

햅틱스 시스템용 3D 재구성을 위한 LoG 방법과 DoG 방법의 성능 분석

성미영[†], 김기권^{**}

요 약

본 연구의 목적은 “로봇의 시각”과 “로봇의 촉각”을 대체할 수 있는 스테레오 비전 기반 햅틱스 시스템에서 가장 적합하고 효과적인 3D 재구성(3D reconstruction) 방법을 제안하는 것이다. 삼차원 영상에 대하여 정확하게 촉감을 전달하려면 스테레오 영상에서 사물의 깊이 정보와 사물의 경계면에 대한 정확한 정보가 필요하다. 본 연구에서는 스테레오 영상에서 사물의 깊이 정보를 정확하게 얻기 위하여 전통적인 스테레오 정합 과정에 경계면 추출 방법인 LoG(Laplacian of Gaussian) 방법과 DoG(Difference of Gaussian) 방법을 혼합 적용하여 3D 영상을 재구성한 결과를 제시한다. 또한 어떤 방법이 햅틱 렌더링을 적용하는데 유용한 지 검증하기 위하여 연산 시간 및 오차 분석 실험을 수행한 결과, 본 연구처럼 비주얼 렌더링에 햅틱 렌더링을 추가하여 사용하는 경우에는 잡음 감소와 경계면 추출 성능이 더 우수한 DoG 방법이 더 효율적인 것으로 판단되었다. 본 논문에서 제안하는 스테레오 비전 기반 햅틱스 시스템을 위한 3D 재구성 방법은 이동형 정찰 로봇의 성능을 높이는 연구 등 여러 산업 분야와 군사 분야에 응용이 가능할 것이다.

Comparison of LoG and DoG for 3D reconstruction in haptic systems

MeeYoung Sung[†], KIM KiKwon^{**}

ABSTRACT

The objective of this study is to propose an efficient 3D reconstruction method for developing a stereo-vision-based haptics system which can replace "robotic eyes" and "robotic touch." The haptic rendering for 3D images requires to capture depth information and edge information of stereo images. This paper proposes the 3D reconstruction methods using LoG(Laplacian of Gaussian) algorithm and DoG(Difference of Gaussian) algorithm for edge detection in addition to the basic 3D depth extraction method for better haptic rendering. Also, some experiments are performed for evaluating the CPU time and the error rates of those methods. The experimental results lead us to conclude that the DoG method is more efficient for haptic rendering. This paper may contribute to investigate the effective methods for 3D image reconstruction such as in improving the performance of mobile patrol robots.

Key words: stereo vision(스테레오 비전), 3D reconstruction(3D 재구성), haptic rendering(햅틱 렌더링), LoG(Laplacian of Gaussian), DoG(Difference of Gaussian)

※ 교신저자(Corresponding Author): 성미영, 주소: 인천시 연수구 송도동 12-1 인천대학교 정보기술대학 427호 (406-772), 전화: 032)835-8496, FAX: 032)835-0780, E-mail: mysung@incheon.ac.kr
접수일: 2012년 1월 11일, 수정일: 2012년 3월 14일
완료일: 2012년 4월 5일

[†] 중신회원, 인천대학교 컴퓨터공학부 교수
(E-mail: mysung@incheon.ac.kr)

^{**} 정회원, 인천대학교 컴퓨터공학과 석사
(E-mail: kwon1983@incheon.ac.kr)

※ 본 연구는 인천대학교 2010년 자체연구비 지원에 의해 수행되었음.

1. 서 론

그 동안 인간의 오감을 디지털화 하는 기술에 대한 연구가 활발히 진행되어 왔다. 오감 중에서도 시각과 청각 분야에 많은 발전이 있었으며, 특히 스테레오 비전 영상처리 기술을 이용하여 인간의 두 눈과 같은 역할을 해주는 시각 센서 개발 등의 흥미로운 연구가 활발하게 수행되고 있다. 로봇의 시각 센서는 주위 환경에 융통성 있게 대처할 수 있는 정보를 제공해주는 로봇 시각의 중추적인 역할을 담당한다.

시각 센서는 주로 카메라를 이용하여 실시간으로 영상을 취득하고, 취득된 방대한 양의 영상 데이터로부터 물체의 크기, 위치, 색깔 등의 정보를 실시간으로 분석하여 물체를 인지해야 한다. 최근에는 이러한 시각 센서를 두 개 이상 사용하여 물체의 삼차원 영상 정보를 실시간으로 추출하는 연구가 학계와 연구소를 중심으로 진행 중이다.

현재 개발된 수준의 시각 센서 기술로는 카메라 및 적외선, 초음파 등의 기본적인 센서에 의해서만 시각적 관찰을 할 수 있어 사용자에게 많은 정보를 전달 할 수 없다. 그러나 사용자가 원격 사물에 대한 정보를 힘 센서나 햅틱 장치를 통해 촉각으로 느낄 수 있게 되면 카메라와 일반 센서로 받는 정보보다 많은 정보를 얻을 수 있으므로 기존 시각 센서에 비해 더 정확하고 사실적으로 원격 환경의 관찰을 할 수 있게 해 줄 수 있다.

최근 물체의 역학을 느끼게 하는 햅틱스(haptics, 촉각) 기술에 대한 관심이 높아지고 있다. 햅틱스는 인간과 컴퓨터가 정보를 주고받는 과정에서 시청각 정보에 국한되지 않고 촉각정보도 이용하게 하는 컴퓨터 촉각기술이다. 만일 컴퓨터 촉각을 로봇에 결합시킬 수 있다면 이동 중인 로봇으로 하여금 극지나 오지를 햅틱 장치를 이용하여 탐사하게 하여 위험을 제거할 뿐 아니라 원격 사물의 강도나 촉감까지 제공하는 새로운 첨단 로봇 기술을 제시할 수 있다.

스테레오 비전 기술을 이용하는 ‘로봇의 시각’과 이제 막 발전이 되기 시작한 햅틱스 기술을 이용하는 ‘로봇의 촉각’이 결합된다면 기존의 로봇 기술을 한 단계 발전시킬 수 있는 계기가 될 것이다.

로봇의 촉각을 잘 구현하려면 로봇의 눈을 대체하는 영상에서 사물의 깊이와 경계 정보를 보다 정확히 얻어 3D 재구성(3D reconstruction)을 해 주어야 한

다. 본 논문에서는 사물의 깊이와 경계를 좀 더 정확히 얻을 수 있도록 스테레오 정합에 경계면 추출이 가능한 LoG(Laplacian of Gaussian)[1,2] 알고리즘과 DoG(Difference of Gaussian)[3] 알고리즘을 혼합 사용하는 방법을 제시한다. 또한 연산시간, 에러율분석을 통하여 어떤 알고리즘이 햅틱 렌더링을 적용하는데 더 효과적이고 유용한지 실험한 내용을 보여준다.

다음 장 관련 기술에서는 스테레오 비전(stereo vision) 기술과 햅틱스 기술에 대하여 간단히 알아본다. 3장에는 본 연구에서 사용한 실제 환경에서 깊이 정보를 추출하고 3D 재구성을 하는 기본적인 방법을 설명하고, 4장에서는 햅틱 렌더링에 적합한 3D 재구성 방법을 실제로 구현된 시스템에서 실험한 결과와 그 분석 내용을 제시한다. 끝으로 5장에서 결론을 맺는다.

2. 관련 기술

이 연구의 궁극적인 목표는 스테레오 비전 기술을 이용하여 이동형 로봇의 시각을 구현하고 햅틱스 기술을 이용하여 이동형 로봇의 촉각을 구현하여, 로봇을 통해 원격의 사물을 보다 효과적으로 검침(probing) 하고자 하는 것이다. 본 장에 스테레오 비전 기술과 햅틱스 기술에 대하여 간단히 소개한다.

2.1 스테레오 비전

스테레오 비전 시스템은 이제 막 연구가 활성화되고 있는 영역이다. 차량 자동 주행 연구를 위하여 차선을 인식하기 위한 방법[4], 감시 카메라에서 사람의 움직임을 인식하기 위한 방법[5], 모바일 로봇에서 장애물 인식 등[6]의 여러 분야에서 연구가 진행되고 있다.

입체 영상 인식은 두 눈에 보이는 영상의 일부 영역이 겹쳐 있다는 사실로부터 출발한다. 생물학적인 시각은 이 겹쳐진 두 영역에 의해서 깊이를 인지하게 되는데, 입체 영상 인식에서도 이런 양안 영상의 차이를 이용하여 삼차원 구조를 인식한다. 양안의 시각(disparity)는 두 눈이 서로 다른 위치에서 관찰한 물체의 위치 차이를 말한다. 스테레오 비전 시스템은 이런 양안의 불균형을 이용하여 깊이(depth) 정보를 추출함으로써 이루어진다[7].

스테레오 비전의 핵심은 왼쪽 카메라와 오른쪽 카메라 사이의 관계를 이용하는 것이고, 이 방식에서는 스테레오 카메라의 보정(calibration) 과정이 중요하다. 이 방식의 단점은 보정 상태에서 조금이라도 벗어나면 스테레오 처리 과정에서 손상된 데이터가 발생한다는 점이지만 한번 보정한 카메라는 카메라의 위치를 조정하지 않는 한 새로운 보정 작업 없이 기존의 데이터를 이용할 수 있다는 장점이 있다.

스테레오 정합(stereo correlation)은 왜곡 없이 보정된 이미지의 왼쪽과 오른쪽 카메라의 특정 지점을 한 이미지와 다른 이미지 사이에서 대응시키는 것이다. 카메라로부터 거리를 측정하기 위해 시차를 이용하여 왼쪽 이미지의 점의 위치를 오른쪽으로 대응시켜 얻어진다. 이것은 이미지를 겹쳐서 왼쪽 영상에 있는 점이 오른쪽에 있는 점과 일치하는 지를 찾는 것이다. 실제 상황에서의 스테레오 보정과 정합에 대하여는 3장에서 좀 더 자세히 알아본다.

2.2 햅틱스

햅틱스는 ‘만지는’이라는 뜻의 그리스어 형용사 ‘haptesthai’에서 온 말로 사람의 오감 중 촉각이라는 채널을 통하여 컴퓨터와 대화할 수 있는 기술을 연구하는 학문 분야이다. 즉, 가상환경 혹은 조종기나 로봇 등을 이용하여 물체를 만지는 환경에서 사용자에게 촉감과 힘, 운동감 정보를 전달하는 방법과 관련된 연구의 총칭이다[8]. 가상환경 또는 현실환경에서 사용자가 물체를 만지고, 옮기고 하는 행동을 가능하게 하는 하드웨어, 소프트웨어 및 심리학적 연구가 모두 햅틱스 분야에 포함된다. 햅틱스 기술에서 사용자에게 전달되는 정보는 크게 힘 피드백(force feedback)과 촉각 피드백(tactile feedback)으로 구분된다. 힘 피드백은 기계 장치를 통해 사용자로 하여금 힘과 운동감을 느낄 수 있게 하고, 촉각 피드백은 물체 표면의 재질을 사실적으로 사용자에게 전달하는 것이다[9].

햅틱 렌더링(haptic rendering)은 가상의 물체들과 상호작용을 하면서 발생하는 벡터를 계산한 후 장치에 전달하여 사용자로 하여금 촉감을 느낄 수 있게 한다[10-13]. 햅틱 렌더링은 경계 영역을 이용하여 힘을 계산한다. 햅틱 렌더링에서 사용하는 경계 영역의 의미는 그래픽 렌더링에서 사용하는 경계 영역과는 차이가 있다. 그래픽 렌더링에서는 각 물체에

대한 경계 영역을 설정하고, 경계 구역간의 교차를 완벽히 회피하여 충돌을 방지하는 것을 목표로 한다. 하지만 햅틱 렌더링은 경계 구역간의 교차 부분을 이용하여 힘을 계산한다.

가상공간 내에 존재하는 자유 구역에서는 햅틱 장치의 탐침 위치(haptic interface position; HIP)는 어떠한 위치에서도 힘을 느낄 수가 없다. 하지만 햅틱 장치의 탐침 위치가 가상 공간상에 존재하는 물체의 경계 영역의 벽에 닿았을 때, 더 이상 자유로이 움직이지 못하고 반작용으로 나타나는 힘을 느끼게 된다.

힘의 크기를 계산하는 것은 경계 영역 표면에 존재하는 이상적인 햅틱 장치의 위치(ideal haptic interface position; IHIP)와 경계 구역을 뚫고 지나간 햅틱 장치 탐침의 실제 위치를 사용하여 계산된다.

경계 영역에 대하여 접촉하였을 때 느끼게 되는 힘의 계산은 (식 1)와 같다.

$$F = \begin{cases} 0 & x > x_w \\ k(x_w - x) & x \leq x_w \end{cases} \quad (1)$$

k 는 물체의 단단한 정도를 나타내고, x 는 촉각 장치의 탐침의 위치를 나타내며, x_w 는 경계 영역의 벽 위치를 나타낸다. F 는 물체에 접촉했을 때 실제적으로 느끼는 힘의 값을 나타낸다. 만약 가상공간에 존재하는 물체가 마찰이 없다고 가정한다면 나타나게 되는 힘의 방향은 법선 벡터(normal vector)의 방향으로 나타난다. 또한 물체가 강체(rigid body)일 경우에는 k 의 값을 가능한 가장 높은 수로 설정한다.

3. 실제 환경에서의 깊이 정보 추출과 3D 재구성 방법

3.1 스테레오 거리측정

이상적인 스테레오 시스템은 두 카메라의 광학축이 완벽하게 평행하게 있고 두 이미지 평면이 동일 평면상에 존재하며 렌즈 왜곡이 없다는 가정을 하게 된다. 그림 2에서 장면상의 점 P 까지의 거리는 3차원 좌표를 결정할 수 있는 정보를 제공해 준다. 여기서 거리 Z 는 (식 2)에 의해서 얻어낼 수 있다.

$$Z = \frac{Bf}{d} \quad (2)$$

이 식에서 f 는 초점거리(focal length), B 는 기준거리(baseline), 그리고 d 는 식 $d = x^l - x^r$ 를 통하여 얻

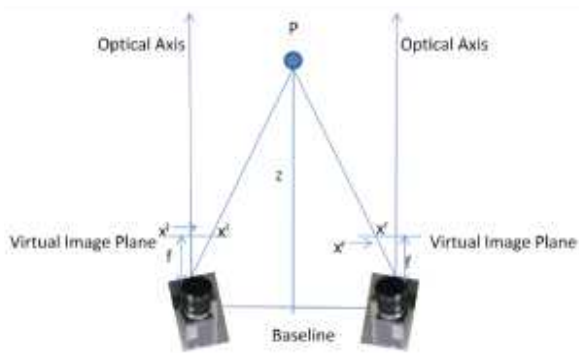


그림 2. 장면상의 P점을 카메라로부터의 거리인 Z값을 알아내기 위한 스테레오 거리측정의 단순 기하학

어낸 시차(disparity) 값이다.

3.2 카메라 보정

보정 과정은 두 카메라 사이의 상대적인 방향과 각도, 초점거리와 같은 카메라의 고유 파라미터와 외부 파라미터 값을 이용하게 된다. 특정 개체를 보정 대상으로 사용하게 되며 대상 개체의 3차원 좌표계를 카메라의 참조로 사용하게 된다[14].

이 논문에서는 Heikkila와 Silven이 사용한 순서에 따라 보정 작업을 진행하였고[15], 보정 알고리즘은 Bouguet 알고리즘[16]을 사용하였다. 보정을 위하여 9×7 체스보드 이미지를 양쪽 카메라 모두에서 보이도록 위치시키고, 자체 제작한 자동 코너 추출 프로그램으로 카메라의 왜곡(distortion) 파라미터를 추출하였다.

그림 3의 오른쪽 이미지는 (식 3)를 이용하여 왼쪽의 체스보드 이미지를 보정한 결과이다.

$$\begin{aligned} x_{corrected} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_{corrected} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{aligned} \quad (3)$$

여기서 (x, y) 는 원본 이미지의 픽셀 좌표이며 $(x_{corrected}, y_{corrected})$ 는 수정된 이미지 픽셀의 수정된 좌표이다. 변수 r 은 이미지 중심으로부터의 거리를

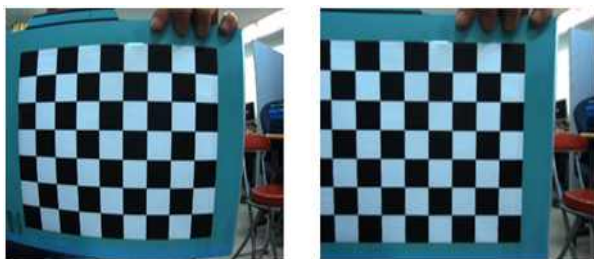


그림 3. 보정 전(좌) 보정 후(우)

말하며, 상수 k_i 는 $r=0$ 에서의 테일러급수(Taylor series)에서 얻을 수 있다. 카메라 보정의 모든 과정은 k_i 값을 풀어내는 것에 있다.

3.3 스테레오 정합

스테레오 정합은 왜곡 없이 보정된 이미지의 왼쪽과 오른쪽 카메라의 특정 지점을 한 이미지와 다른 이미지 사이에서 대응시키는 것이다. 카메라로부터 거리를 측정하기 위해 시차가 필요하고 이를 위해 왼쪽 이미지의 점의 위치를 오른쪽의 위치로 대응시켜야 한다. 이것은 이미지를 겹쳐서 왼쪽 영상에 있는 점이 오른쪽에 있는 점과 대응되는 지를 찾는 것이다.

스테레오 정합을 계산하고 대응점을 찾는 방법은 윈도우 내부의 SAD(sum of absolute distances)를 계산하는 블록 매칭 기법을 사용한다. SAD는 주어진 윈도우 내부에서 두 영상의 픽셀 값들의 절대 값을 모두 합한 값을 의미한다. 블록 매칭 방법은 세 가지 과정으로 이루어지는데 첫째는, 영상의 밝기를 정규화하고 질감을 향상 시키는 전처리 과정이다. 둘째는, 수평 에피폴라 직선(epipolar line)을 따라 SAD 윈도우를 이동시키며 대응점을 찾는 과정이다. 그리고 마지막은 잘못된 대응점을 제거하는 후처리 과정이다.

첫 번째 정규화 과정에서 왼쪽과 오른쪽 이미지는 정규화 되고 밝기 차이 영향을 줄일 수 있다. 서로 다른 여러 크기의 윈도우에 위치한 픽셀들은 (식 4)에 의하여 대체된다.

$$\min[\max(I_c - \bar{I}, I_n), I_n] \quad (4)$$

이 식에서 \bar{I} 는 윈도우 내부의 평균값을 의미하고 I_n 은 상한 값을 설정하는 변수이며, I_c 는 윈도우 중앙에 위치한 픽셀을 의미한다.

두 번째는 대응점을 찾는 과정이다. 보정 후에는 이론상으로 왼쪽 영상에서 찾은 특징 점들에 대하여 오른쪽 영상에서 동일한 행에서 나타날 가능성이 높다. SAD 윈도우가 왼쪽 이미지에 설정되고 점수를 계산한다. 그리고 알고리즘은 오른쪽 이미지에서 왼쪽과 같은 위치에서부터 최대시차에 가까워질 때까지 x축을 따라서 왼쪽으로 이동한다. 시차는 초기 픽셀에 대한 대응점을 찾은 픽셀까지의 픽셀의 개수이다. SAD 값을 계산하는 (식 5)는 아래와 같다.

$$SAD(r, c) = \sum_{y=-w}^{y=w} \sum_{y=-w}^{y=w} |R(y+r, x+c+d) - L(y+r, x+c)| \quad (5)$$

방정식 (r, c)는 오른쪽 이미지의 대응점 검색을 할 좌표이고, d는 오른쪽 이미지의 초기 이미지에 대한 시차 값이다. 그리고 w는 각 좌표에 위치할 윈도우의 사이즈이다. 이 방정식에서 점수는 각 좌표에서 픽셀 주위를 둘러싸는 이웃하고 있는 픽셀의 강도를 기준으로 계산된다. 오른쪽 검색 범위에서 이미지의 점수 값이 가장 작은 값이 가장 왼쪽 이미지에 대한 가장 적합한 매칭 점이 되는 것이다. 이 시차 값을 이용하여 깊이 값을 계산할 수 있는데, 카메라에서 가까운 점은 시차 값이 작게 되고 카메라에서 멀수록 시차 값이 커지게 되는 것을 이용하는 것이다.

대응 결과로 각각의 픽셀의 왼쪽 이미지에 대한 오른쪽 이미지의 시차 값을 이미지화 시킨 것을 시차 지도(disparity map)이라 한다. 그림 4는 시차지도의 예를 보여주고 있다. 어둡게 나타나는 부분이 카메라와 가까운 부분이고 밝게 나타나는 부분은 카메라에서 먼 부분이다. 검은색 부분은 두 이미지 사이에 매칭 점이 없다는 뜻이다.

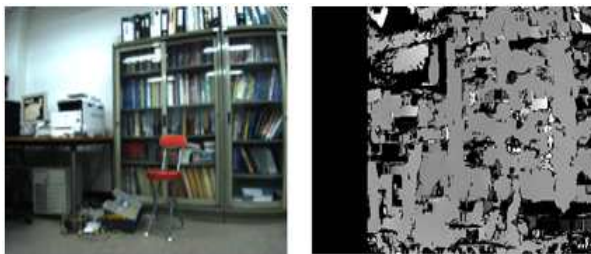


그림 4. RGB이미지(좌) 시차지도(우)

3.4 3D 재구성(3D Reconstruction)

2D 좌표를 3D 좌표로 재구성하기 위해서는 각 픽셀의 시차 값을 이용하는데, 2D 좌표의 각 픽셀에 대하여 상응하는 3D 좌표 (X, Y, Z)를 얻을 수 있다. Z 좌표 값은 기준선 B, 초점거리 f, 시차 d를 이용하는 (식 6)과 같은 거리 방정식을 이용한다.

$$Z = \frac{Bf}{d(sensor_element_size)} \quad (6)$$

시차를 거리로 변환하기 위해서는 얻어낸 시차 값에 센서의 개별 요소 크기를 곱한 값을 이용한다. X와 Y 값은 얻어낸 Z 값을 거리에 따른 픽셀 크기를

계산하여 얻어낼 수 있다. 먼저 화각(field of view: FOV) 값을 (식 7)에 따라 계산한다.

$$FOV = 2 * \arctan\left(\frac{size_of_sensor}{2f}\right) \quad (7)$$

센서의 크기는 가로나 세로 방향의 FOV를 계산하기 위한 카메라 센서의 전체 크기를 말한다. FOV 값이 계산되면 지정된 거리에서의 픽셀의 크기는 (식 8)를 이용하여 계산된다.

$$size_of_pixel = 2 * \sin\left(\frac{FOV}{2}\right) * Z \quad (8)$$

세로 방향 FOV와 가로 방향 FOV를 이용함으로써 이미지가 포함하는 가로와 세로 방향 거리를 측정할 수 있다. 또한 이를 이용하여 가로와 세로 방향의 총 픽셀 수로 나뉘 카메라로부터의 z값을 얻어낼 수 있다. 마지막으로 왼쪽 카메라의 중심 좌표인 c_x 와 c_y 를 투영의 중심좌표 (c_x, c_y)로 사용함으로써 실 세계의 X와 Y 좌표를 픽셀 좌표 x와 y로 투영시킬 수 있다. 다음 수식은 어떻게 x와 y가 계산되는지 보여준다.

$$X = \frac{2 * \sin\left(\frac{FOV_{horizontal}}{2}\right)}{image_width} * (x - c_x) \quad (9)$$

$$Y = \frac{2 * \sin\left(\frac{FOV_{vertical}}{2}\right)}{image_width} * (x - c_y)$$

이 방정식을 사용하면 픽셀 좌표와 해당 시차를 쉽게 점운(point cloud)으로 재구성 할 수 있다. 그림 5는 스테레오 카메라로부터 재구성된 3D 점운을 보여주고 있다.



그림 5. 2D 이미지(좌) 3D 재구성(우)

3.5 경계면 추출 알고리즘

햅틱스 시스템에서는 최종적으로 햅틱 렌더링으로 사용자에게 촉감을 부여해야 한다. 물체의 경계면에 대한 정보가 정확할수록 햅틱 렌더링을 정확하게 할 수 있다. 이전 절에서 설명한 3D 재구성 과정을

거쳐 생성된 3D 점운에 대하여 햅틱 렌더링을 바로 적용하게 되면 햅틱 렌더링을 정확하게 하기 어렵다. 그림 5에서 보듯이 일반적인 스테레오 처리 방법으로는 노이즈가 생겨 경계면에 대한 정보를 정확히 얻어내기 힘들기 때문이다.

수십 년 동안 수많은 경계면 추출 알고리즘들이 연구되어 왔다. 1차 도함수(first-derivative)를 이용하는 Roberts 방법, Kirsch Compass 방법, Prewitt 방법, Sobel 방법 등이 있다. 기울기의 변화량을 추적하여 좀 더 정확하게 경계면을 추출하는 2차 도함수(second-derivative) 방법에는 라플라시안, Laplacian of Gaussian, Difference of Gaussian, 그리고 Gaussian 방법들이 있다.

이들 방법의 장단점을 아래 표에 요약하였다.

Laplacian 마스크는 주변 밝기와 차이 값을 이용하여 경계면을 추출하기 때문에, 노이즈에 약하고 경계면보다는 영상 내의 가는 선이나 고립점에 강하게 반응한다. 따라서 Laplacian 방법은 좋은 경계면 추출 알고리즘이 아닐 수 있다. 이러한 단점을 극복하기 위해 영상의 노이즈를 Gaussian 둔화(blurring) 시킨 후 적용함으로써 더 강인하게 경계면을 추출하는 방법이 있다. 바로 LoG와 DoG가 그것이다.

LoG의 장점은 단순히 Laplacian 마스크를 이용하여 경계면을 추출하는 것보다 Gaussian 필터를 이용하여 경계면을 추출하는 것보다 Gaussian 필터를 이용하여 경계면을 추출하는 것보다 좀 더 굵은 경계면을 추출할 수 있다는 점이다.

DoG는 Gaussian 필터의 분산 값을 약간 차이 나

게 주어 Gaussian 영상의 차이를 이용하여 경계면을 추출하는 방법이다. 분산 값이 $\sigma_1/\sigma_2=1.6$ 정도 비율일 때 DoG는 LoG와 가장 유사한 형태를 보여준다. DOG는 LOG에 비해 연산 속도가 빠른 장점이 있다.

위의 방법들 중 경계면 추출 성능이 가장 우수한 방법은 Canny 방법이다. 그러나 Canny 방법은 계산이 복잡하고 연산시간이 많이 소모되므로 햅틱스 시스템에는 적합하지 않다. 왜냐하면 햅틱스 시스템에서는 1 KHz 이상의 햅틱 렌더링과 30 Hz 이상의 비주얼 렌더링이 동시에 수행되어야 하는 상황에서 동영상 프레임의 경계면들을 실시간으로 분석해야 하기 때문이다.

본 연구에서는 경계면 검출 성능이 우수하고 비교적 연산 부담이 적은 LoG 방법과 DoG 방법을 실제 햅틱스 시스템에 적용해 보고 어느 방법이 햅틱 렌더링에 더 적합한지 알아보려고 한다. 이를 위하여 아래 두 알고리즘을 블록 매칭 전 단계의 좌측과 우측 영상에 대하여 각각 적용하는 실험을 하였다.

3.5.1 LoG

Laplacian 연산자가 노이즈에 민감한 점을 Gaussian 필터를 이용하여 개선한 LoG 방법은 아래 (식 10)과 같은 형태의 마스크를 이용함으로써 수행될 수 있다.

$$LoG(x,y) = -\frac{1}{\pi\sigma} \left[1 - \frac{x^2+y^2}{2\sigma^2} \right] e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (10)$$

여기서 σ 는 편차 값으로 값이 클수록 함수의 폭이

표 1. 경계 검출 방법의 장점과 단점[17]

	연산 방법	장 점	단 점
1차 도함수	Loberts Kirsch Prewitt Sobel	<ul style="list-style-type: none"> • 단순 • 에지와 에지방향 추출 	<ul style="list-style-type: none"> • 노이즈에 민감 • 부정확
2차 도함수	Laplacian	<ul style="list-style-type: none"> • 에지와 에지방향 추출 • 모든 방향에 같은 특성 	<ul style="list-style-type: none"> • 일부 에지에 반응 • 노이즈에 민감
	Laplacian of Gaussian (Marr-Hildreth[1])	<ul style="list-style-type: none"> • 정확한 에지 검출 • 픽셀 주변의 넓은 범위 조사 	<ul style="list-style-type: none"> • 커브, 코너, 그레이 레벨 변화 큰 경우 부정확 • 라플라시안 필터 한계로 에지방향 탐지 어려움
	Difference of Gaussian	<ul style="list-style-type: none"> • 정확한 에지 검출 • 빠른 연산 • 근사 LoG 	<ul style="list-style-type: none"> • 전체적 대비 감소
	Gaussian (Canny[18], Shen-Castan)	<ul style="list-style-type: none"> • 에러율 검사에 통계 사용 • SNR(signal to noise) 향상 • 노이즈 상황에서 좋은 성능 	<ul style="list-style-type: none"> • 계산 복잡 • 연산 시간 소모

넓어지고 작을수록 함수의 폭이 커진다. 아래 그림 6은 스테레오 카메라에 LoG를 적용한 영상이고, 그림 7은 그 결과 영상으로 블록 매칭과 3D 재구성 과정을 거친 영상이다.

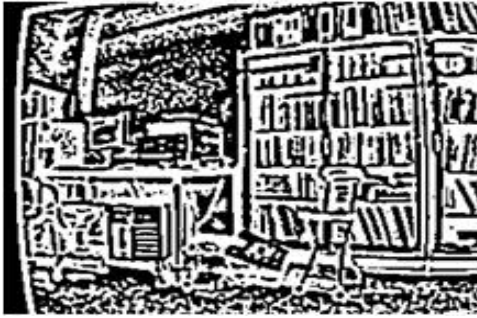


그림 5. LoG를 적용한 영상

것을 볼 수 있다. 햅틱 렌더링 또한 그래픽 렌더링과 유사한 방법으로 렌더링 되기 때문에 햅틱 렌더링을 적용하였을 경우 확실한 경계면을 느낄 수 있을 것으로 판단된다.



그림 7. DoG를 적용한 영상



그림 6. LoG로 3D 재구성한 영상



그림 8. DoG로 3D 재구성한 영상

3.5.2 DoG

Gaussian 영상 사이의 차이를 이용하는 DoG 연산은 다른 편차(σ) 값에서 두 개의 가우시안 마스크를 차 연산 마스크로 회전(convolution)함으로써 수행된다. DoG 알고리즘은 영상에서 정보를 읽어 들이는 망막과 유사한 기능을 가지는 것으로 알려져 있다. DoG는 다음 (식 11)으로 계산된다.

$$DoG = G_{\sigma_1} - G_{\sigma_2} = \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\sigma_1} e^{-(x^2+y^2)/2\sigma_1^2} - \frac{1}{\sigma_2} e^{-(x^2+y^2)/2\sigma_2^2} \right] \quad (11)$$

DoG 연산은 σ_1 과 σ_2 값을 변화시키면서 검출할 에지의 넓이를 조절할 수 있다. 다음 그림 8은 영상에 DoG를 적용한 결과인데 경계 부분을 확실하게 얻어낼 수 있었다. 그림 9는 DoG 적용 후 3D 재구성을 거친 3D 영상이다.

그림 8에서 보듯이 LoG 방법보다 노이즈가 상당 부분 감소하였으며 정합 실패의 결과도 많이 줄어든

4. 구현된 햅틱스 시스템에서의 실험 및 분석

본 연구를 위하여 스테레오 카메라를 통하여 입력된 2D 장면을 3D 장면으로 렌더링 한 후 햅틱 렌더링을 적용하는 실험을 하였다.

3D 재구성의 첫 번째 단계는 카메라와 렌즈를 선택하는 것이다. 그 중 스테레오 정합에 영향을 주는 카메라의 스펙은 해상도, 화각, 초점거리이고, 렌즈의 초점 조절 기능은 스테레오 보정에 영향을 주기 때문에 초점 조절 기능을 사용하지 않는 것이 좋다. 해상도는 스테레오 시스템의 정확도에 대하여 중요한 역할을 한다. 해상도는 카메라 센서의 픽셀 배열에 따라 결정되고 보통 CCD 카메라를 사용한 사각형의 빛 감지 센서를 가지고 있다.

화각과 초점 거리는 역의 관계를 가지고 있다. 초점거리가 증가하게 되면 화각이 좁아지게 되며 유효 해상도는 증가하게 된다. 초점거리가 감소하게 되면

보여 지는 이미지의 양이 증가하게 되지만 유효 해상도는 감소하게 된다. 스테레오 비전 시스템에서 화각은 겹쳐지는 두 카메라 사이에서 겹쳐지는 이미지의 영역을 결정하는데 중요한 역할을 하게 된다. 결과적으로 3D 좌표를 얻을 수 있는 영역은 이 겹쳐진 영역에서만 가능하다.

4.1 연산 속도 비교

실시간으로 처리되어야 하는 동영상의 특성 때문에 연산속도에 관한 문제는 중요한 이슈가 된다. 이 연구에서는 기본 블록 매칭 방법에 새로운 연산을 추가하는 방법을 사용하였기 때문에 연산시간 상승은 피할 수 없는 문제이다. 따라서 추가된 LoG 연산과 DoG 연산 중 어느 방법이 효율적으로 빠르게 되는지 확인하기 위하여 카메라로부터의 영상의 획득에서부터 획득된 영상을 정합하고 점운을 만드는 데까지 걸리는 시간을, LoG 방법, DoG 방법, 그리고 경계면 추출을 추가적으로 하지 않는 기본(Basic) 방법과 비교함으로써 어떤 연산이 더 빠르게 수행되는지 확인하였다.

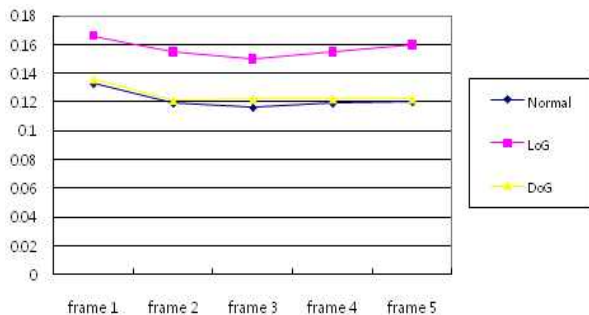


그림 9. 연산 시간 측정(단위 : millisecond)

실험 결과는 그림 8에서 보는 바와 같이 예상대로 LoG 방식이 연산 시간이 많이 소요되는 것으로 나타났고 DoG 방법은 차 연산만을 바탕으로 하는 방법이 기 때문에 기본적인 방법보다는 약간 높게 나오기는 했지만 그 증가 폭은 미미하게 나타났다. LoG 방법도 크게 시간이 증가된 것은 아니지만 실시간 연산을 해야 하는 동영상 응용에서는 가능한 한 연산 시간이 적을수록 많은 데이터를 처리할 수 있기 때문에 DoG 방법이 더 적합하다고 볼 수 있다. 또한 그래픽 렌더링과 유사하게 동작하는 햅틱 렌더링의 특성을 고려할 때, DoG 방법이 LoG보다 경계 면을 더 확실

하게 추출하기 때문에 햅틱 렌더링에 있어서도 DoG 방법이 더 우수한 것으로 판단된다.

4.2 계산된 거리의 오차

위에서 설명한 것과 같이 스테레오 시스템에서 해상도와 정확도는 깊이 값을 복원하는데 있어 중요한 역할을 한다. 정확도는 이미지에서 더 정확한 매칭 좌표를 식별하는 것을 말한다. 올바른 스테레오 정합 과정에서 얻어낸 깊이 해상도를 ΔZ 라 하면 다음(식 12)으로 구할 수 있다.

$$\Delta Z = \frac{Z^2}{fB} \Delta d \tag{12}$$

식 13에서 Z는 정합과정에서 얻어낸 거리를 말하고, f는 렌즈의 초점 거리, B는 기준선, 그리고 n는 센서의 픽셀 크기를 뜻한다. 일부 저가의 센서의 픽셀은 서로 다른 가로와 세로 크기를 나타내게 된다. 여기서는 표 1과 같은 3.5mm의 초점거리와 0.00375mm의 픽셀크기를 가지는 것으로 계산한 후 실측된 값과 스테레오 정합 후 얻어낸 깊이 값을 비교하여 보았다.



그림 10. 영상의 테스트 영역

그림 11의 영상내의 1번과 2번 부분에 대하여 기본(Basic) 방법과 LoG, DoG 각각의 방법으로 거리에 대한 오차율을 계산하였다.

표 2. 거리에 대한 오차율

	1번 지점			2번 지점		
	실측값 (m)	계산값 (m)	오차율 (%)	실측값 (m)	계산값 (m)	오차율 (%)
Basic	1.72	1.74	1.16	2.54	2.51	1.19
LoG	1.72	1.70	1.17	2.54	2.59	1.96
DoG	1.72	1.70	1.17	2.54	2.58	1.57

위 표의 실측값은 실제 카메라와의 거리이며 계산된 값은 위의 (식 13)에 따라 계산된 값이다. 이 결과로 볼 때, LoG와 DoG를 적용하여 개선한 각 방법은 기본(Basic) 방법에 비하여 약간 오차 값이 증가한 부분이 있었지만 큰 차이는 없는 것으로 나타났다. 그러므로 개선된 방법을 사용하여도 스테레오 정합 과정에는 크게 영향을 미치지 않는 것을 알 수 있었다. 따라서 개선된 방법은 오차의 영향이 크게 생기지 않을뿐더러 더 나은 3D 재구성 결과를 기대할 수 있다고 여겨진다.

4.3 햅틱 렌더링에 사용되는 3D 좌표의 개수

햅틱 렌더링 좌표는 재구성된 3D 좌표를 이용하여 렌더링 되는데 렌더링을 구성하는 3D 좌표가 증가함에 따라 실제 환경을 사용자에게 정확하게 전달할 수 있다. 따라서 LoG와 DoG 각각의 방법에 따라 햅틱 렌더링을 하고 그 햅틱 렌더링에 사용되는 3D 좌표의 수를 비교하여 어느 방법이 햅틱 렌더링에 더 적합한지 비교해 보았다.

그림 12에서 보는 바와 같이 DoG 방법을 사용하는 경우 햅틱 렌더링 되는 좌표의 개수가 가장 많고 실제로 경계 면과 평면에 있어 가장 실제와 유사하게 나타나는 것을 확인할 수 있다. 그리고 LoG 방법을 사용하는 경우에는 평면의 노이즈 감소는 효과적이거나 경계 면을 너무 굵게 추출함으로 인해 평면이 무너지는 현상이 발생하였다.

2D 장면에서는 LoG 방법이 경계면 추출 없는 기본(Basic) 방법보다 우수할 수 있으나 이 논문에서와 같이 3D 좌표를 얻어내고 햅틱 렌더링을 적용하는 경우에 있어서는 오히려 기존(Basic) 방법보다 더 떨어지는 것으로 나타났다. 따라서 본 연구의 결과에서는 DoG 방법이 스테레오 비전 기반 햅틱스 시스템에

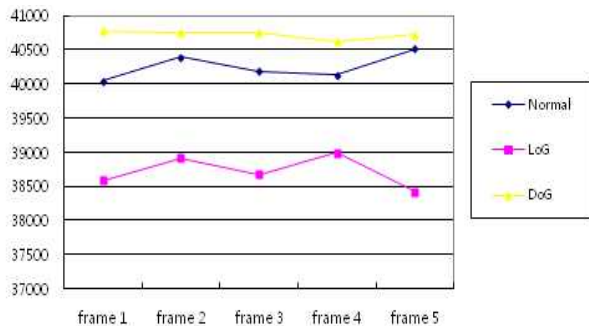


그림 11. 햅틱 렌더링에 사용되는 3D 좌표의 개수 비교

가장 적합한 것으로 나타났다.

5. 결 론

본 논문은 로봇의 시각과 촉각 구현에 응용될 수 있는 스테레오 비전 영상처리 기술과 햅틱스 기술의 융합에 대한 연구이다. 지금까지의 햅틱스 시스템 연구는 3D 그래픽스 객체에 대한 렌더링이 대부분이었다. 본 연구에서 추구하는 실시간 영상에 대한 직접적인 햅틱 렌더링은 3D 그래픽스 객체에 대한 햅틱 렌더링 보다 훨씬 복잡하고 어렵다.

스테레오 카메라로 입력된 실제 비주얼 환경을 3D로 구성하는 과정에서 입체감을 주는 스테레오 영상 처리, 그리고 촉감을 느끼게 하는 햅틱 렌더링 기술을 사용하였다. 햅틱 렌더링에서의 정확한 촉감 전달을 위하여, LoG와 DoG 방법을 적용하여 경계면을 확실하게 추출하고 스테레오 정합을 수행함으로써 보다 정확하게 촉감을 제시하는 방법을 제안하였다.

LoG와 DoG 두 가지 방법에 대한 실험과 분석을 통하여 아래의 결과를 얻었다.

- 연산 시간이 중심이 될 때는 LoG와 DoG 방법을 사용하지 않고 경계면 추출 없는 기본(Basic) 방법으로 스테레오 정합을 수행하는 것이 유리하다.
- DoG 방법을 사용하였을 경우 연산시간이 다소 증가하기는 하지만, 그림 13에서 보는 바와 같이 노이즈와 정합이 실패하는 영역의 면에 있어서 LoG 방법보다 DoG 방법이 더 우수한 것으로 나타났다. 또한 DoG 방법이 햅틱 렌더링을 위한 3D 재구성에 있어서 더욱 실제와 유사하게 렌더링이 되는 것을 확인하였다.

이상으로 볼 때, 본 연구와 같이 햅틱 렌더링을 추가하여 사용하는 경우에는 노이즈 감소와 경계면의 추출이 확실한 DoG 방법이 가장 효과적인 것으로 결론 내려진다.



(a) LoG 방법

(b) DoG 방법

그림 12. 햅틱 렌더링

스테레오 영상에 햅틱 렌더링을 적용하는 연구는 아직 초보적인 단계이다. 본 논문의 독창성은 스테레오 영상에 직접적인 촉감을 실시간으로 제시하기 위하여 요구되는 효과적인 경계면 추출 방법에 대한 실험적인 결과를 제시한다는 점이다.

본 논문에서 제안하는 스테레오 비전 기반 햅틱스 시스템을 위한 3D 재구성 방법은 로봇 자율 주행 및 장애물 인식 등의 지능형 로봇 분야의 응용뿐만 아니라 영화, 애니메이션, 방송, 가상현실(virtual reality) 등의 여러 산업 분야나 군사 분야에서 요소기술로 응용이 가능할 것이다.

참 고 문헌

- [1] D. Marr and E. Hildreth, "Theory of Edge Detection," *Proce. of the Royal Society of London. Series B, Biological Sciences*, Vol. 207, No.1167, pp. 187 - 217, 1980.
- [2] Steve R. Gunn, "On the Discrete Representation of the Laplacian of Gaussian," *Pattern Recognition*, Vol.32, No.8, pp. 1463-1472, 1999.
- [3] Difference of Gaussians, http://en.wikipedia.org/wiki/Difference_of_Gaussians, 2011.
- [4] Massimo Bertozzi, "GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection," *IEEE Transactions on Image Processing*, Vol.7, No.1, pp. 62-81, 1998.
- [5] Liang Zhao, "Stereo-and Neural Network-Based Pedestrian Detection," *IEEE Transactions on Intelligent Transportation Systems*, Vol.1, No.3, pp. 148-154, 2000.
- [6] Don Murray and James J. Little, "Using Real-Time Stereo Vision for Mobile Robot Navigation," *Autonomous Robots*, Vol.8, No.2, pp. 161-171, 2000.
- [7] J. Brian, Rogers and Ian P. Howard, *Binocular Vision and Stereopsis*, Oxford Psychology Series, Oxford University Press, 1995.
- [8] M.A. Srinivasan and C. Basdogan, "Haptics in Virtual Environments: Taxonomy, Research Status and Challenges," *Computers & Graphics*, Vol.21, Issue 4, pp. 393-404, 1997.
- [9] T.H. Massie, *Design of a Three Degree of Freedom Force-Reflecting Haptic Interface*, BS Thesis, MIT EECS Department, 1993.
- [10] William R. Mark, Scott C. Randolph, Mark Finch, James M. Van Verth, and Russell M. Taylor II, "Adding Force Feedback to Graphics Systems: Issues and Solutions," *Proc. of The 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*, pp. 447-452, 1996.
- [11] K. Salisbury, F. Conti, and F. Barbagli, "Haptic rendering: Introductory Concepts," *IEEE Computer Graphics and Applications*, Vol.24, No.2, pp. 24-32, 2004.
- [12] 이환문, 김기권, 성미영, "만질 수 있는 비디오를 위한 햅틱 렌더링 기술", 멀티미디어학회논문지, 제13권, 제5호, pp. 691-701, 2010.
- [13] A. Yamamoto. et al. "Electrostatic Tactile Display with Thin Film Slider and Its Application to Tactile Telepresentation Systems," *Transaction on Visualization and Computer Graphics*, Vol.12, No.2, pp. 168-177, 2006.
- [14] Gary R. Bradski, and Adrian Kaehler, "Robot-Vision Signal Processing Primitives," *IEEE Signal Processing Magazine*, Vol.25, Issue 1, pp. 130-133, 2008.
- [15] Janne Heikkila and Olli Silven, "A Four-Step Camera Calibration Procedure with Implicit Image Correction," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 1106-1112, 1997.
- [16] Gary R. Bradski, Vadim Pisarevsky, and Jean-Yves Bouguet, *Open Source Computer Vision Library*, Springer, 1st edition, New York/Heidelberg, 2006.
- [17] Mohsen Sharifi, Mahmoud Fathy, and Maryam Tayefeh Mahmoudi, "A Classified and Comparative Study of Edge Detection Algorithms," *Proc. of the International Conference on Information Technology: Coding and*

Computing (ITCC '02), pp. 117-120, 2002.

- [18] John Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.8, No.6, pp. 679-698, 1986.



성 미 영

1982년 서울대학교 학사
1987년 프랑스 INSA de Lyon 컴
퓨터공학 석사
1990년 프랑스 INSA de Lyon 컴
퓨터공학 박사
1990년~1993년 한국전자통신연
구소 선임연구원

1993년~현재 인천대학교 컴퓨터공학부 교수
관심분야 : 멀티미디어, 가상현실, 햅틱스



김 기 권

2009년 인천대학교 컴퓨터공학과
학사
2009~2011년 인천대학교 컴퓨터
공학과 석사

관심분야 : 멀티미디어, 컴퓨터비
전, 햅틱스