜란^{***} 유지상^{****}

(Yoon-Hyang Rho) (Byoung-Chul Ko) (Hyeran Byun) (Ji-Sang Yoo)

요약 일반적으로 2D 스테레오 영상으로부터 3차원 모델링을 위해서는 정확한 변위 측정이 필수적이다. 기존의 스테레오 영상에서 변위 측정 방식은 전체 영상에 대하여 정합 연산을 수행함으로써 많은 연산 시간과 함께 높은 오정합 확률의 문제가 있다. 본 논문에서는 스테레오 영상에서의 변위 벡터가 전체 탐색 범위 안에 골고루 분포되어 있지 않고 배경과 물체의 변위에 해당하는 값만을 갖는다는 특성을 이용하여 스테레오 영상을 웨이블릿 변환을 하고 1/4 크기로 줄어든 저주파 영역으로부터 영역 기반 방법을 이용하여 대략적인 변위 영역을 구한다. 대략적인 변위 벡터로부터 변위 히스토그램을 생성하고, 이를 이용하여 전경과 배경을 분할 한 뒤, 다시 전경 영상만을 원 영상으로 복원하여 화소의 밝기값이 아닌 2차 미분값을 이용한 화소기반 방법을 통해 조밀한 변위를 구하는 2단계 하이브리드 방법을 제안한다. 또한, 분할된 전경 영역으로부터, 특징점들을 뽑아내고 변위 벡터와 카메라 파라미터를 이용하여 특징점들의 깊이 정보를 추정해 내는 3차원 모델링 과정을 제시한다.

본 논문에서 제안한 방법을 적용할 경우, 기존의 영역 기반 방법의 문제점인 계산 시간 문제를 상당 부분 단축시킬 수 있고, LOG 필터를 통한 2차 미분값을 이용한 화소기반 방법을 추가함으로써, 정밀한 변위를 구할 수 있다. 또한 교차 일치성 검사를 통해 잘못된 변위를 제거하고, 폐색 영역들을 검사할 수 있다. 아울러 3차원 모델링 과정에서, 기존의 Delaunay 삼각측량법의 문제점인 오정합 문제를 전경/배경 분할 알고리즘을 제안함으로써 효과적으로 해결 할 수 있다.

Abstract Generally, it is essential to estimate exact disparity for the 3D modeling from stereo images. Because existing methods calculate disparities from a whole image, they require too much computational time and bring about the mismatching problem. In this article, using the characteristic that the disparity vectors in stereo images are distributed not equally in a whole image but only exist about the background and object, we do a wavelet transformation on stereo images and estimate coarse disparity fields from the reduced lowpass field using area-based method at first-step. From these coarse disparity vectors, we generate disparity histogram and then separate object from background area using it. Afterwards, we restore only object area to the original image and estimate dense and accurate disparity by our two-step pixel-based method which does not use pixel brightness but use second gradient. We also extract feature points from the separated object area and estimate depth information by applying disparity vectors and camera parameters. Finally, we generate 3D model using both feature points and their z coordinates. By using our proposed method, we can considerably reduce the computation time and estimate the precise disparity through the additional pixel-based method using LOG filter. Furthermore, our proposed foreground/background method can solve the mismatching problem of existing Delaunay triangulation and generate accurate 3D model.

· 본 논문은 정보 통신부에서 지원하는 대학 기초 연구 지원 사업으로 수행된 것임

† 비 회 원 : 연세대학교 컴퓨터과학과
yhrho@aipiri.yonsei.ac.kr

** 학생회원 : 연세대학교 컴퓨터과학과
soccer1@aipiri.yonsei.ac.kr

*** 중신회원 : 연세대학교 컴퓨터과학과 교수
hrbyun@aipiri.yonsei.ac.kr

**** 비 회 원 : 광운대학교 전자공학과 교수
jsyoo@daisy.kwangwoon.ac.kr

논문접수 : 2000년 5월 16일

심사완료 : 2001년 5월 18일

1. 서론

최근 들어, 실세계와 가상 세계를 혼합한 증감현실 (Augmented Reality), 3차원 화상 회의, 원격 의료 서비스 등과 같이 가상환경을 이용한 멀티미디어 응용 분야들이 새로운 관심분야로 부각되고 있다. 하지만, 실제로 사용되고 있는 3차원 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 2차원 영상을 이용한 3차원 모델링 방법들이 활발히 연구되어 오고 있다. 이 방법은 사람의 양안 시차를 이용하여, 카메라의 위치를 이동시키고, 위치에 따라 조금씩 다르게 인식되는 사물의 모양을 통해서, 3차원 정보를 얻는 방법이다. 이와 같이 2차원 스테레오 영상 혹은 다시점 영상을 통해서 사물의 3차원 정보를 얻기 위해서는 다음과 같은 단계가 필요하다[1]. 첫째, 영상으로부터 배경과 사물을 구분해 내야 한다. 이 작업을 위해서 우선, 2차원 스테레오 영상으로부터 변위를 측정한다. 그런 뒤에 일정한 변위가 발생한 영역만을 사물로 간주하고 이 변위와 카메라와 카메라 사이의 거리, 카메라와 물체와의 거리등을 이용하여 3차원 깊이 정보를 계산한다. 하지만, 이 과정은 정합 과정이 높은 정밀도의 요구와 많은 연산 시간으로 인하여 고속 연산을 요구하는 실시간 구현에 장애가 되는 단계이다[2]. 따라서, 효율적인 3차원 모델링을 위해서는 정확하고 빠른 변위 측정이 필수적이다. 두 번째로, 카메라 캘리브레이션 (calibration) 작업이 필요하다. 이 작업을 통해서, 작업에 사용된 카메라들의 외부 파라메터를 비교해 봄으로써 동일 직선상에 평행한 광축을 갖도록 카메라가 배치되었음을 확인 할 수 있다. 세 번째로, 분할된 사물로부터 특징점을 추출하는 작업이 필요하다. 이렇게 추출된 특징점들은 마지막 네 번째 단계에서, 모델링과 렌더링 작업의 성능을 좌우하므로, 적절한 수의 특징점을 추출해 내도록 해야 한다.

본 논문은 일반적인 영상보다는 전경과 배경의 구분이 어느 정도 드러나는 영상에 대하여 좀더 빠르고 정확한 변위를 추출하기 위한 것으로 웨이블릿 변환에 기반한 2단계 하이브리드 변위 측정 방법을 제안한다. 그리고 이를 이용하여, 배경과 전경을 분할한 뒤, 전경으로부터 적절한 수의 특징점을 추출한다. 마지막으로 이 특징점들을 이용하여 z좌표를 계산하고 3차원으로 모델링하는 알고리즘을 제안한다.

본 논문에서는 전경과 배경의 영상차의 특성을 이용

하여 웨이블릿에 기반한 2단계 하이브리드 변위 측정 방법을 이용하는데 이것은 단순한 기존의 영역 기반 방식인 SSD, SAD, NCC와 유사한 히스토그램 분포를 가지면서도 기존의 영역 기반방법의 단점이었던 많은 계산 시간을 단축시킬 수가 있었고 2차 미분값을 이용한 화소기반 방법을 사용함으로써 폐색 영역에 덜 민감하고 보다 정확한 변위를 측정 할 수 있었다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제안하는 2단계 하이브리드 변위 측정 알고리즘을 설명하고, 본 논문에서 제안하는 방법과 기존의 방법을 비교 분석한다. 또한, 측정된 변위 벡터를 이용하여 변위 히스토그램을 생성하고 이를 이용하여 전경과 배경을 분리하는 방법을 설명한다. 3장에서는 2장에서 추출된 전경 영상에 대하여, 특징점을 추출하고, 이 특징점에 대하여 변위 벡터와 카메라 파라메터를 이용하여 z 좌표를 추정해 내는 과정을 설명한다. 4장에서는 기존의 3차원 모델링 알고리즘을 개선시킨 새로운 3차원 모델링 과정을 소개하고, 마지막으로 5장에서, 결론 및 향후 연구 방향을 소개한다.

2. 변위 측정 알고리즘

영상으로부터 3차원 정보를 추출하기 위해서는 두 영상의 대응점을 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, 특징 기반 방법, 화소 기반 방법 등이 있다.

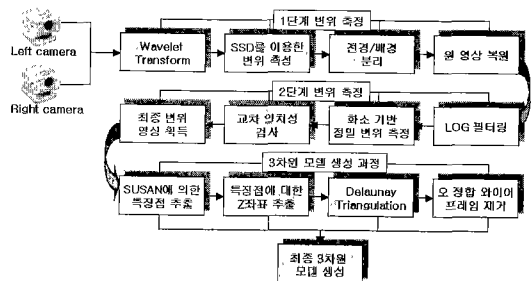


그림 1 본 논문의 3차원 모델 생성 과정

이중에서 영역 기반 방법은 기준 화면의 모든 화소에 대해 대응점을 구하기 때문에 조밀한 변위 정보를 얻을 수 있고 연산이 단순하다는 장점을 가지고 있다. 반면에, 잡음에 민감하고 상관 윈도우 안의 모든 점을 계산

해야 하므로 연산 시간이 많이 걸리고 폐색 영역이나 화소값의 차이가 작은 영역에서는 정확도가 떨어진다. 또한, 정합의 정확도가 창의 크기에 크게 좌우되므로 영상에 따라서 적절한 크기의 정합 창을 선택해야하는 단점이 있다. 이러한 단점을 보완하기 위해, 윈도우의 크기를 변화 시키며 변위를 구하는 방법[3]과, 영상을 피라미드 구조로 만들고 낮은 해상도에서 얻은 결과를 상위 해상도에서 이용하는 계층적 알고리즘 방법[4]이 제안되었다.

특징 기반 방법은 영상의 모든 화소에 대해 정합을 하지 않고 영상의 밝기 정보로부터 추출된 심볼의 특징을 이용한다[5],[12]. 이러한 심볼 특징에는 경계선, 끝점, 모서리 및 특별한 연산자에 의해 선택된 점들이 있으며 이러한 특징들은 영상 자체의 밝기보다는 밝기 변화에 관련됨으로써 영상의 물리적 변화를 잘 나타낸다. 또한, 영상에서 특징들의 개수가 일반적으로 화소 수보다 적으므로 전체적인 계산 처리 시간도 영역기반 방법에 비해 빠르다. 하지만, 특징들이 넓게 퍼져 있으므로, 구해지지 못한 나머지 화소들에 대해서 후처리로 보강을 시켜줘야 하는 단점이 있다[1].

화소 기반 방법은 영상내의 각각의 화소에 대해 밝기 값을 이용하여 대응점을 구하는 방식으로, 조밀한 변위를 측정할 수 있다[1],[11]. 이 방법의 초기 형태는 좌·우 영상의 화소 밝기 값의 차를 최소화하는 방법을 사용하였지만, 영상 안에 잡음이 존재한다거나 광학적 변화가 클 경우 잘못된 변위를 찾아낼 수 있다. 하지만, 이 문제는 평활화 과정을 통해 잡음을 제거함으로써 해결될 수 있고[6], 화소들의 유사성 측정 과정에서 폐색 영역을 발견해 낼 수 있는 장점이 있다.

2.1 웨이블릿 변환을 이용한 2단계 하이브리드 변위 측정 알고리즘

일반적으로 스테레오 영상에서 변위 벡터는 전체 탐색 범위 안에 골고루 분포되어 있지 않고 배경과 물체의 변위에 해당하는 값만을 갖지만, 기존의 방법은 전체 탐색 범위에 모두 정합 연산을 수행함으로써 많은 연산 시간과 함께 높은 오 정합 확률의 문제가 있었다[2]. 이러한 문제를 해결하기 위해 영상을 다단계 구조로 분리하고 낮은 해상도에서 구한 변위를 높은 단계의 해상도에 적용하는 계층적 방법(hierarchical block matching) 방법[3],[4],[7]을 사용함으로써, 계산 시간을 단축시키고 정확한 변위를 측정하고자 하였다. 하지만 이 방법은 낮은 해상도에서 한 화소의 변위가 높은 해상도에서 여러 화소에 똑같은 값으로 영향을 줌으로 오히려 경계 부분에서 정확한 변위를 예측하기가 힘들게 된다. 따라

서 본 논문에서는 웨이블릿 변환을 이용해 축소된 영상으로부터 영역기반 방법을 이용하여 대략적인 변위 벡터를 추출하고, 이를 이용하여 원 영상으로부터 전 단계에서 변위가 발생한 영역에 대해서만, 다시 화소의 밝기가 아닌 2차 미분값을 이용한 화소 기반 방법을 적용시킴으로써, 영역 기반 방법의 단점인 계산 시간을 단축시키고, 변형된 화소 기반 방법을 통해 폐색 영역에 덜 민감감하면서 정밀한 변위를 측정할 수 있는 하이브리드(hybrid) 방법을 제안한다.

본 논문에서는 실험을 위해 웹(<http://www-syntim.inria.fr/syntim/analyse/paires-eng.html>)으로부터 얻은 그림 2와 같은 320x120 크기의 영상 2개를 사용하였다.



(a) 1번 영상(left, right)



(b) 2번 영상(left, right)

그림 2 실험에 사용된 영상

2.2.1 웨이블릿 변환

웨이블릿 변환은 일반적으로 영상에 대하여 다 해상도의 표현, 뛰어난 에너지 압축 효과, 그리고 넓은 분야의 적용성등을 제공함으로[8], 영상 처리, 영상검색, 비디오 인덱싱과 같은 컴퓨터 비전 분야에 광범위하게 사용되어 오고 있다.

일반적으로 웨이블릿 변환으로 얻어지는 저주파 영역은 원 영상을 그대로 몇 배 작게 줄여 놓은 것이므로 모든 웨이블릿 부영역(subimage)중에서 가장 중요한 정보를 포함하고 있을 뿐만 아니라 잡음이 제거된 영상이다[8]. 따라서 본 논문에서는 Daubechies 웨이블릿 필터(tap-4)를 사용하여 영상을 1/4 크기로 줄여든 1단계로 분해하고, 분해된 저주파 영역으로부터 영역 기반 방법 중 SSD(Sum of Squared Distance)를 이용하여 대략적인 변위 발생 영역을 추정해 낸다.

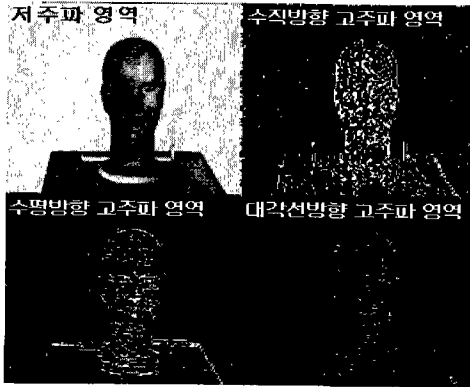


그림 3 1단계 웨이블릿 변환 결과

2.2.2 저주파 영역으로부터 1단계 변위 측정

우선, 웨이블릿 변환된 영상의 저주파 영역으로부터 영역 기반 방법 중 SSD를 이용하여 대략적으로 변위가 발생하는 영역을 추정해 낸다.

$$SSD(x, y, d) = \sum_w [L_{x,y} - R_{x+d,y}]^2 \quad (1)$$

이때, $R_{x,y}$, $L_{x,y}$ 는 각각 우 영상 및 좌 영상의 (x, y) 위치에서의 화소 밝기 값을 나타내고, w 는 대응하는 좌, 우 영상의 정합 창 크기를 나타낸다. SSD는 조명 변화나 부분 폐색 등과 같은 잡음 등에 민감하다는 단점을 가지지만, 식이 간단하여, 빠른 시간 안에 변위 발생 영역을 추정해 낼 수 있는 장점이 있다. 본 논문에서는, SSD를 위한 탐색 창 크기를 8×4 , 정합 창 크기를 4×4 로 정의하여 실험하였다. 여기서 탐색 창과 정합 창의 y 축 길이가 동일한 것은 스테레오 영상 자체가 y 축의 길이를 고정시키고, x 축의 위치만을 이동시켜 만들어진 것이므로 y 축의 이동이 없다는 가정 하에서 정해졌다. 저주파 영역으로부터 SSD를 이용하여 변위를 측정 한 뒤에 변위 영상에 대하여, 잡음 등에 의해 잘못된 찾아진 변위를 제거해 주기 위해 3×3 미디안 필터를 적용시킨다. 다음으로 웨이블릿 변환된 영역으로부터 구해진 변위를 원 영상의 변위로 변화 시켜 주는 작업이 필요하다. 본 논문에서 사용한 저주파 영역은 원 영상에 대해 $1/4$ 크기로 줄어든 영상이므로 그림 4와 같이, 웨이블릿 변환과 SSD 방법으로 줄어든 변위를 원 영상의 변위 벡터로 복원하는 작업이 필요하다. 그런 다음, 변위 영상안에 포함되어 있는 구멍등을 채워 주기 위해 보간법을 사용한다. 본 논문에서는 이 작업을 위해 쌍선형 보간법(bilinear interpolation)을 사용하였다. 이 방법은 계산량이 적고, 비교적 정확한 보간이 가능하므로 좋은 결과를 얻을 수 있다. 이렇게 하여 생성된 초기 변

위 영상은 그림 5와 같다.

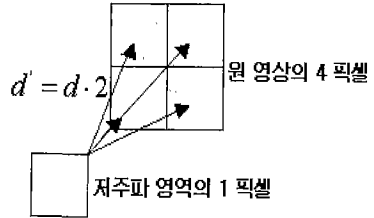


그림 4 변위 벡터 보정

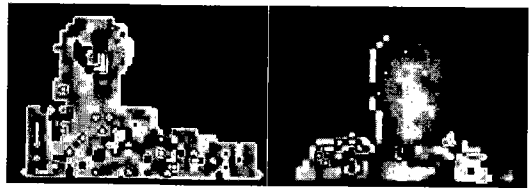


그림 5 보정된 1단계 변위 영상

다음으로, 측정된 변위를 이용하여 영상의 전경과 배경을 구분하는 작업이 필요하다. 이 과정은 변위 히스토그램에 의해 이루어지는데, 변위 히스토그램은 변위 영상에서 특정 변위값을 가지는 화소의 수를 히스토그램 형태로 나타낸 것이다. 변위 히스토그램을 이용한 전경과 배경 구분은, 일반적으로 스테레오 영상에서 변위 정보가 전체 탐색 영역에 고르게 분포되어 있지 않고, 배경과 전경의 변위에 해당하는 값만을 갖는다는 특성을 이용한 것이다. 즉, 스테레오 영상에서 물체는 일반적으로 카메라의 앞부분에 위치함으로 변위 벡터가 크게 나타나고, 배경의 경우는 뒷부분에 위치함으로 0에 가까운 값을 가지게 된다. 따라서, 이러한 특성을 이용하여, 계산된 변위 히스토그램으로부터 물체인 전경과 배경을 구분해 낼 수 있다.

변위 히스토그램 분석을 위해서, 본 논문의 1단계 보정된 영상의 변위 히스토그램과, 원 영상을 이용하는 영역 기반 방법의 대표적 3가지 방법인 SSD, SAD(Sum of Absolute Distance), NCC(Normalized Correlation Coefficient)를 비교하여 보았다(그림 6,7).

$$SAD(x, y, d) = \sum_w |L_{x,y} - R_{x+d,y}| \quad (2)$$

$$NCC = \frac{\sum_w [L_{x,y} - \bar{L}_{x,y}]^2}{\sqrt{\sum_w [L_{x,y} - \bar{L}_{x,y}]^2} \times \sqrt{\sum_w [R_{x,y} - \bar{R}_{x,y}]^2}} \quad (3)$$

$L_{x,y}$: 좌 영상의 (x, y) 위치에서의 화소 밝기 값

$R_{x+d,y}$: 좌 영상에서 d 만큼 이동된 위치에서의 우영상 화소 밝기 값

$\bar{R}_{x,y}$: 우 영상에서 정합창 내의 화소 평균값

$\bar{L}_{x,y}$: 좌 영상에서 정합창 내의 화소 평균값

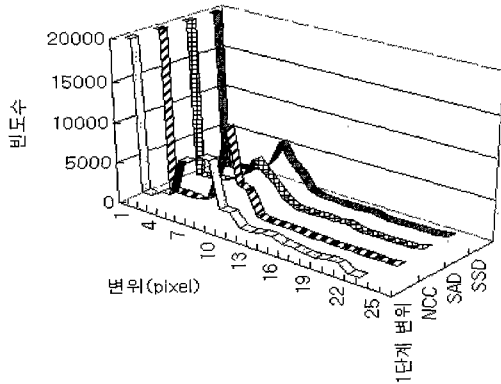


그림 6 1번 영상의 변위 히스토그램

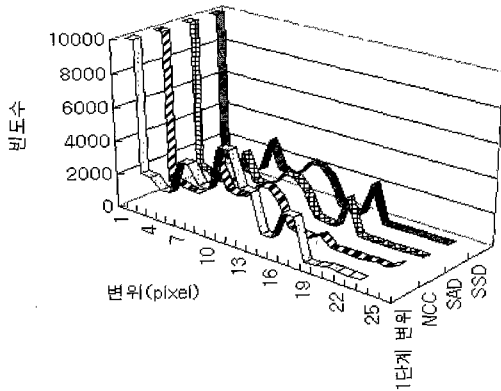


그림 7 2번 영상의 변위 히스토그램

그림 6과 그림 7에서 보는 것과 같이 줄어든 영상으로부터 구해진 변위 벡터를 원 영상으로 복원시켰더라도, 원 영상을 사용한 다른 방법인 SSD나, SAD, NCC와 유사한 변위 히스토그램 분포를 가지는 것을 알 수 있다. 따라서, 1단계 변위를 이용하여 대략적인 천경과 배경을 구분해 낼 수 있다.

2.2.3. 화소 기반 방법을 이용한 2단계 변위 추정

1단계 저주파 영역에서 추정된 초기 변위는 원 영상으로부터 변위 추정을 위한 후보 영역으로 사용된다. 즉, 원 영상을 이용하는 2번째 단계에서는, 영상의 전

영역에 대하여 변위를 추정하는 것이 아니라, 1단계에서 얻어진 초기 변위 발생 영역에 대해서만 화소 기반 방법을 적용시켜 계산 시간을 단축시킨다.

이때, 배경이 분리된 전경 영상에 대하여, LOG 필터를 적용시킨다. LOG 필터를 통해 영상의 화소 값은 밝기 값이 아닌 2차 미분값으로 변화된다. LOG 필터는 카메라로부터 영상을 획득하는 과정에서 생기는 잡음이나, 빛의 변화로 인한 영상의 명암 변화와 같은 현상도 제거해 주고 또한 영상의 특징들을 향상 시켜 주는 역할을 하므로 [9] 화소 기반 방법으로 인한 단점을 보완해 줄 수 있다.

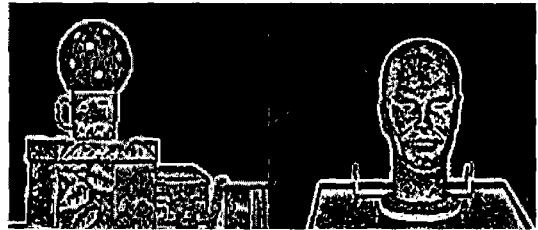


그림 8 LOG 필터 적용 결과

LOG 필터를 적용시킨 뒤에, 화소 기반 변위 추정 알고리즘을 적용한다. 일반적으로 화소 기반 변위 추정 방식은 영상 안에 포함된 잡음이나, 명도 차에 민감하게 반응하지만, 본 논문에서는 전처리로 LOG 필터를 적용함으로써 이러한 오류를 방지하였다.

본 논문에서는 식(4)를 이용하여 정밀한 변위를 추정하였다. 여기서 w 는 화소의 변위 추정 범위를 의미하는 것으로, 동일한 y 축 에피폴라(epipolar) 직선상에서 좌측 기준 영상으로부터 앞뒤로 $w/2$ 범위 안의 모든 점의 2차 미분값을 비교하여, 가장 유사한(차이값이 가장 작은) 화소를 찾아내어, 좌측 영상에서의 화소 좌표와 우측 영상의 찾아진 화소 좌표간의 거리를 변위로 정의한다.

$$d = |x_L - x'_R| \text{ if } \min_{x_R - \frac{w}{2} \leq x' \leq x_R + \frac{w}{2}} |I_L(x, y) - I_R(x', y)| \quad (4)$$

x_L, x_R : 좌측, 우측 영상에서 x 좌표

I_L : 좌측 영상의 화소 2차 미분값

I_R : 우측 영상의 화소 2차 미분값

그런 뒤에, 추정된 변위 벡터를 보정 시켜주는 작업을 시행한다. 이때도 마찬가지로, 쌍선형 보간법을 이용한다.

다음으로, 쌍방향으로 교차 일치성 검사(cross-consistency check)과정 [2],[3]이 필요하다. 이것은 좌측

영상으로부터 계산된 변위와, 우측 영상을 기준으로 계산된 대응점의 변위는 서로 일치해야 한다는 조건으로 이를 만족하지 못하는 변위를 갖는 화소는 정합에서 제외시킨다. 또한, 교차 일치성 검사 과정을 통해 제외된 화소를 중심으로 오른쪽에 있는 제외되지 않은 화소의 변위값과 비교하여 미리 정의된 임계값(본 논문의 실험에서는 5로 설정)를 초과할 경우 해당 영역을 폐색 영역으로 간주하고, 특징점 추출 영역에서 제외한다.

본 논문에서는, 2단계 하이브리드 방법의 성능을 비교하기 위해서 영역 기반 방법 중 가장 많이 사용되는 SSD와의 변위 결과를 비교하여 보았다. 그림에서 보는 것과 같이 일반적인 SSD의 경우, 영역안에 구멍등이 그대로 존재하고, 교차 일치성 검사를 하지 않아 변위값이 일정하게 분포하지 못하는 결과를 나타낸다.

본 논문에서 제안하는 알고리즘에 의해 최종적으로 구해진 변위 영상은 그림 11,12와 같다.

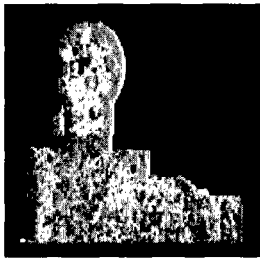


그림 9 SSD를 이용한 1번 변위 영상

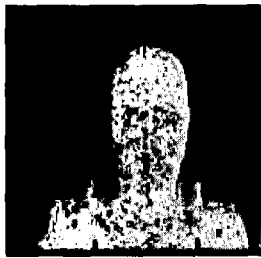


그림 10 SSD를 이용한 2번 변위 영상

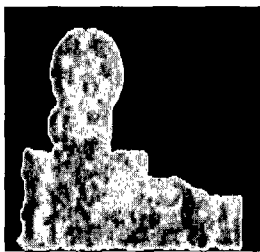


그림 11 최종적으로 구해진 1번 변위 영상

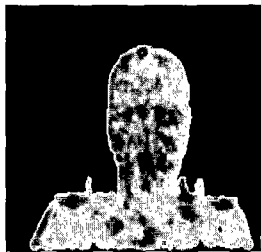


그림 12 최종적으로 구해진 2번 변위 영상

표 1은 좀더 정확한 성능 비교를 위하여 일반적인 영역 기반 방법에 의한 SAD, SSD, NCC와 영역 기반 방법의 가장 큰 단점인 블록의 크기를 고려하여, 블록 안의 경계선 정보에 따라 블록의 크기를 조절하는 경계선

에 의한 가변 블록 SAD와 SSD[13], 그리고 본 논문에서 제안하는 2단계 하이브리드 방법에 의한 변위 측정 성능을 각각 비교하였다. 이때, 2단계 하이브리드 방법의 계산 시간은 웨이블릿 변환 시간을 포함한 시간이다. 본 실험은 PentiumIII-350 PC 환경에서 이루어졌다.

표 1 2단계 하이브리드 방법의 성능 비교

방 법	1 번 영상		2 번 영상	
	계산 시간 (초)	최대 변위 (픽셀)	계산 시간 (초)	최대 변위 (픽셀)
2단계 하이브리드	2	20	2	22
경계선에 의한 가변 블록 SAD	98	29	96초	17
경계선에 의한 가변 블록 SSD	105초	36	95초	36
SAD	13	23	13	22
SSD	99	23	70	22
NCC	223	23	265	23

표에서 보는 것과 같이, 본 논문에서 제안하는 2단계 하이브리드 방법을 사용할 경우, 계산 시간을 상당 부분 단축시킬 수 있을 뿐만 아니라, 2차 미분값을 이용한 화소 기반 방법을 사용함으로써 폐색 영역 등에 덜 민감하고 보다 정확한 변위를 측정 할 수 있었다.

또한 다음으로는, 일반적으로 배경이 있는 복잡한 영상에 대한 실험을 해보기 위하여 보편적으로 스테레오 영상 실험에서 많이 사용되어지는 그림 13과 같은 수족관(aquarium) 영상을 가지고 변위를 측정하여 보았다. 그림 14는 중복 블록을 이용한 변위 측정(OBD-Overlapped Block Disparity)[14]의 결과이고 그림 15는 본 논문에서 제안한 방법에 의한 실험 결과이다. 결과 영상에서 보는 것과 같이 배경이 있는 일반적인 영상에서도 본 논문을 이용할 경우 정확하고 세밀한 변위를 얻을 수 있는 것을 볼 수 있다.



그림 13 수족관 실험 영상 3(left, right)

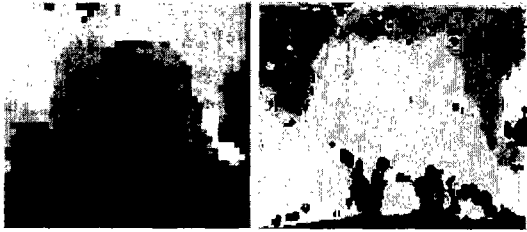


그림 14 OBD를 이용한 수축 그림 15 WT_SSD를 이용한 수축 변위 영상(188초) 수축 변위 영상(5초)

3. SUSAN 에지 검출기를 이용한 특징점 추출

특징점이란, 영상안의 물체가 가지고 있는 두드러지는 윤곽선이나 모서리에 해당하는 점으로, 이 특징점을 이용하여 움직임 추정, 3차원 인식, 모델링과 같은 분야에 적용시킬 수 있다. 일반적으로 영상 안에서 특징점을 추출하는 방법은 3가지로 분류 할 수 있다. 첫째로 윤곽선 맵을 이용하는 방법은, 먼저 영상으로부터 윤곽선을 뽑아낸 다음 이들을 체인 코드로써 표현하고, 지역적 극소 곡률을 갖는 점을 특징점으로 삼거나 혹은, 그 체인 코드에 폴리곤 추정을 수행하여 직선을 얻어낸 다음 직선들의 교차점을 특징점으로 삼는 방법이다. 두 번째는, 영상의 그레이 값을 직접 다루는 방법으로, 크게 밝기값 비교를 통한 edgeness와 방향성 미분을 기반으로 하는 cornerness로 분류 할 수 있다. 여기서, 방향성 미분을 이용하는 방법은 잡음에 민감한 단점을 가지고, 반대로, 밝기값 비교를 이용한 방법은 계산이 빠르고 잡음에 안정적인 장점을 가진다. 세 번째로, 모델 기반의 접근 방법은, 특징점을 뽑아 내기 위해 특정 모델을 사용하는 방법으로, 부화소 단위의 정확도를 갖는 모서리의 지역화가 가능하다.

본 논문에서는 3차원 모델링을 위하여 특징점을 추출하기 위해 두 번째 방법의 일종인 SUSAN 에지 검출기 [10]를 사용한다. SUSAN 에지 검출기는 윤곽선 맵이나 단순히 그레이 레벨의 밝기 값을 이용한 방법보다 잡음에 덜 민감하고, 비교적 정확한 특징 값을 추출해 낼 수 있다[3].



그림 16 특징점을 추출하기 위한 원형 마스크

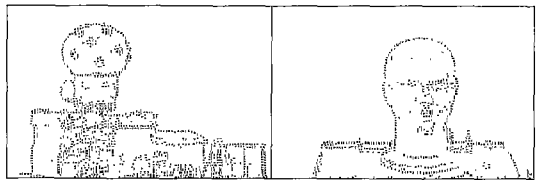
그림 16에서 볼 수 있듯이 SUSAN 방법은 원형 마스크를 사용한다. 이때, 배경이 제외된 전경 영상 안의 한 점 $p(x,y)$ 가 마스크 중심이 되고, 이 중심점의 최소값과 마스크 안의 다른 점 q 와의 차이 값은 식(5)를 통해 계산된다.

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

$$N(p) = \sum_w C(p, q)$$

식(5)을 통해 계산된 마스크 안의 $C(p, q)$ 의 합 $N(p)$ 가 마스크 크기의 반($w/2$)보다 작을 때 $P(x, y)$ 점은 특징점이 된다. 본 논문에서는 실험에 의해 $t=25$ 로 정의하고, 특징점이 물체의 지역적인 부분에 과다하게 추출되는 것을 방지하기 위해 원 논문에서 사용한 마스크의 크기 보다 큰 11×11 의 마스크를 사용하였다.

본 논문에 의해 추출된 특징점은 그림 17과 같다. 이때, 특징점이 되기 위한 조건을 $w/2$, $w/3$, $w/4$ 와 같이 조정함으로써 특징점의 수를 조절 할 수 있다.



(a) 1번 영상(특징점:996) (b) 2번 영상(특징점:579)

그림 17 SUSAN 방법에 의한 특징점 추출 결과

3.1 변위 벡터를 이용한 특징점 깊이 정보 추출

일반적으로 스테레오 영상은(2대의 카메라를 사용할 경우) 두 대의 카메라 초점거리(f)가 서로 같게 배치시킨다. 여기서 좌측, 우측 카메라의 각 영상 좌표계는 X_L, Y_L, Z_L 과 X_R, Y_R, Z_R 이라고 하고, 두 카메라의 광축을 각각 Z_L, Z_R 로 하여 서로 평행하도록 위치시킨다. 이때, 좌측 렌즈의 중심과 우측 렌즈의 중심 사이의 거리를 b , 이 두 중심 위치를 연결하는 선분을 기준선(base line)으로 정의한다[5]. 그림 18에서 보는 것과 같이 카메라를 배치함으로써, 대응점을 찾는 탐색 범위가 동일 선상으로 제한 될 수 있게 하는 에피폴라 구속 조건을 만족할 수 있으며, 이로부터 3차원 정보도 쉽게 얻을 수 있다.

좌우 카메라의 초점 거리를 f , 카메라 사이의 거리를 b , 3차원 공간상의 임의의 한 점 $E(x, y, z)$ 라 할 때, P 가 좌,우 카메라의 영상면에 투영된 점의 좌표를 각각

(x_L, y_L) , (x_R, y_R) 이라 한다면, x 축 좌표는 삼각형 비례 관계를 이용하여 다음과 같은 식으로 나타낼 수 있다.

$$x_L = \frac{fx}{z}, \quad x_R = \frac{f(x-b)}{z} \tag{6}$$

위 두 식으로부터 z 를 구하면,

$$z = \frac{bf}{x_L - x_R} = \frac{bf}{d} \tag{7}$$

로 나타낼 수 있고, 이 식을 이용하여 깊이 정보를 계산 할 수 있다. 본 논문에서는 2.2.3절에서 구해진 변위 벡터(d)중 특징점에 해당하는 변위 벡터에 대해서만 해당 좌표의 x, y 좌표와 d 를 이용하여 z 좌표를 추출한다. 이때, b 와 f 는 스테레오 영상 획득 과정에서 정해진 값 ($b : 40\text{mm}, 40\text{mm}$, $f : 10\text{mm}, 10\text{mm}$)을 사용하였다.

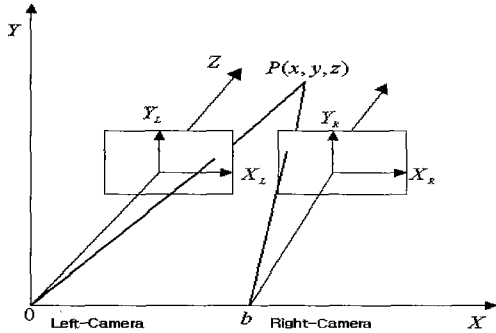


그림 18 스테레오 영상을 얻기 위한 카메라의 공간적 모형

4. 특징점과 Delaunay 삼각 측량(triangulation)을 이용한 3차원 모델링

추출된 특징점과 이 특징점들의 x, y, z 좌표 값들을 이용하여 실제적으로 3차원 모델링 하는 과정이 필요하다. 본 논문에서는 3차원 모델 생성을 위하여 Delaunay 삼각 측량법을 사용하였다.

이 방법은 우선, 3장에서 추출된 최상위 특징점을 초기 와이어 그리드 생성을 위한 시작 노드로 사용하고 3.1절에서 구해진 해당 특징점의 z 좌표를 이용하여 가장 근사한 점들을 잇는, 초기 3차원 와이어 프레임을 생성한다. 초기 와이어 프레임이 생성된 후, 추가적인 노드들은 삼각측량 정면 조각과 물체의 다른 특징점들과의 유클리디언 거리 계산에 의해 결정된다. 이때, 거리는 삼각 측량의 조각이 놓여 있는 방향의 수직 방향으로 계산되고, 물체의 특징점과 삼각 측량의 정면 조각과

의 거리가 미리 정의된 임계값을 초과한다면, 이 점을 새로운 노드로 선언하고, 와이어 프레임에 추가시킨다. 새로운 노드를 갖게 된 삼각 측량의 조각은 다시 새로운 노드에 꼭지점을 갖는 여러 개의 부 삼각형으로 나누어지게 된다.

하지만, 3차원 모델링을 위해 일반적으로 사용되는 이 방법은 몇 가지의 문제점을 가지고 있다. 그림 19에서 보는 것과 같이 대부분의 경우, 와이어 프레임이 정확히 생성되지만, 일부는 삼각 측량의 조각과 특징점들간의 거리 차만을 이용하므로, 오 정합된 와이어 프레임을 생성 할 수 있다.

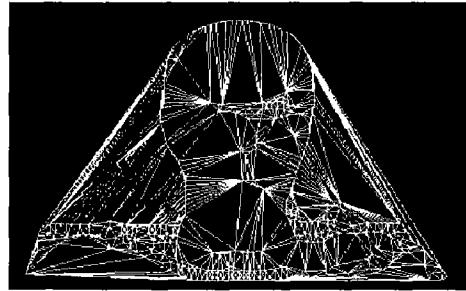


그림 19 오 정합된 Delaunay 삼각 측량법 결과

따라서, 정확한 텍스처 매핑을 위해서는 필수적으로 정확한 와이어 프레임이 생성되어야 하므로, 이러한 오 정합된 와이어 프레임을 제거해 주기 위해 본 논문에서는 전경/배경 분할 알고리즘을 제안한다.

4.1 오 정합 와이어 프레임 제거를 위한 전경/배경 분할 알고리즘

우선, 변위 히스토그램으로부터 얻어진 배경 영역을 버퍼에 저장 시켜 둔다. 그런 뒤에 Delaunay 삼각측량법을 이용하여 얻어진 각각의 삼각 측량 조각으로부터 중심점의 x, y 좌표 값을 구한다. 이렇게 구해진 각 삼각 측량 조각의 중심 좌표 값을 버퍼에 저장된 배경 영역과 매칭 시키는 작업을 실시한다. 그리고 만약, 중심 좌표가 배경 영역에 해당된다면, 이 삼각 측량 조각은 제거되거나 또는 새로운 부 삼각형으로 선언되지 못하게 된다.

$$tp_i = \begin{cases} \text{remove} & \text{if } C(w_x, w_y) \in \text{buffer}(x, y) \\ \text{branch} & \text{otherwise} \end{cases} \tag{8}$$

$C(w_x, w_y)$: 삼각 측량 조각의 중심 좌표

tp_i : i 번째 삼각 측량 조각

이 작업은 Delaunay 삼각측량법 과정에 몇 줄의 코드만을 추가하여 이루어짐으로써 전체적인 와이어 프레임

생성 시간에는 거의 영향을 주지 않는다. 이 과정을 통해 최종적으로 생성된 3차원 모델은 그림 20,21과 같다.

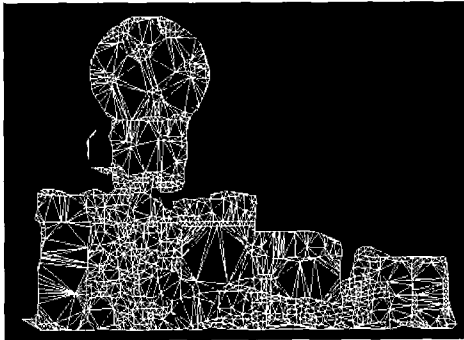


그림 20 3차원 모델링 결과(1번 영상)

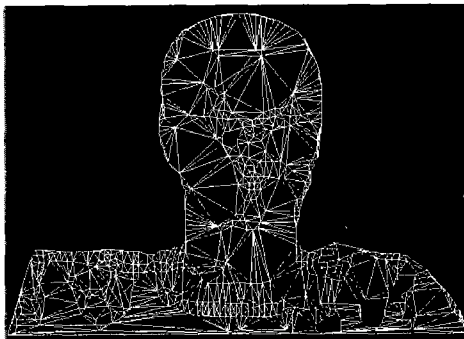


그림 21 3차원 모델링 결과(2번 영상)

5. 결론 및 향후 연구 방향

본 논문에서는 웨이블릿 변환을 이용하여 1/4 크기로 줄어든 저주파 영역으로부터 영역 기반 방법을 이용하여 대략적인 변위 영역을 구하고, 이를 이용하여, 전경과 배경을 분할 한 뒤, 다시 원 영상으로 복원하여 화소 기반 방법을 통해 조밀한 변위를 구하는 2단계 하이브리드 방법을 제안하였다. 또한, 분할된 전경 영역으로부터, 특징점들을 뽑아내고 변위 벡터와 카메라 파라미터를 이용하여 특징점들의 깊이 정보를 추정해 내어 3차원 모델링하는 과정을 제시하였다.

본 논문에서 제안하는 2단계 하이브리드 방법을 이용할 경우, 기존의 영역 기반 방법의 문제점이었던 계산 시간 문제를 상당 부분 단축시킬 수 있었고, LOG 필터를 통한 2차 미분값을 이용한 화소 기반 방법을 추가함으로써, 정밀한 변위를 구할 수 있었다. 또한 교차 일치성 검사를 통해 잘못된 변위를 제거하고, 폐색 영역들을

검사할 수 있었다. 아울러, 3차원 모델링 과정에서, 기존의 Delaunay 삼각측량법의 문제점이었던, 오 정합 문제를 전경/배경 분할 알고리즘을 추가함으로써 효과적으로 해결 할 수 있었다.

본 논문의 추가적인 연구 방향을 살펴보면, 우선 정확한 카메라 보정 작업이 이루어져야 한다. 현재 본 논문에서는 웹상에서 공개된 데이터를 사용하고 있으므로, 영상이 작아지거나, 변형될 경우 카메라 캘리브레이션 수치들도 변하게 된다. 따라서, 정확한 깊이 정보를 추정해 내기 위해서는 정확한 카메라 파라미터들을 얻을 수 있어야 한다. 또한 자연스러운 3차원 모델 생성을 위한 비교적 잡음에 덜 민감하면서도 빠르고 곡면처리도 할 수 있는 정확한 특징 값을 추출 할 수 있는 방법에 대한 연구가 계속 되어야 한다.

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

마지막으로, 동영상상을 이용한 3차원 모델링 작업이 추가되어야 한다. 이렇게 함으로써, 증감 현실이나, 3차원 화상회의 등과 같은 실제적인 응용 분야에 쉽게 적용할 수 있을 것이다.

참 고 문 헌

- [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, 김형곤, 차균현, "지 해상도 변위 히스토그램을 이용한 고성능 변위 정보 추출 알고리즘", 전자공학회 논문지 제 35권 제 3호, pp. 131-143, 1998.
- [3] Ebroul Izquierdo M, "Stereo Matching for Enhanced Telepresence in Three Dimensional Video Communications," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, No. 4, August 1997.
- [4] Christian Menard and Norbert Brandle, "Hierarchical Area-Based Stereo Algorithm for 3D Acquisition," International Workshop on Stereoscopic and Three Dimensional Imaging, September 1995.
- [5] 최우영, 박래홍, "Stereo vision 및 응용", 전자공학회지, 제21권, 제10호, pp. 1010-1020, October 1994.
- [6] B.K.P. Horn, Robot Vision, Cambridge Mass, MIT Press, 1986.
- [7] L.H. Quam, "Hierarchical warp stereo," Proceeding of DARPA Image Understanding Workshop, New

- Orleans, LA, pp. 149-155, 1984.
- [8] M.K. Mandal, T. Aboulnasr, S. Panchana than, "Image Indexing Using Moments and Wavelets," IEEE Transactions on Consumer Electronics, Vol. 42, No. 3, pp. 557-565, August 1996.
- [9] Takeo Kanade, Hiroshi Kano and Shigeru Kimura, "Development of a Video-Rate Stereo Machine," Proceeding of International Robotics and Systems Conference, 1995.
- [10] Stephen M. Smith, "SUSAN-A new approach to low level image processing," International Journal of Computer Vision 23(1), pp. 45-78, 1997.
- [11] Stan Birchfield and Carlo Tomasi, "A Pixel Dissimilarity Measure that is Insensitive to Image Sampling," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 20, No. 4, April 1998.
- [12] Angela Rojas, Antonio Calvo and Jose Murioz, "A dense disparity map of stereo images," Pattern Recognition letters, Vol. 18, pp. 385-393, 1997.
- [13] Jen-Luc Lotti and Gerard Giraudon, "Adaptive window algorithm for aerial image stereo," Pattern Recognition, Vol. 1, No. 3, October 1994.
- [14] Woontack Woo and Antonio Ortega, "Overlapped Block Disparity Compensation With Adaptive Windows for Stereo Image Coding," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 2, March 2000.



유 지 상

1985년 2월 서울대학교 전자공학과 졸업(공학사). 1987년 2월 서울대학교 대학원 전자공학과 졸업(공학석사). 1993년 5월 Purdue 대학교 전기공학과 졸업(Ph.D.). 1993년 9월 ~ 1994년 8월 현대전자산업(주) 산전연구소 선임연구원. 1994년 9월 ~ 1997년 8월 한림대학교 전자공학과 조교수. 1997년 9월 ~ 현재 광운대학교 전자공학과 조교수. 관심분야는 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리



노 윤 향

2000년 전북 대학교 컴퓨터 과학과 졸업(이학사). 2000년 ~ 현재 연세대학교 대학원 컴퓨터과학과 석사과정 재학중. 관심분야는 스테레오 영상, 3차원 모델링, 컴퓨터 비전, 변위 추출, 영상 처리, 패턴 인식



고 병 철

1998년 경기대학교 전자계산학과 졸업(이학사). 1998년 ~ 2000년 연세대학교 대학원 컴퓨터과학과 졸업(공학석사). 2000년 ~ 현재 연세대학교 대학원 컴퓨터과학과 박사과정 재학중. 관심분야는 영상검색, 비디오 인덱싱, 패턴 인식, 인

공지능

변 헤 관

정보과학회논문지 : 소프트웨어 및 응용
제 28 권 제 3 호 참조