

스테레오정합과 신경망을 이용한 3차원 잡기계획

이현기[†] · 배준영* · 이상룡**

(2002년 9월 18일 접수, 2003년 5월 26일 심사완료)

3D Grasp Planning using Stereo Matching and Neural Network

Hyun-Ki Lee, Joon-Young Bae and Sang-Ryong Lee

Key Words : Stereo Matching(스테레오정합), Neural Network(신경망), Grasp Planning(잡기계획)

Abstract

This paper deals with the synthesis of the 3-dimensional grasp planning for unknown objects. Previous studies have many problems, which the estimation time for finding the grasping points is much long and the analysis used the not-perfect 3-dimensional modeling. To overcome these limitations in this paper new algorithm is proposed, which algorithm is achieved by two steps. First step is to find the whole 3-dimensional geometrical modeling for unknown objects by using stereo matching. Second step is to find the optimal grasping points for unknown objects by using the neural network trained by the result of optimization using genetic algorithm. The algorithm is verified by computer simulation, comparing the result between neural network and optimization.

1. 서론

인간은 물건을 보고 직관적으로, 경험적으로 물건을 잡게 된다. 하지만 로봇의 경우는 그렇지 못하다. 인간의 조작에 의해서나 계산에 의한 결과를 가지고 물건을 잡게 된다. 하지만 인간이 로봇을 조작할 수 없는 장소나, 물건에 대한 정보가 전혀 없는 공간에서 로봇이 작업해야 하는 상황이 된다면 계산적으로, 인간의 조작에 의한 작업을 할 수 없게 된다.

이런 문제들을 해결하기 위해 1980년부터 잡기계획(Grasp Planning)이라는 제목으로 많은 연구들이 행하여져 왔다. 잡기계획에 있어서 가장 중요시되는 것은 손가락의 위치와 힘의 적용에 대한 안정성이다.

잡기계획에 있어서 가장 기초가 되고 중요한 토대가 되는 것은 Nyguen⁽¹⁾에 의해 소개되어진 "Force Closure"의 개념이다. 마찰을 고려하고, 물체

와 그리퍼와의 접촉을 점 접촉으로 고려하여 안정적인 잡는 점(grasping points)을 찾는 방법을 제시하였다. 유연한 잡는 시스템을 구현하기 위해서 여러 논문에서 비전 센서를 장착한 시스템을 연구하였다. Hauck⁽²⁾는 임의의 물체를 두대의 카메라를 이용하여 3-차원 공간상에서 2-손가락 그리퍼의 잡는 점을 찾는 알고리즘을 개발하였고, Morales⁽³⁾는 임의의 물체를 한대의 고정된 카메라를 이용하여 2-손가락 그리퍼의 최적의 잡는 점을 찾는 알고리즘을 개발하였다.

이와는 별도로 3-손가락 이상의 다손가락(multi-finger)의 최적 잡는 점에 대한 연구도 많이 되어져왔다. 다손가락은 2-손가락과는 다르게 2-차원 해석이 불가능하여, 알고 있는 물체들을 이용하여 3 차원에서 해석을 수행하였다. Borst⁽⁴⁾는 "Grasp Quality"⁽⁵⁾를 이용하여 빠르고, 강건한 잡는 점을 찾는 알고리즘을 제시하였고, Katada⁽⁶⁾는 리아프노프 안정도해석(Lyapunov Stability) 이론을 도입하여 안정한 잡는 점을 찾고, "Evolution Programming"을 이용하여 최적의 점을 찾는 알고리즘을 개발하였다. 이런 기존의 연구들은 각자 한계를 가지고 있었다.

첫째로 비전을 이용하는 경우중에서 한대의 카

[†] 책임저자, 회원, 경북대학교 기계공학과

E-mail : mradam1004@hotmail.com

TEL : (053)952-2229 FAX : (054) 952-0648

* 회원, 경북대학교 대학원 기계공학과

** 회원, 경북대학교 기계공학부

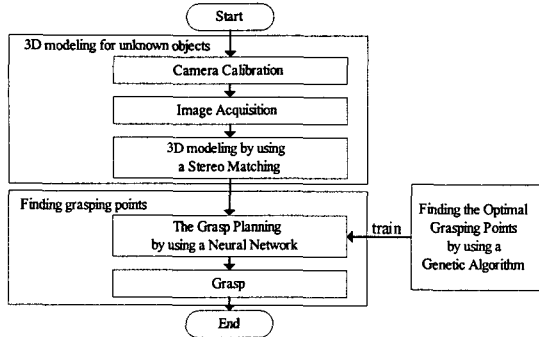


Fig. 1 Block diagram of proposed approach

메라를 이용하는 경우 다손가락 핸드의 경우 해석이 전혀 불가능하다. 두번째로 두대의 카메라를 이용하는 경우 연직 상방에서만 물체를 촬영하고 대부분 2-손가락 핸드의 해석이 그치는 경우가 많이 있었다. 마지막으로 다손가락 해석을 하는 연구들을 보면 대부분 알고 있는 물체에 대한 해석을 하였다. 이를 극복하기 위해 두대의 카메라를 이용하여 물체를 3차원으로 모델링하고, 유전자알고리즘을 이용하는 연구를 시행하였다.⁽⁷⁾ 하지만 이 연구도 최적의 잡는 점을 찾는 데 시간이 상당히 많이 소요되었다. 그래서 본 연구에서는 최적의 잡는점을 찾는 기존의 연구의 결과를 그대로 가져오면서 빠른 시간안에 결과를 가져다 주는 알고리즘을 개발하고자 하였다.

본 논문에서는 임의의 물체를 두대의 카메라를 이용하여 3-차원으로 해석하고, 해석된 물체를 바탕으로 3-손가락 그리퍼의 최적의 잡는 점을 찾는 것을 목적으로 하고 있다. 따라서 논문의 방향은 Fig. 1 과 같이 크게 두 가지로 나눌 수 있다.

첫번째 부분은 물체에 대한 정보가 없을 경우 두대의 카메라를 이용한 스테레오정합을 이용하여 잡으려고 하는 물체의 정보를 찾는 과정이다. 이 과정에서 기존에 많이 쓰이고 있는 SMW⁽⁸⁾ 방법이 도입되었다. 두번째 부분은 유전자 알고리즘의 결과를 이용하여 학습된 신경망을 이용하여 최적의 잡는점을 찾는 부분이다.

2. 3 차원 영상처리 기술

본 논문에서는 Fig. 2 에서 보는 바와 같이 두대의 카메라를 이용하여 3-차원 물체를 인식시키는 방법인 스테레오 정합을 사용하였다. 이 방법

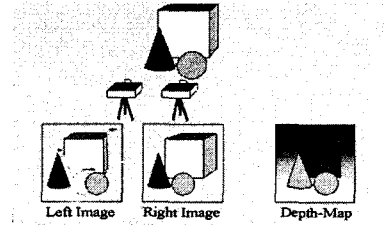


Fig. 2 Principle of stereo matching

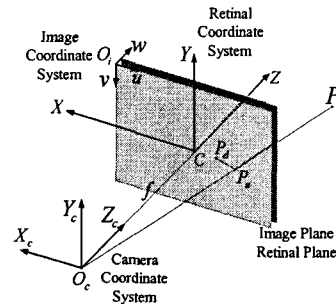


Fig. 3 Pinhole camera model

은 인간의 시각체계를 형상화한 것이라 볼 수 있다.

스테레오 정합에는 여러 종류가 있다. 먼저 정합 요소에 따라 특징 기반법과 영역 기반법으로 나눌 수 있고, 둘째로 카메라의 내부, 외부 파라미터들을 알고 있느냐, 모르고 있느냐에 따라 Calibrated, Uncalibrated Stereo Matching 으로 나눌 수 있다. 본 연구에서는 특징 기반법을 기초로 한 Calibrated Stereo Matching 을 수행하였다.

2.1 Pinhole 카메라 모델

일반적인 카메라는 Fig. 3 과 같이 모델링 할 수 있다. Fig. 3 에서 f 는 초점거리(focal distance), O_c 는 초점점(focal point), c 는 영상의 중심(image center), P 는 실제의 3-차원 점이고, P_i 는 P 점이 실제 영상에서 보이는 점의 위치를 가리키고, P_c 는 P 점이 camera coordinate 에서 이상적으로 위치하는 점의 위치를 가리킨다.

위의 모델에서 아래와 같은 관계를 구할 수 있다.

$$sm_i = M_{int} M_{ext} \begin{bmatrix} M \\ 1 \end{bmatrix} \quad (1)$$

식 (1)에서 s 는 일반적인 상수로 물체와의 깊이 정보를 가리키고, m_i 는 image coordinate 에서의 p_i 의 좌표를 가리키고, M 은 global coordinate 에서의 p_i 의 좌표를 가리킨다. 또, M_{in} 은 카메라의 초점

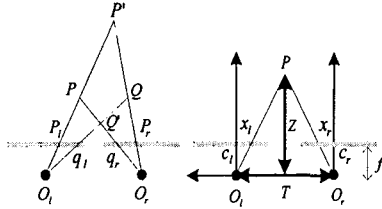


Fig. 4 Geometry of stereo camera

거리와 광중심으로 이루어진 내부 파라미터 (intrinsic parameter)를 가리키고, M_{ext} 은 카메라의 회전과 병진의 파라미터로 이루어진 외부 파라미터 (extrinsic parameter)를 가리킨다.

여기에서 M_{int} 은 아래와 같이 쓸 수 있다.

$$M_{int} = \begin{bmatrix} -f/s_x & f/s_x \cos \alpha & u_0 \\ 0 & -(f/s_y)/\sin \alpha & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

여기에서 s_x 와 s_y 는 CCD 카메라의 픽셀의 가로, 세로 사이즈를, $c = [u_0 \ v_0]^T$ 카메라의 광중심의 좌표를, α 는 카메라 픽셀의 기울어진 각도를 가리킨다.

$M_{ext} = [R \ T^T]$ 는 카메라의 회전과 병진의 파라미터로 이루어져 있다.

2.2 Calibrated Stereo Camera

Fig. 4 에서 특별한 경우의 스테레오 정합 모습을 보여준다. 이 경우 두 카메라의 광축이 평행하고, 좌, 우측 영상이 맺히는 평면이 동일 평면상에 존재한다. f 는 각 카메라의 초점거리를 나타내고, T 는 두 평행 광축 사이의 거리를 나타내고, O_l 과 O_r 은 투영의 중심점이고, $x_l - x_r$ 은 디스페리티 (disparity)라고 하여 스테레오 정합에 있어서 실제 거리를 구하는 중요한 파라미터가 된다. Fig. 4 에서 실제물체의 점과 좌, 우측 영상에 나타난 점의 기하학적인 구조를 이용하여 실제점의 거리를 구할 수 있다.

$$Z = \frac{T \times f}{x_l - x_r} \quad (3)$$

2.3 Stereo matching

본 논문에서 상용한 스테레오정합의 알고리즘을 간단히 그려보면 아래 Fig. 5 와 같다.

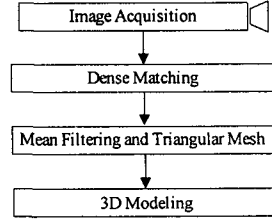


Fig. 5 Sequence of stereo matching

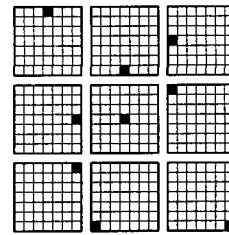


Fig. 6 Nine different correlation windows. The pixel for which disparity is computed is highlighted

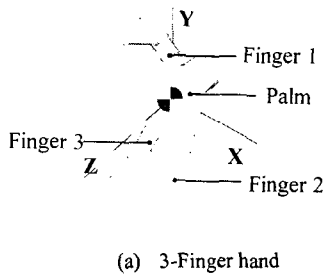
Fig. 5 에서 먼저 기초화 되어진 두 대의 스테레오 카메라를 이용하여 스테레오 영상을 얻는다.⁽¹⁵⁾ 다음으로 밀도 있는 정합 과정을 거쳐야 한다. 기초화(Calibrate)된 좌, 우측 영상은 X 축으로만 정합 점들을 조사해 나가면 된다. 그리고 앞서 구한 정합 점들의 정보를 이용하여 좌측의 한 점에 대한 우측의 디스페리티 정보를 이용하여 아주 빠른 시간 내에 밀도있는 깊이 지도를 구할 수 있다.

Fusiello, Robert and Trucco⁽⁸⁾가 이런 알고리즘을 제안했다. 좌측영상의 한 점에서 우측영상의 디스페리티내에 드는 모든 점들에 대한 유사도를 측정한다. 이 유사도는 SSD(Sum of Squared Difference) 에러로 측정 되어지고,

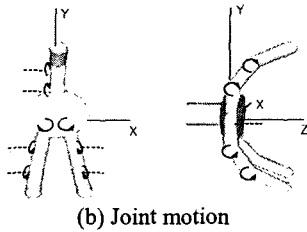
$$c(x, y, d) = \frac{\sum_{i=-n}^n \sum_{j=-m}^m [I_l(x+i, y+j) - I_r(x+i+d, y+j)]^2}{\sqrt{\sum_{i=-n}^n \sum_{j=-m}^m I_l(x+i, y+j)^2 \sum_{i=-n}^n \sum_{j=-m}^m I_r(x+i+d, y+j)^2}} \quad (4)$$

이 SSD 에러가 가장 작은 점이 정합점이 된다. 좀더 정밀하게 정합점을 구하기 위해서, 에러를 곡선으로 보간해서 주위 점들과의 에러를 작게하는 방법을 사용하였다. 이 방법은 아래의 Fig. 6 과 같이 여러 개의 비교 윈도우를 사용하는 SMW⁽⁸⁾ 방법이다.

밀도 있는 정합을 한 후에 밀도 깊이 지도 (Dense Depth Map)를 보면 많은 노이즈 값들이 있다. 이런 값들을 줄이기 위해 평균필터(Mean Filter)가 적용 되었고, 3 차원 모델링에 적용하기



(a) 3-Finger hand



(b) Joint motion

Fig. 7 3-Finger hand modeling

위해 삼각형 메쉬(Triangular Mesh)가 적용 되었는데 Delaunay Triangulation Method⁽⁹⁾가 사용 되었다.

3. 신경망을 이용한 잡기 계획

3.1 3-손가락 핸드의 모델링

본 연구에서는 Fig. 7 과 같이 3-손가락 핸드를 모델링하여 모사실험을 수행하였다. 3-손가락 핸드 는 다손가락에 해당하면서 해석이 가장 용이하고, 다손가락의 특성을 잘 표현할 수 있는 장점이 있다.

본 연구에서 사용된 3-손가락 핸드는 총 8 개의 회전운동을 하는 조인트로 구성된 8 자유도 핸드 이다. 하지만 손가락을 각각 해석하면 손가락 1 은 2 자유도, 손가락 2 와 3 은 각각 3 자유도를 가진다. 위의 3-손가락 핸드에 대해서 기구학, 역기구학, 자코비안 행렬, 손가락 끝단 힘에 대한 모델링을 하였다.

본 연구에서는 손바닥의 중심에 핸드의 좌표계를 설치하였고, 이것을 중심으로 각 손가락의 파라미터들을 구하였다. 손바닥과 손가락의 첫마디 는 물체의 XY 평면과 평행하게 만들었고, 손의 중심 좌표와 물체의 중심좌표 점을 일치 시켜 계산하였다. 각 손가락의 계산식과 파라미터들의 계산식은 아래와 같다.

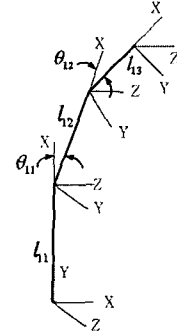


Fig. 8 Modeling of finger 1

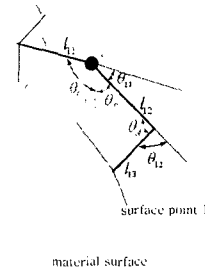


Fig. 9 Inverse kinematics of finger 1

손가락 1 을 아래 그림과 같이 핸드의 중심을 중심으로 모델링을 하면, 손가락 끝단의 위치는 아래의 식과 같이 나타난다.

$$\begin{aligned} x_1 &= 0 \\ y_1 &= l_{11} + l_{12} \cos \theta_{11} + l_{13} \cos(\theta_{11} + \theta_{12}) \\ z_1 &= l_{12} \sin \theta_{11} + l_{13} \sin(\theta_{11} + \theta_{12}) \end{aligned} \quad (5)$$

손가락 1 에대한 역기구학을 구해보면

$$\begin{aligned} \theta_{11} &= 180 - (\theta_b + \theta_c) \\ \theta_{12} &= 180 - \theta_a \end{aligned} \quad (6)$$

으로 나타나고, 자코비안 매트릭스와 손가락 끝단 힘은 아래의 식 (7)과 (8)로 나타난다.

$$J_1 = \begin{bmatrix} 0 & 0 \\ -l_{12} s_{11} - l_{13} s_{11+12} & -l_{13} s_{11+12} \\ l_{12} c_{11} + l_{13} c_{11+12} & l_{13} c_{11+12} \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} \quad (7)$$

$$F_1 = (J_1^T)^{-1} \begin{bmatrix} \tau_{11} \\ \tau_{12} \end{bmatrix} \quad (8)$$

다음으로 손가락 2 와 3 을 위의 손가락 1 과 같

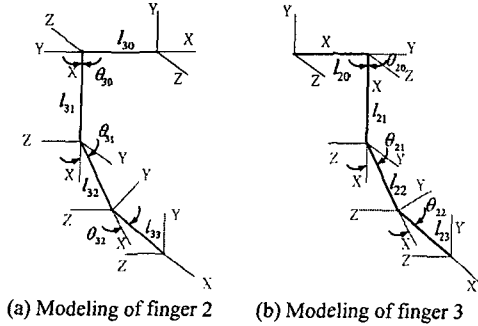


Fig. 10 Modeling of finger 2 and 3

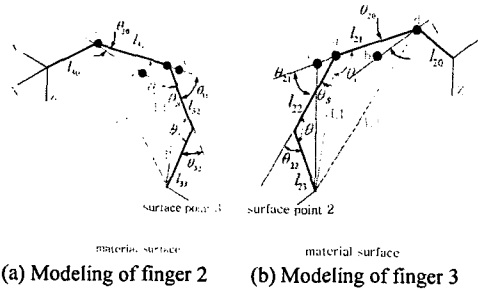


Fig. 11 Inverse kinematics of finger 2 and 3

은 방법으로 구해보면 아래와 같이 구할 수 있다. 손가락 2 와 3 은 구조는 동일하나 조인트의 방향이 반대라는 특징이 있다. 먼저 손가락 2 와 3 을 모델링해 보면, 손가락 2 와 3 의 끝단 위치는 아래와 같은 식으로 구할 수 있다.

손가락 2 의 끝단 위치 :

$$\begin{aligned} x_2 &= l_{20} + [l_{21} + l_{22} \cos \theta_{21} + l_{23} \cos(\theta_{21} + \theta_{22})] \sin \theta_{20} \\ y_2 &= -[l_{21} + l_{22} \cos \theta_{21} + l_{23} \cos(\theta_{21} + \theta_{22})] \cos \theta_{20} \\ z_2 &= l_{22} \sin \theta_{21} + l_{23} \sin(\theta_{21} + \theta_{22}) \end{aligned} \quad (9)$$

손가락 3 의 끝단 위치 :

$$\begin{aligned} x_3 &= -l_{30} + [l_{31} + l_{32} \cos \theta_{31} + l_{33} \cos(\theta_{31} + \theta_{32})] \sin \theta_{30} \\ y_3 &= -[l_{31} + l_{32} \cos \theta_{31} + l_{33} \cos(\theta_{31} + \theta_{32})] \cos \theta_{30} \\ z_3 &= l_{32} \sin \theta_{31} + l_{33} \sin(\theta_{31} + \theta_{32}) \end{aligned} \quad (10)$$

다음으로 역기구학을 구하기 위해 다음과 같이 손가락 2 와 3 을 모델링 하였다.

위의 모델링을 이용하여 손가락 2 와 3 의 역기구학을 다음의 식과 같이 구하였다.

손가락 2 의 역기구학:

$$\begin{aligned} \theta_{20} &= \text{sign}(cd \times bd) \cdot \cos^{-1} \left(\frac{cb^2 - dc^2 - bd^2}{-2dcdb} \right) \\ \theta_{21} &= 180 - (\theta_B + \theta_C) \\ \theta_{22} &= 180 - \theta_A \end{aligned} \quad (11)$$

손가락 3 의 역기구학:

$$\begin{aligned} \theta_{30} &= \text{sign}(bd \times cd) \cdot \cos^{-1} \left(\frac{cb^2 - dc^2 - bd^2}{-2dcdb} \right) \\ \theta_{31} &= 180 - (\theta_B + \theta_C) \\ \theta_{32} &= 180 - \theta_A \end{aligned} \quad (12)$$

다음으로 조인트의 값과 토크의 값들을 이용하여 손가락 2 에 대한 자코비안 매트릭스 식 (13)와 손가락 끝단힘 식 (14)과 손가락 3 에 대한 자코비안 매트릭스 식 (15)와 손가락 끝단힘 식 (16)를 아래와 같이 구할 수 있다.

$$J_2 = \begin{bmatrix} (l_{21} + l_{22}c_{21} + l_{23}c_{21+22})c_{20} & -l_{22}s_{20}s_{21} - l_{23}s_{20}s_{21+22} & -l_{23}s_{20}s_{21+22} \\ (l_{21} + l_{22}c_{21} + l_{23}c_{21+22})s_{20} & l_{22}c_{20}s_{21} + l_{23}c_{20}s_{21+22} & l_{23}c_{20}s_{21+22} \\ 0 & l_{22}c_{21} + l_{23}c_{21+22} & l_{23}c_{21+22} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (13)$$

$$F_2 = (J_2^T)^{-1} \begin{bmatrix} \tau_{20} \\ \tau_{21} \\ \tau_{22} \end{bmatrix} \quad (14)$$

$$J_3 = \begin{bmatrix} -(l_{31} + l_{32}c_{31} + l_{33}c_{31+32})c_{30} & l_{32}s_{30}s_{31} - l_{33}s_{30}s_{31+32} & l_{33}s_{30}s_{31+32} \\ (l_{31} + l_{32}c_{31} + l_{33}c_{31+32})s_{30} & l_{32}c_{30}s_{31} + l_{33}c_{30}s_{31+32} & l_{33}c_{30}s_{31+32} \\ 0 & l_{32}c_{31} + l_{33}c_{31+32} & l_{33}c_{31+32} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad (15)$$

$$F_3 = (J_3^T)^{-1} \begin{bmatrix} \tau_{30} \\ \tau_{31} \\ \tau_{32} \end{bmatrix} \quad (16)$$

3.2 유전자 알고리즘을 이용한 잡기계획

유전자 알고리즘을 이용하여 모델링 된 물체에 있어서 최적의 잡는 점을 이용하는 방법이다.⁽⁷⁾ 유전자 알고리즘의 결과를 이용하여 신경망을 학습시키는 데이터로 이용한다.

강제로 가정된 임의의 물체에 위와 같이 힘이 작용한다면, 물체를 안정적으로 잡기 위해서는 힘의 평형이 맞아야 한다.

Fig. 12 에서 물체에 작용하는 모든 힘을 각각 f_1, f_2, \dots, f_n 이라 하고, 무게중심으로부터 힘이 작용하는 점까지의 위치 벡터를 각각 $\rho_1, \rho_2, \dots, \rho_n$ 이라 하면 식 (17)과 식 (18) 같은 방정식을 만족하여야 한다.

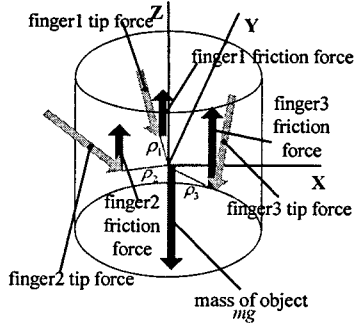


Fig. 12 Grasp equilibrium

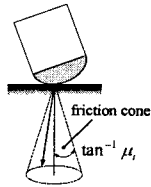


Fig. 13 Coulomb friction model

$$\text{힘의 평형} : \sum_{i=1}^n f_i = -mg \quad (17)$$

$$\text{모멘트의 평형} : \sum_{i=1}^n \rho_i \times f_i = 0 \quad (18)$$

여기에서 m 은 물체의 질량, g 는 중력 가속도를 가리킨다.

또, Fig. 13 과 같이 물체를 들기 위해서는 손가락 끝단과 물체 사이에 마찰력이 작용해야 한다.

$$|n_i \times f_i| \leq \mu_i (f_i \cdot n_i), \quad i = 1, \dots, n \quad (19)$$

여기에서 n_i 는 물체의 표면에 수직한 단위 벡터이고, μ_i 는 마찰계수를 가리킨다.

최적화 알고리즘을 적용하기 전에 먼저 목적함수와 제한조건을 정의한다. 목적함수는 아래와 같이 정의할 수 있다.

$$F = D(p) \quad (20)$$

여기에서 $D(p) = \sum_{i=0}^n (p_i)^2$ 이고, n 은 손가락의 갯수를, p_i 는 각 손가락에서 손가락 끝단에 작용하는 힘으로 정의한다. 식 (20)은 물체를 잡을 때 힘이 가장 적게드는 위치를 찾는다는 것을 의미한다. 제한조건은 잡기평형(Grasp Equilibrium) 조건을 이용하였다.

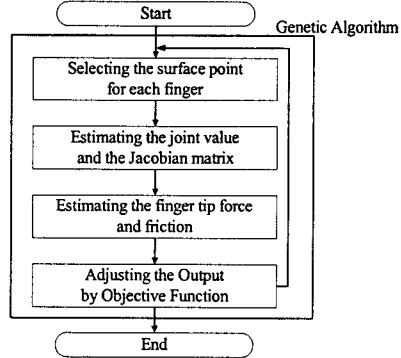


Fig. 14 Flow chart of optimization

위의 조건들에 의해 최적화해야할 파라미터들은 아래와 같다.

- ① 손가락 1,2,3 에 대한 표면점의 높이
- ② 손가락 2,3 에 대한 표면점의 각도
- ③ 각 조인트에 작용하는 토크
- ④ 3-손가락 핸드의 높이와 각도

위의 식들은 비선형 방정식으로 표현이 되고, 위의 최적화 파라미터를 가지고 아래와 같은 수식으로 표현할 수 있다.

$$\text{Minimize} \quad F(y) \quad (21)$$

$$\text{Subject to} \quad g_i(y) \leq 0, \quad i = 1, \dots, l \quad (22)$$

여기에서 l 은 제한조건의 개수이다. 식 (21)과 식 (22)를 풀기 위해 식 (21)과 식 (22)를 벌칙 함수로 변형시켰다.

$$L(y) = F(y) + P \sum_{i=1}^l (g_i^+(y))^2 \quad (23)$$

위의 식 (23)에서 $F(y)$ 는 목적함수, P 는 벌칙 값(penalty value)으로 충분히 큰 값을 선택하고, $g_i^+(y) = \max\{0, g_i(y)\}$ 는 i 번째 제한조건에 대한 위반의 크기를 나타낸다.

식 (23)에서 최적의 값을 찾기 위해서 본 연구에서는 Fig. 14 와 같은 순서로 최적화를 하였고, 최적화를 위해 유전자 알고리즘을 사용하였다.

3.3 신경망을 이용한 잡기계획

신경망을 적용하기 전에 먼저 신경망의 학습이 필요하다. 신경망의 결과로 얻어져야 하는 것은 유전자 알고리즘을 이용하였을 경우와 동일하다.⁽⁷⁾ 유전자 알고리즘에 있어서 출력 파라미터는 모두 24 개로 Table 1 에 그 값들이 자세히 나와있다. 손바닥의 자세와 각 손가락의 위치, 조인트의 값들

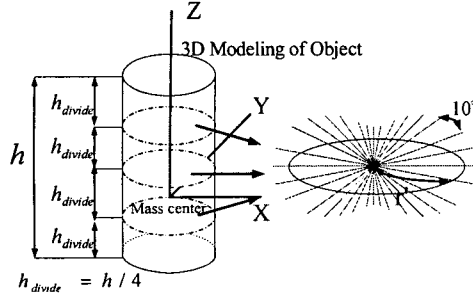


Fig. 15 Input parameter of neural network

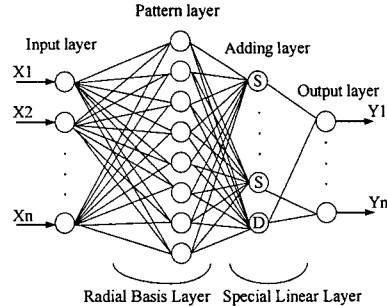


Fig. 16 GRNN

과 토크 들이다. 입력 파라미터는 모델링 된 물체의 형상과 관계가 있다. 따라서 본 연구에서는 물체에서 형상을 정의 하기 위해 아래와 같은 입력 파라미터들을 설정 하였다.⁽¹⁶⁾

입력의 첫 번째와 두 번째 파라미터는 모델링 된 물체의 최고, 최저 깊이의 값으로 모델링된 물체의 높이와 연관이 있다. 이 파라미터는 3-손가락 핸드의 높이와 손가락들의 표면점의 높이와 연관이 있다. 물체의 단면의 모습을 구하기 위한 것으로 Fig. 15 와 같이 나머지를 설정 하였다.

물체의 높이를 네부분으로 나누고, 나뉘어진 세 단면에서 각 10 도 마다의 반지름의 길이를 입력 파라미터로 사용하였다. 위의 그림에서 최상부와 최하부 단면은 노이즈가 많이 생기므로 이용하지 않았다. 결과적으로 110 개의 입력 파라미터와 24 개의 출력 파라미터를 신경망에 적용하였다.

다음으로 신경망을 설정하여야 한다. 본 연구에서는 아래의 그림과 같은 GRNN(Generalized Regression Neural Networks)⁽¹⁰⁾ 모델을 선정 하였다. 이 GRNN 모델은 예측 모델로서 학습이 용이하고, 값을 예측하는 곳에 있어서 뛰어난 기능을 가지고 있다. GRNN 모델은 아래의 그림과 같이 4 개의 층으로 이루어져 있다. Pattern 층은 입력층과 연결이 되어 있고 입력의 개수와 동일하고, Adding 층은 출력의 개수보다 한 개 많은 뉴런을 가지고 있다. 그림에서 S-adding 뉴런은 합하는 요소이고, D-adding 뉴런은 나누는 요소로 한 개가 존재한다.

4. 모사실험과 결과

Fig. 1 에서 본 순서대로 모사실험을 진행하였다. 먼저 두대의 카메라의 기초화가 필요하다. 두대의 카메라의 기초화에는 카메라의 내부, 외부 파라미

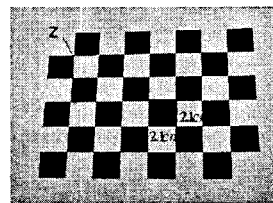


Fig. 17 Planar pattern for camera calibration

터가 모두 필요하지만 본 연구에서 부파라미터는 이미 알고 있으므로 내부 파라미터만 구하였다. 카메라의 초기화는 고전적인 방법으로 평면패턴⁽¹¹⁻¹⁴⁾을 이용하는 방법을 사용하였다.

사용 되어진 평면패턴은 Fig. 17 과 같다. 8×6 개의 2.1cm 길이의 블록들을 이용하였다. 그 결과는 아래와 같다. 두 광축사이의 거리는 2.0cm 로 설정 하였다. 위의 평면 패턴을 이용하면 식 (1)에서 인트린식 파라미터는 아래와 같이 구할 수 있다.

$$M_{in} = \begin{bmatrix} 730.3701172 & 0.0000000 & 306.0815735 \\ 0.0000000 & 733.8732910 & 252.6554108 \\ 0.0000000 & 0.0000000 & 1.0000000 \end{bmatrix} \quad (24)$$

다음으로 알고 있는 물체에 대한 학습이 필요하다. 학습을 위해서 알고 있는 물체에 대해서 유전자 알고리즘을 통해 입력과 출력에 관한 결과를 얻어 내었다. Fig. 18 은 학습을 하기 위해 선택되어진 다양한 물체의 모양이다. 본 연구에서는 원통형 물체에 대해서만 학습을 실시 하였다. 이것은 일반적인 물체의 경우 원통형 물체가 많고, 원통형의 적용후에 일반적인 물체에도 적용이 가능함을 보이기 위함이다. 아래의 원통형 물체들은 높이의 차이, 원통의 반지름의 차이, 타원의 크기의 차이, 각도의 차이를 가지고 있다. 모두 99 개의 학습 데이터를 구하였다.

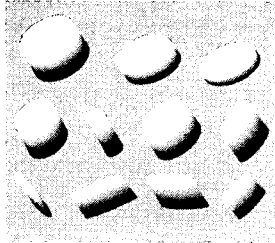
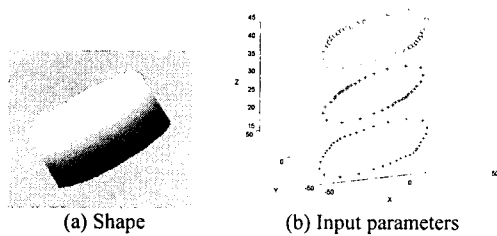
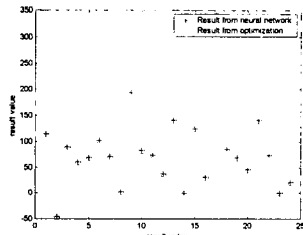


Fig. 18 Known object to get the training data



(a) Shape

(b) Input parameters



(c) Comparing the result from between neural network and optimization

Fig. 19 The input shape and parameters for test to train the network and result

신경망을 학습한 후에 학습에 사용되지 않은 물체를 이용하여 신경망이 바르게 학습이 되었는지 확인하여 보았다.

Fig. 19 에서 볼 수 있듯이, 유전자 알고리즘을 이용하여 구한 출력 파라미터와 신경망을 이용하여 구한 결과와는 거의 차이가 없음을 볼 수 있고, 실제적으로 0.1%정도의 오차를 가지고 있음을 알 수 있다. 실제적인 값을 보면, 최적화를 이용하여 얻은 결과⁽⁷⁾인 Table 1 과 신경망을 이용하여 얻은 결과인 Table 2 에서의 값에는 거의 차이가 없음을 볼 수 있다. Fig. 15(b)는 모델의 좌표계에서 본 입력력파라미터를 표시하고, Fig. 15(c)에서 result index 는 출력 변수들을 순서대로 표시한 것이고, result value 는 출력 값을 의미한다.

다음으로 학습된 신경망을 가지고 Fig. 20 과 같이 실제적인 물체에 적용을 시켜 보았다. 지름은 61.42 mm, 높이는 81.00 mm 인 원통형 물체를 선정하였다. 3-손가락 핸드의 끝에 고무가 달려 있다고

Table 1 The result of optimization parameters

variables	palm	Surface point		
		Finger1	Finger2	Finger3
Angle (degree)	-45.696	90.000	194.514	240.669
Height (mm)	114.335	30.077	33.016	34.5327

variables		Finger1	Finger2	Finger3
Angle (degree)	Joint0		73.669	67.952
	Joint1	68.608	37.747	46.353
	Joint2	102.529	141.439	140.975
Torque (g mm)	Joint0		-0.736	0.428
	Joint1	70.341	125.009	120.166
	Joint2	3.204	31.311	22.0987

Table 2 The result of neural network

Variables	Palm	Surface point		
		Finger1	Finger2	Finger3
Angle (degree)	-46.723	90.000	194.514	241.554
Height (mm)	114.452	29.247	30.436	34.355

Variables		Finger1	Finger2	Finger3
Angle (degree)	Joint0		73.669	67.952
	Joint1	68.608	37.747	46.353
	Joint2	102.529	141.439	140.975
Torque (g mm)	Joint0		-0.633	0.524
	Joint1	71.156	126.279	121.155
	Joint2	3.203	32.299	21.184

가정하여 마찰력은 1.0 으로 선정하고, 물체의 밀도는 모든 부분에서 동일하다고 보았다. 3-손가락 핸드의 각 링크의 길이는 50.00 mm 로 하였으며, 벌칙값은 1.0e+20 으로 설정하였다.

먼저 기초화된 두 대의 카메라를 이용하여 Fig. 21 과 같이 스테레오 영상을 얻었다. 카메라는 물체의 연직 상방에 위치 시켰다. 해상도는 640x480 의 컬러 사진을 획득하였다.

2 장에서 설명한 방법을 이용하여 위의 물체의 3 차원 모델링을 구해보면 Fig. 22 와 같다. 우측 아래의 그림은 모델링 표면을 확대하여 본 삼각형 매쉬의 모습을 보여주고 있다.

위의 모델링의 결과 최상면과 하면에 노이즈들이 많이 있는 것을 볼 수 있다. 이것은 특징점이 없는 부위에서 많이 나타난다. 점 매칭법의 단점으로 볼 수 있다.

위의 모델링의 결과를 앞에서 학습된 신경망에 적용하면 Fig. 23 과 같은 결과를 얻을 수 있었다.



Fig. 20 Cylindrical object used for simulation

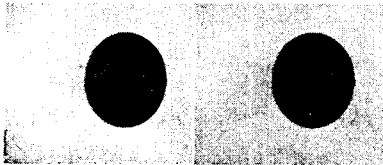


Fig. 21 Calibrated stereo images

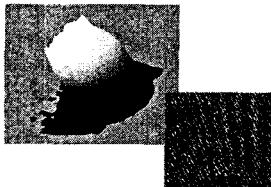
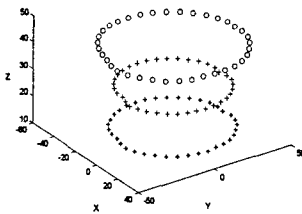
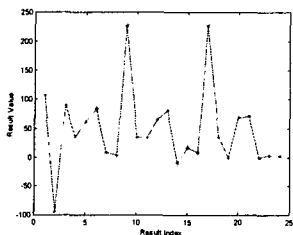


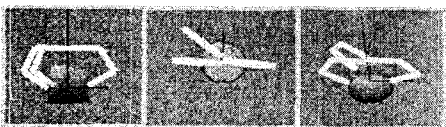
Fig. 22 3D geometrical modeling



(a) Input parameter to find the grasping points



(b) Result value from trained GRNN neural network



(c) Simulation of grasping with the result from neural network

Fig. 23 Input parameters and result for neural network

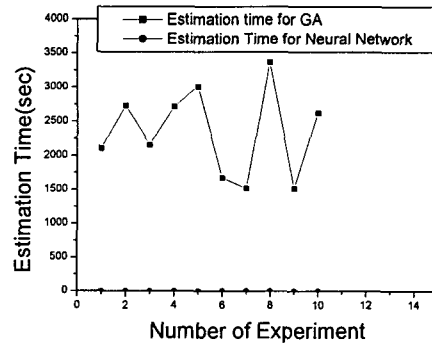


Fig. 24 The compare of the estimation time between using GA method and Neural Network method

Fig. 23(a)와 같이 물체 좌표계에서 입력 파라미터를 구할 수 있었고, Fig. 23(b)에서와 같이 유전자 알고리즘의 결과와 비교하였을 때도 그다지 결과의 차이가 없음을 볼 수 있었다. Fig. 23(b)에서 result index 는 24 개의 출력 파라미터이고, result value 는 각각의 출력 값들을 나타낸다. Fig. 23(c)에서는 최종 결과를 볼 수 있다.

5. 결론

본 연구에서는 스테레오 정합을 이용하여 물체에 대한 정보가 없는 경우 물체의 3-차원 모델링을 구함으로써 물체에 관한 정보를 구할 수 있었고, 3-손가락 핸드를 이용하여 최적의 잡는 점을 구하는 알고리즘을 구할 수 있었다. Fig. 24 는 최적화 방법을 이용했을 경우와 신경망(Neural Network)을 이용했을 경우에 있어서 계산 시간을 비교한 그림이다. 각기 다른 물체 10 개에 대해서 실험을 수행 하였다. 최적화의 경우 30 분에서 40 분 정도의 계산 시간이 소요되지만 신경망(Neural Network)의 경우 1 초나 2 초정도의 계산 시간이 소요 되었다. 따라서 계산 시간이 많이 개선 되었음을 알 수 있다. 또 어떤 물체에 대해서 한번 학습이 되고 나면 후에 다시 그 물체를 봤을 경우 빠른 시간안에 결과를 도출해 낼 수 있다.

하지만 학습을 위한 데이터가 상당히 많이 필요하다는 점과 학습을 위한 시간이 많이 소요된다는 점에서는 단점이 있다. 하지만 실시간 학습이 가능 하여진다면 실제 작업 환경하에서의 적용이 가능하다.

참고문헌

- (1) Nguyen, V.-D., 1988, "Constructing force-closure grasps," *The International Journal of Robotics Research*, 7(3).
- (2) Hauck, A., Rüttinger, J., Sorg, M., and Färber, G., 1999, "Visual Determination of 3D Grasping Points on Unknown Objects with a Binocular Camera System," *In Proceeding of the IEEE/RSJ International Conference on Intelligent robots and System*, pp. 272~278.
- (3) Morales, A., Recattala, G., Pedro J. Sanz, and Angel P. del Pobil, 2001, "Heuristic Vision-Based Computation of Planar Antipodal Grasps on Unknown Objects," *In Proceeding of the IEEE International Conference on Robotic & Automation*, Seoul, Korea.
- (4) Borst, Ch., Fischer, M., and Hirzinger, G., 1999, "A Fast and Robust Grasp Planner for Arbitrary 3D Objects," *In Proc. of IEEE International Conference on Robotics & Automation*, Detroit, Michigan.
- (5) Ferrari, C., and Canny, J., 1992, "Planning Optimal Grasps," *In Proceeding of the IEEE International Conference on Robotics & Automation*, pp. 2290~2295, Nice, France.
- (6) Katada, Y., Svinin, M., Ohkura, K., and Ueda, K., 2001, "Optimization of Stable Grasps by Evolutionary Programming," *In Proceeding of the 32nd ISR*.
- (7) Hyun-Ki Lee, Myun-Hee Kim and Sang-Ryong Lee, 2002, "Optimization of 3D Grasping Points with Whole 3D Modeling for Unknown Object," *In Proc. Of the 2nd China-Korea Joint Workshop on Robotics*, Shenyang, China.
- (8) Fusiello, A., Trucco, E., and Verri, A., 2000, "A Compact Algorithm for Rectification of Stereo Pairs," *Machine Vision and Applications*, 12(1):16~22.
- (9) Laszlo, J., "Computational Geometry and computer graphics in C++," Prentice Hall Press.
- (10) Hyun-Hyup Lee and Kyung-Il Mun, "Fuzzy-Neuro by using a MATLAB," Ajin Press.
- (11) Maybank, S. J., and Faugeras, O., 1992, "A Theory of Self-Calibration of a Moving Camera," *International Journal of computer Vision*, 8(2): 123~151.
- (12) Zhang, Z., 1998, "A Flexible New Technique for Camera Calibration," Technical Report MSRTR-98-71, Microsoft Research.
- (13) Zhang, Z., 1999, "Flexible Camera Calibration by Viewing a Plane from Unknown Orientations," *In Proc 7th International Conference on Computer Vision*, Kerkyra, Greece, pp. 666~673.
- (14) Heikkilä, J., and Silvén, O., 1997, "A Four-step Camera Calibration Procedure with Implicit Image Correction," *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR'97)*, San Juan, Puerto Rico, pp. 1106~1112.
- (15) Kim, G. B., Chung, S. C., 2001, "A Stereo Matching Algorithm with Projective Distortion of Variable Windows," *Transactions of KSME, A*, Vol. 25, No. 3, pp. 461~469.
- (16) Park, K., 1999, "3D Object Recognition and Accurate Pose Calculation Using a Neural Network," *Transactions of KSME, A*, Vol. 23, No. 11, pp. 1929~1939.