

자기조직 신경 회로망을 이용한 스테레오 비전 시스템의
대응문제 해결에 관한 연구

조 영 빈*, 권 대 갑*

A Study on Correspondence Problem of Stereo Vision System using
Self-Organized Neural Network

Y. B. Cho*, D. G. Gweon*

ABSTRACT

In this study, self-organized neural network is used to solve the correspondence problem of the axial stereo image. Edge points are extracted from a pair of stereo images and then the edge points of rear image are assigned to the output nodes of neural network. In the matching process, the two input nodes of neural networks are supplied with the coordinates of the edge point selected randomly from the front image. This input data activate optimal output node and its neighbor nodes whose coordinates are thought to be correspondence point for the present input data, and then their weights are allowed to be updated. After several iterations of updating, the weights whose coordinates represent rear edge point are converged to the coordinates of the correspondence points in the front image. Because of the feature map properties of self-organized neural network, noise-free and smoothed depth data can be achieved.

Key Words : Stereo Vision (스테레오 비전), Correspondence Problem (대응문제), Self-Organized Neural Network (자기조직 신경회로망), Edge Detection (에지추출), 3D Data Extraction (3차원 정보추출)

기호설명

$g(x), g(y)$: 가우시안 저주파 통과 연산자	$I(x, y)$: 원래 영상의 밝기값
$g_x(x), g_y(y)$: Canny 에지 연산자의 근사함수	$I_x(x, y), I_y(x, y)$: 영상 밝기값의 방향 미분
σ	: 백색 가우시안 잡음의 분산	$P_o(x_o, y_o, z_o)$: 실 좌표계
		$P_f(x_f, y_f, z_f)$: 앞 카메라의 좌표계

* 한국과학기술원

$P_b(x_b, y_b, z_b)$: 뒤 카메라의 좌표계	η	: 신경 회로망의 학습율
P_{max}, P_{min}	: 작업공간 매개변수	ε	: 신경 회로망의 입력 벡터
f	: 유효 초점거리	ΔW_{ij}	: 신경회로망의 연결 가중치의 변화
O_i	: 신경회로망의 i 번 출력 단자의 출력값	$\Delta(i, i^*)$: 이웃단자 함수(Neighborhood function)
d	: 에시폴라 선으로 부터의 거리	R_b, T_b	: 뒤 카메라에 대한 앞 카메라 좌표의 회전 행렬과 병진 행렬
S	: 영상의 밝기의 차이		
λ_1, λ_2	: d 와 s 사이의 가중치		
W_{ij}	: 신경 회로망의 연결 가중치		

1. 서 론

스테레오 비전 시스템을 이용한 3차원 거리 정보 추출은 정상적 조명 상태에서 서로 다른 위치에 놓인 2대의 카메라를 사용하는 매우 일반적인 방법이지만 대응 문제의 해결이 가장 큰 어려움으로 남아 있다. 처리 속도의 향상을 위해서 특징점에 대해서만 대응점을 구하는 것이 일반적이므로 좋은 성능의 특징 추출 연산자를 위한 많은 연구가 있어왔다. 이러한 연산자로는 영상내의 한점에서 4방향에 대한 밝기 변화의 표준 편차가 큰 지점을 특징점으로 선택하는 Interest operators⁽¹⁾나 농담 화면에서 밝기의 불연속점을 추출하는 에지 연산자가 있다. Marr-Hildreth⁽²⁾는 2차 미분값의 부호가 변하는 점을 에지로 하기 위해 Laplacian of Gaussian 연산자를 제안하였으며 Canny⁽³⁾는 최적의 1차 미분 연산자를 수치해석적으로 구하였다. Baker, Binford⁽⁴⁾, Kanade⁽⁵⁾는 영상 데이터의 1차 미분을 구하고 스캔 라인 (Scan line) 방향의 최대값 부분을 에지점으로 하였다.

대응은 스테레오 비전에서 가장 중요한 단계로 많은 연구가 진행되어 왔다. 대응점은 첫째 두 카메라의 배치에 의해 결정되는 Epipolar 선상에 존재해야 하며 둘째 특성(Local properties)이 유사하고 셋째 전체적인 제한 조건(Global constraints)을 만족해야 한다. 이 세가지 조건을 만족하는 여러가지 알고리즘이 제시되어 왔는데 카메라 배치는 Parallel, Axial, Verging 스테레오 모델이 있다^{(6) (7) (8)}. 주로 사용되는 특성은 에지이며 제한 조건은 주변 상황을 비교하여 가능한 후보점의 확률을 계속 높여 가는 반복적이고 상호 보완적인 방법이 일반적이다. 대표적인 예로 생물학적 근거를

바탕으로 제시된 Marr-Poggio⁽⁹⁾ 이론이 있으며 Grimson⁽¹⁰⁾에 의해 구현되었다. Mayhew-Frisby⁽¹¹⁾는 3개의 서로 다른 크기의 필터 (3Spatial-frequency tuned channel)를 종합적으로 이용하는 방법을 제시하였고 Burt, Julesz⁽¹²⁾는 Disparity gradient limits을 사용하였다. 이 밖에 Hopfield 신경회로망의 최적화 기법을 이용하여 대응문제를 해결하려는 연구도 있었다⁽¹³⁾.

물체의 3차원 정보를 정확히 추출하기 위해 카메라의 파라미터를 추정하는 많은 연구가 있었는데 그중 Tsai는 렌즈의 왜곡과 CCD 카메라의 Scanning rate까지도 보정하는 효과적인 알고리즘을 제시하였다⁽¹⁵⁾.

본 연구에서는 정상적인 조명상태에서 물체의 3차원 정보를 추출하기 위하여 공통 시야(Common field of view)가 큰 축성(axial) 스테레오 카메라 모델을 사용하였으며 대응점 결정을 위한 특징으로 에지와 그 밝기 값을 사용하였다. 대응점 결정을 위하여 자기 조직 신경회로망을 이용하는 알고리즘을 개발하고 여러가지 대상물에 대한 실험을 통하여 성능을 평가하였다. 또한 이로부터 추출된 거리정보의 정확성을 평가하기 위하여 원뿔 모양의 대상물에 대한 실험도 행하였다.

2. 스테레오 비전 시스템

2.1 스테레오 비전 시스템의 전체 과정

3차원 거리의 측정에는 다음의 각 단계를 차례로 거쳐 이루어진다.

- (1) 영상 획득 (Image acquisition)
- (2) 특징 추출 (Feature extraction)
- (3) 영상 매칭 (Image matching)

- (4) 카메라 모델링 (Camera modeling)
- (5) 거리 결정 (Distance determination)

2.2 영상 획득

영상의 획득은 CCD카메라로 부터 얻어지고 카메라로부터 나온 신호는 프레임 그래버 (Frame grabber)에서 이산화되어 기억된다. 필터링의 경우 작업 시간을 단축시키기 위해 별도의 프레임 프로세서(Frame processor)를 사용하였으며 전체 콘트롤은 IBM PC 80486을 이용하였다. 이 카메라 시스템은 Sony사의 스칼라 타임의 로봇 팔목에 장착되어지고 공압 실린더에 의해 카메라의 축방향으로 움직일 수 있게 되어 있다. 영상처리 시스템의 블록 선도를 Fig.1에 나타내었다.

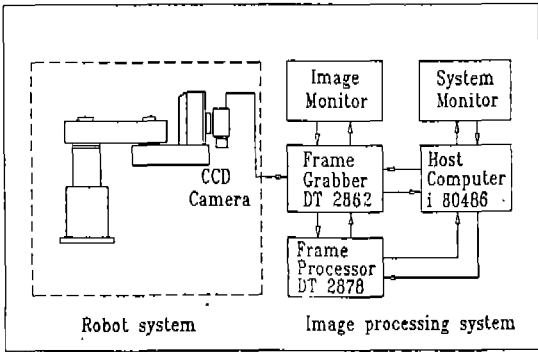


Fig.1 Schematic diagram of the system

2.3 에지의 추출

에지 추출을 위한 가장 기본적인 방법은 1차 미분 연산자의 결과에 적절한 임계값을 이용하는 것인데 적절한 임계값을 찾기가 어려운 단점이 있다. 또다른 방법으로는 2차 미분 연산자인 LoG 연산자를 사용하여 그 결과의 부호 변화 부분을 에지로 하는 것이다. 그러나 2차 미분 연산자는 잡음 성분이 증폭되어 안정적이지 못한 것으로 알려져 있다(14).

미분 연산자는 보통 경험적으로 유도되었는데 Canny는 분산이 σ 인 백색 잡음 (White gaussian noise)이 포함된 영상에서 스텐 에지를 추출하는 최적의 에지 연산자를 수치 해석적으로 구하였다(3). 또한 이 함수가 가우시안 함수로 단순화될 수 있으며 그 성능의 차이가 미소함을 밝혔다.

본 연구에서는 이를 구현하기 위해 다음의 함수로 필터를 디자인하고 $\sigma=1$ 을 사용하였다.

$$g(x) = \exp(-\frac{x^2}{2\sigma^2}) \tag{1}$$

$$g_x(x) = -\frac{x}{\sigma^2} \exp(-\frac{x^2}{2\sigma^2}) \tag{2}$$

영상의 그래디언트 값은 다음과 같이 구하였다.

$$\text{grad } I(x, y) = |I_x \hat{x} + I_y \hat{y}| \tag{3}$$

여기서

$$I_x = g_x(x) * (g(y) * I(x, y)) \tag{4}$$

$$I_y = g_y(y) * (g(x) * I(x, y)) \tag{5}$$

이렇게 구해진 영상으로 부터 에지 방향의 극대점만을 추출하기 위하여 3X3 윈도우를 사용하였다. 이 윈도우내에서 에지 방향으로 존재하는 두 점의 에지 강도를 비교하고 작은 값은 에지에서 제외시킴으로서 선명한 영상을 얻어내었다. 이렇게 구해진 에지는 한개의 픽셀 단위로 이어지게 되고 에지의 강도가 약하거나 강하더라도 균일한 에지 영상의 추출이 가능해지게 된다.

2.4 축성 스테레오 모델

축성 스테레오의 경우는 두 카메라의 축이 서로 일치하며 단지 축 방향의 변위만이 존재하는 카메라 모델로서 공통된 시야가 넓고 은점(occluding point)의 수가 작아서 대응점 문제의 해결이 매우 수월하며 또한 탐색의 범위가 1/2로 줄어들며 카메라의 상대 변위를 늘려도 대응점을 찾기 어려워지는 경향은 발생하지 않는다.

축성 스테레오 모델은 카메라의 상대위치가 축방향으로 ΔZ 만큼 떨어져 있는 것으로 Fig.2에 그 개략도가 나와 있다. 공간성의 한점 $P(x_0, y_0, z_0)$ 은 앞 카메라

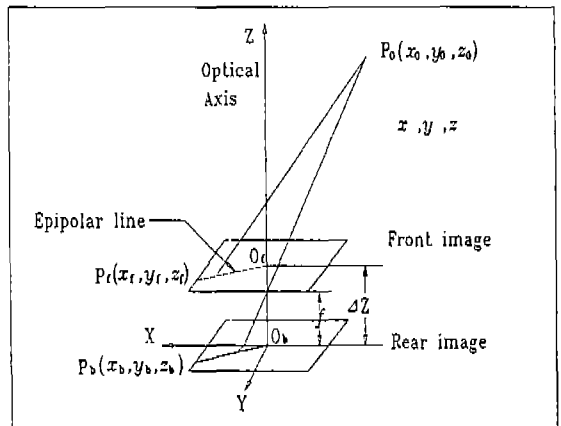


Fig.2 Axial Stereo system

에서 $P_f(x_f, y_f, z_f)$ 점에 뒤 카메라에서는 $P_b(x_b, y_b, z_b)$ 에 상이 맺히게 된다. $O_f P_f$ 와 $O_b P_b$ 두 선은 서로 평행하므로 닳은 삼각형의 조건에 의해 다음의 식이 만족된다.

$$\frac{z_0 - f - \Delta z}{f} = -\frac{x_0}{x_f} = -\frac{y_0}{y_f} \quad (6)$$

$$\frac{z_0 - f}{f} = -\frac{x_0}{x_b} = -\frac{y_0}{y_b} \quad (7)$$

위의 두식 (6), (7)로부터 에피플라 제한 조건은 다음의 식으로 표현되어진다.

$$\frac{y_b}{x_b} = \frac{y_f}{x_f} \quad (8)$$

식 (8)로부터 에피플라 선이 중심을 향하고 있음을 알 수 있다. 작업 공간을 다음과 같이 나타내면

$$P_{\min} \cdot \Delta z + f < z_0 < P_{\max} \cdot \Delta z + f \quad (9)$$

앞의 식 (6), (7), (9) 으로부터 앞면 영상점 (x_f, y_f) 의 대응점 (x_b, y_b) 은 다음 식과 같은 범위내에만 존재한다는 것을 알 수 있다.

$$\frac{P_{\min} - 1}{P_{\min}} |x_f| < |x_b| < \frac{P_{\max} - 1}{P_{\max}} |x_f| \quad (10)$$

$$\frac{P_{\min} - 1}{P_{\min}} |y_f| < |y_b| < \frac{P_{\max} - 1}{P_{\max}} |y_f| \quad (11)$$

이 식으로부터 대응점 결정을 위한 탐색의 범위가 작업공간과 카메라의 상대적 위치에 의해 결정됨을 알 수 있다. 실제 이 범위는 조립 시스템을 설계할 때 결정할 수 있으므로 효율적인 탐색을 위해 이용된다.

3. 대응점 결정 알고리즘

3.1 자기조직 신경 회로망의 패턴인식

입력은 있으나 요구되는 출력이 어떤 것인지 정해지지 않는 신경 회로망을 자율 신경회로망(Unsupervised neural network)이라 하는데, 이런 신경 회로망은 입력 데이터의 패턴과 특징을 분류하고 입력의 상관 관계를 정의하는 기능을 가진다. 특히 자기조직 신경 회로망의 경우에는 출력이 입력의 기하학적 사영(Topological Mapping)이 되는 피쳐 맵핑(Feature mapping)기능을 가지는데 비슷한 입력 패턴이 비슷한 출력 단자를 자극하는 것으로 이런 기능이 만들어진다.

이 경우 반드시 입력 데이터에 리던던시(Redundancy)가 있어야만 한다. 이를 학습시키기 위해서 연결 가중치(Connection weight)를 벡스칸 모자 형태로 하거나 선택된 출력단자와 이웃 단자도 같이 학습을 시키는 코호넨(Kohonen) 알고리즘을 사용하기도 한다.

3.2 자기조직 신경회로망의 스테레오 매칭

자기조직 신경회로망을 대응 문제에 사용하려면 입력 데이터의 수가 출력 단자의 수보다 더 많아 리던던시가 있어야 된다. 축성 스테레오 비전의 경우 앞에서 본 영상은 크고 뒤에서 본 영상은 작으므로 뒤 영상의 에지점을 자기조직 신경회로망의 출력 단자로 하고 앞 영상의 에지좌표를 입력으로 한다면 입력 데이터의 리던던시가 보장된다. Fig.3에서 보는 바와 같이 뒤 카메라 영상으로 부터 에지를 구하고 이 에지점들을 자기조직 신경회로망의 각 출력 단자에 대응시킬 경우 출력 단자의 배열은 뒤 카메라의 영상과 같게 된다. 그리고 앞 카메라 영상의 에지점을 무작위로 추출하여 각각의 좌표값을 2개의 입력단자에 대입한다. 초기에 출력 단자에 연결된 웨이트는 출력 단자의 좌표값인 뒤 카메라 영상의 에지좌표와 같은 값을 가지고 있으나 입력 단자로부터 앞 카메라의 영상 좌표가 계속해서 입력됨에 따라 최적의 대응점으로 수렴해 나가게 된다. 대응점이 될 수 있는 에지의 후보점은 반드시 에피플라 선상에만 존재하므로 이 선 주변에 있는 출력 단자만이 선택되어진다. 또한 식 (10), (11)로부터 탐색 범위가 정해지며 에지점의 밝기값도 출력 단자를 결정할 때 고려 대상이 된다. 이러한 선택 조건에 의해 주어진 입력 좌표에 대한 조적의 출력 단자 i^* 는 다음의 식과 같이 선택된다.

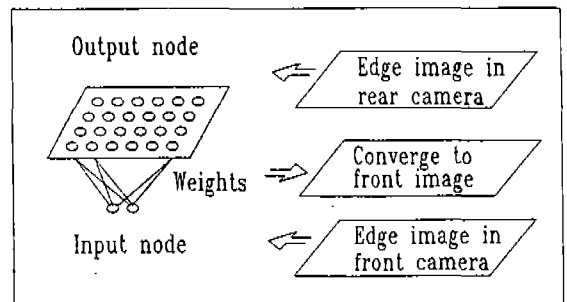


Fig.3 Self-organized neural network used in this vision system

$$O_i^* \geq O_i \quad \text{for all } i \quad (12)$$

여기서

$$O_i = \frac{1}{1 + \lambda_1 d^2 + \lambda_2 s^2}$$

$$s : s = \left(\frac{I_b - I_f}{I_b} \right)^2$$

또한 λ_1 과 λ_2 는 에피폴라 제한조건과 밝기값의 차이의 영향을 조정하는 변수이다. 주어진 입력에 대해 출력값이 가장 큰 단자(i^*)가 선택된 후 그 주위의 단자와 함께 웨이트가 수정받을 학습 법칙은 다음과 같다.

$$\Delta w_{ij} = \eta \cdot \Lambda(i, i^*) \cdot (\xi_j - w_{ij}) \quad (13)$$

여기서

η : 학습율

$\Lambda(i, i^*)$: 이웃단자 함수(Neighborhood function)

$$\Lambda(i, i^*) = \exp\left(-\frac{|r_{i^*} - r_i|}{2\sigma^2}\right) \quad (14)$$

여기서 이웃값의 학습 정도를 나타내는 식 (14)의 유클리디안 거리는 출력 단자 사이의 거리이다. 위의 학습 법칙은 대응 문제의 해결을 위해 전체적 대응 제한 조건(Global matching constraints)으로 평활 제한

조건(Smoothness constraints)를 주던 여러 알고리즘과도 일치하는 것이다. 이러한 대응과정은 출력 단자에 연결된 웨이트들의 값이 대응하는 입력값에 수렴할 때까지 반복된다. 결국 대응과정이 완료된 후 뒤 카메라의 영상점(출력 노드)은 이에 상응하는 앞 카메라의 영상점(출력 노드에 연결된 웨이트)과 대응되어 진다. 이러한 대응 과정의 흐름도를 Fig. 4에 나타내었다.

4. 카메라 보정 및 좌표변환

(x_f, y_f) 와 (x_b, y_b) 사이의 대응관계가 결정된 후 거리 정보는 식 (6)과 (7)로부터 쉽게 구해질 수 있다. 그러나 두 카메라의 상대 위치는 Fig. 2에서 보는 바와 같이 이상적이 아니므로 축 방향의 ΔZ 성분 이외도 다른 축 방향으로의 병진 변위와 회전 변위가 존재한다. 또한 렌즈와 프레임 그래버에 의한 왜곡도 정확한 거리 정보 추출을 위해서 고려되어야 한다. 본 연구에서는 Tsai⁽¹⁵⁾에 의해 제안된 방법에 의해 카메라를 보정하였다. 공간상에 놓여진 기준 좌표계에 대한 앞, 뒤 두 카메라의 회전 행렬과 병진 행렬을 구한 후⁽¹⁵⁾ 두 대응점 사이의 관계는 다음과 같이 구해진다.

$$P_f = R_b^f \cdot P_b + T_b^f \quad (15)$$

여기서 첨자 f 는 앞 영상을 의미하고 b 는 뒤 영상을 의미한다. 만일 대응점 (x_f, y_f) 와 (x_b, y_b) 가 주어지면, P_f, P_b 다음과 같이 표현된다.

$$P_f = \left(\frac{x_f}{f} z_f, \frac{y_f}{f} z_f, z_f \right)^T = \vec{a} \cdot z_f$$

$$P_b = \left(\frac{x_b}{f} z_b, \frac{y_b}{f} z_b, z_b \right)^T = \vec{b} \cdot z_b \quad (16)$$

여기서

$$\vec{a} = \left[\frac{x_f}{f}, \frac{y_f}{f}, 1 \right]^T$$

$$\vec{b} = \left[\frac{x_b}{f}, \frac{y_b}{f}, 1 \right]^T$$

위 식에서 대응점 (x_f, y_f) 와 (x_b, y_b) 는 신경회로망을 이용한 대응점 결정과정에서 구해지고 f, R_b, T_b 는 카메라 보정 과정에서 구해지므로 벡터 a 와 b 는 기지의 값이고 z_f 와 z_b 만이 미지의 값이 된다. 식 (15)와 (16)을 정리하면 다음과 같은 식을 구할 수 있다.

$$\begin{bmatrix} z_f \\ z_b \end{bmatrix} = \begin{bmatrix} \vec{a} & -R_b^f \cdot \vec{b} \end{bmatrix}^{-1} \cdot T_b^f \quad (17)$$

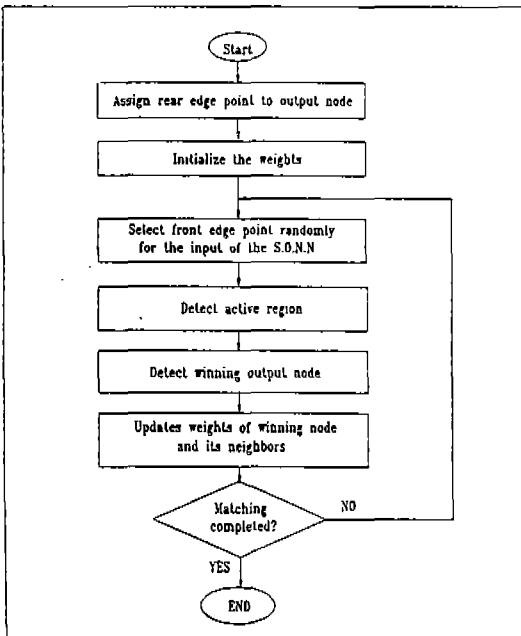


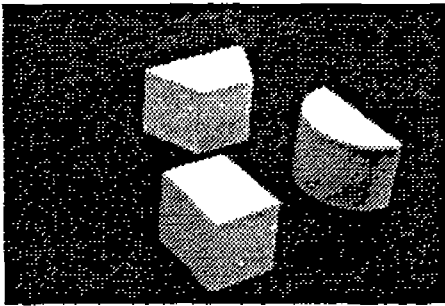
Fig. 4 Flow chart of matching algorithm

이 식으로부터 거리 정보의 정확한 계산이 가능해지게 된다.

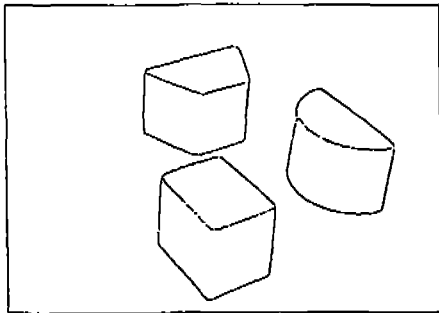
5. 실 험

5.1 특징 추출의 결과

본 연구에서는 특징점으로 에지를 추출하였다. 이 실험에서는 단순한 모양의 대상물에 대해서 에지를 추출하였는데 에지 방향으로 국소 최대점만을 에지로 하는 알고리즘에 의해 구한 에지영상을 Fig. 5에 나타내었다. 에지 추출에 임계값을 사용하면 너무 굵고 혹은 끊어져서 에지가 나타나지만 에지 변화 방향의 극대점만을 에지로 하는 알고리즘은 에지가 약한 부분이나 강한 부분이 모두 매끈하게 한줄로 나타나게 할 수 있었다.



(1) Original image



(2) Edge image

Fig.5 The results of edge detection algorithm

5.2 카메라 보정 결과

Tsai⁽¹⁵⁾에 의해 제안된 카메라 보정 방법을 사용하기 위해 30개의 점을 가진 높이 45mm, 윗면 160

×100mm인 금속 블럭 2개를 이용하였다. 보정점 (Calibration point)은 직경이 10mm인 원으로서 2진 영상으로 받아들인 후 이 점의 중심을 구함으로써 잡음의 영향을 최소화하였다. 우선 렌즈의 왜곡을 무시하여 대략적인 매개변수를 추정하고 렌즈의 왜곡을 고려하여 모든 보정점에 대한 최소 자승 에러를 구하는 방법으로 모든 매개변수를 구하였다. 에러는 카메라 보정을 통해 구해진 매개변수를 사용하여 시편상의 보정점이 카메라의 영상내의 좌표로 역변환 되었을 경우 실제의 영상 좌표와의 평균 에러로 정의된다⁽¹⁵⁾. 카메라 보정과정 중 에러의 변화를 Fig. 6에 나타내었다. 영상내의 모든 좌표에서 0.3픽셀 이하로 에러를 줄일 수 있으므로 카메라의 왜곡에 의한 모든 에러를 수정할 수 있다.

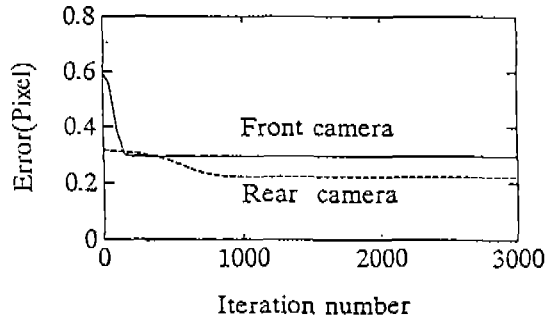
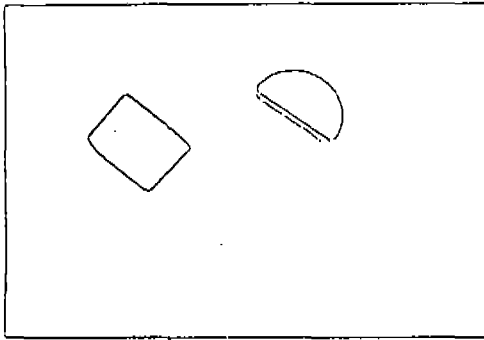


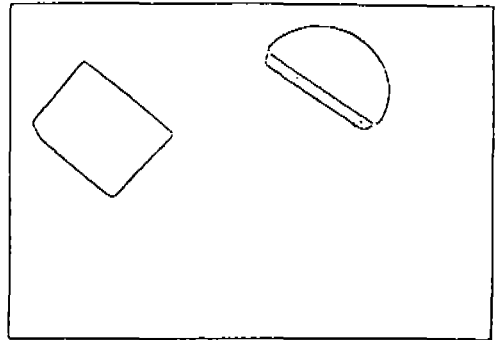
Fig.6 Error Minimization in Camera Calibration

5.3 신경 회로망의 성능 평가

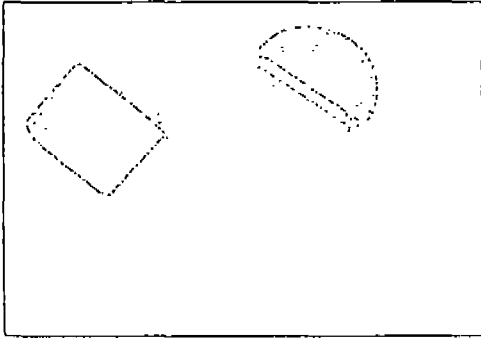
자기 조직 신경회로망의 성능을 평가하기 위해서 우선 육면체와 반원기둥같은 간단한 대상물의 대응문제에 대해 실험하였다. Fig. 7의 첫번째 그림은 뒤 카메라의 에지 영상이고 두번째 그림은 앞 카메라의 에지 영상이다. 그 다음은 앞 카메라에서 추출된 에지 영상의 좌표가 신경 회로망의 2개의 입력 단자에 들어감에 따라 출력 단자와 연결된 연결 가중치가 대응점으로 수렴해 가는 과정을 보여준다. 이 예와 같이 간단한 영상에 대해서는 5번 정도의 반복 입력만으로도 충분히 대응점에 수렴하는 것을 알 수 있었다. 다음은 천공 편지, 사무용 칼, 그리고 볼펜이 어지럽게 널려있는 조금 복잡한 영상에 대해 실험하였다. Fig. 8에서 보는 바와 같이 몇번의 반복 입력에 의해 대응문제가 잘 해결되고 있음을 알 수 있다. 이 경우는 10여번의 반복 입력후에야 대응점에 수렴하였다.



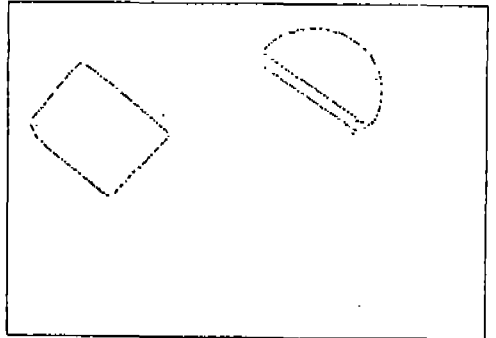
(1) Rear image



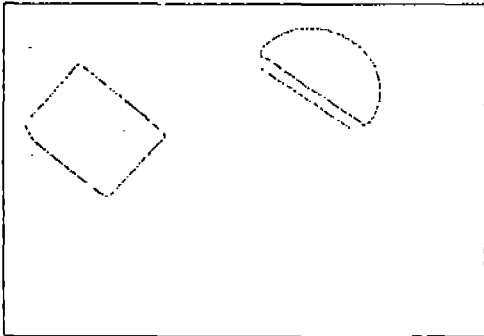
(2) Front image



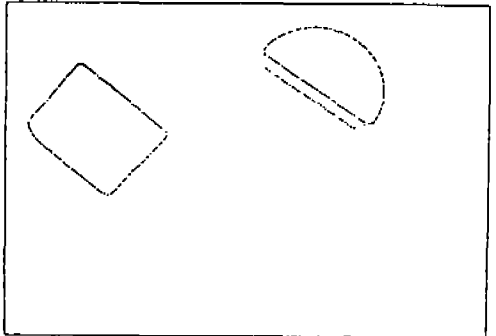
(3) Weight image (1st iteration)



(4) Weight image (2nd iteration)

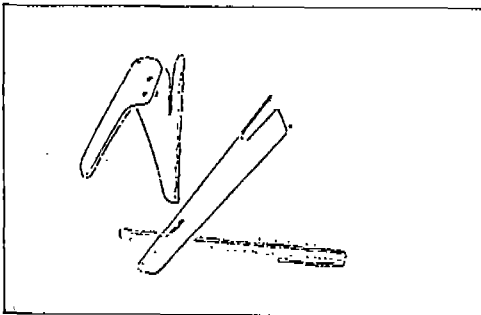


(5) Weight image (5th iteration)

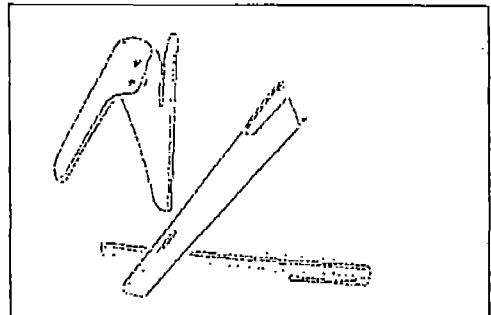


(6) Weight image (10th iteration)

Fig.7 Matching process of self-organized neural network for simple object



(1) Rear image



(2) Front image

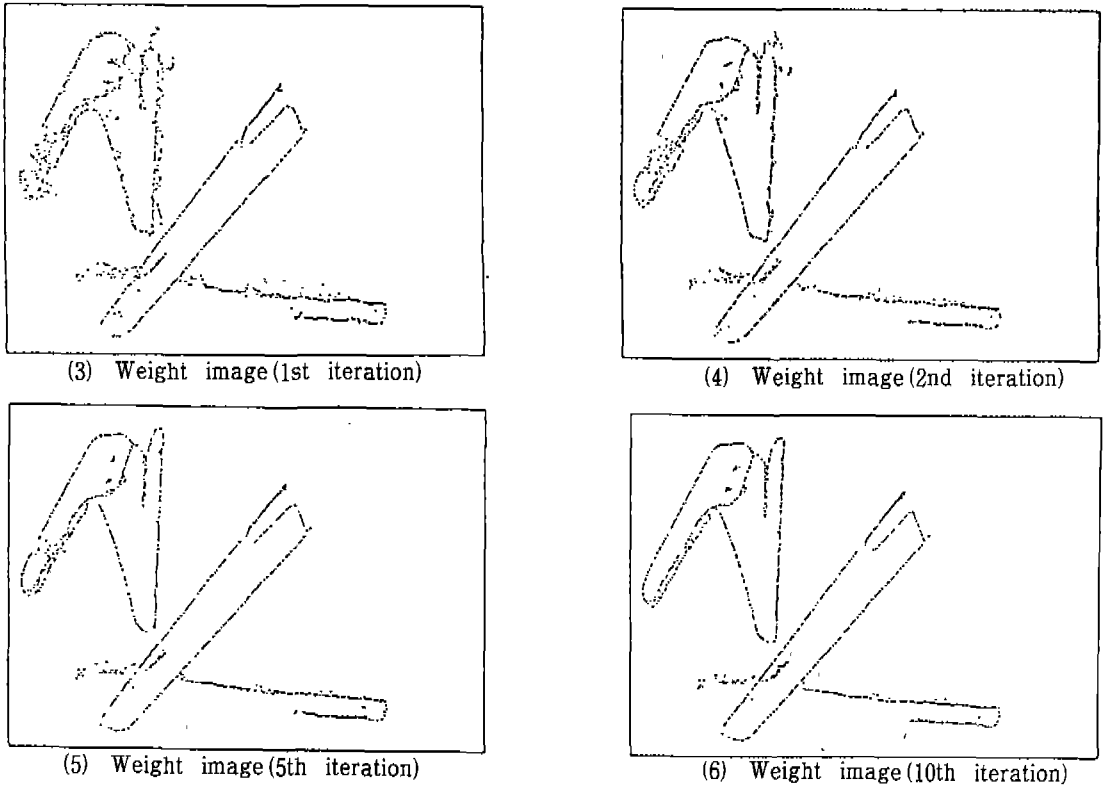


Fig. 8 Matching process of self-organized neural network for complex object

5.4 3차원 거리 정보 추출의 결과

3차원 거리 정보의 정확성을 실험하기 위해서 Fig. 7에서 사용된 반원기둥의 경우를 다시 조사하였다. 테이블면의 z축 좌표는 0이고 z축의 방향은 테이블의 윗면을 향하도록 좌표계를 설정하였다. 따라서 반원 기둥의 밑면은 거리가 0이어야 하고 윗면은 반원기둥의 높이인 48mm이어야 한다. 대응이 끝난 후 거리 정보를 추출한 결과를 Fig. 9에 나타내었다. 또한 거리 정보의 정밀도를 정량적으로 보기 위하여 Fig. 10와 같이 일정 높이에만 에지가 나타나는 모델을 가지고 실험하였다. 모델은 원뿔 모양이며 바닥으로부터 27mm와 45mm 사이에 점을 가지고 있어서 에지를 추출할 경우 0mm, 27mm, 45mm 높이에서만 에지점이 존재하게 된다. Fig. 10에서 보는 바와 같이 추출된 거리 정보가 실제 모델의 에지 높이 부근에 분포하는 것을 알 수 있는데 거리 정보의 에러는 작업 물체의 거리에 비해 약 2~3% 이내에 있어서 정확성이 높은 것을 알 수 있다. 이것은 신경 회로망의 에러에 대한 둔감성의 영향이기도

하고 에지의 대응이 subpixel 정밀도로 얻어지기 때문이다.

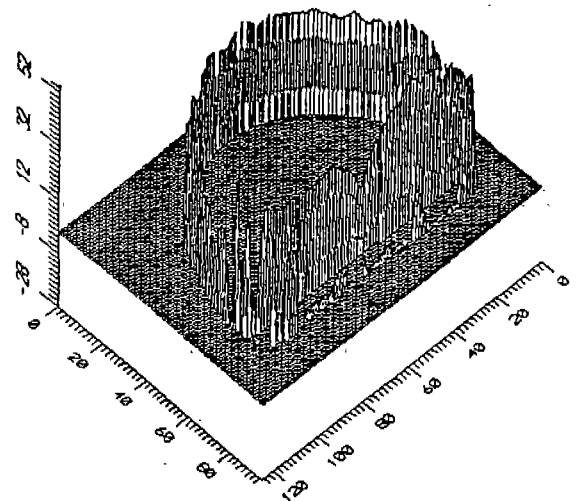


Fig. 9 Depth display in case of half cylinder used in fig. 7

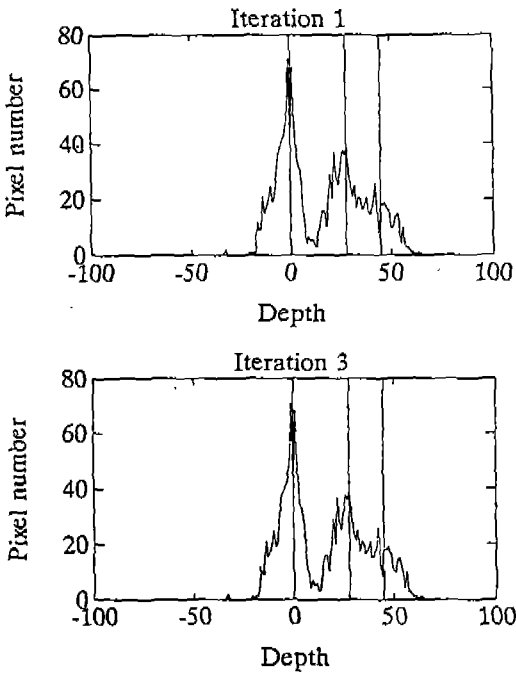
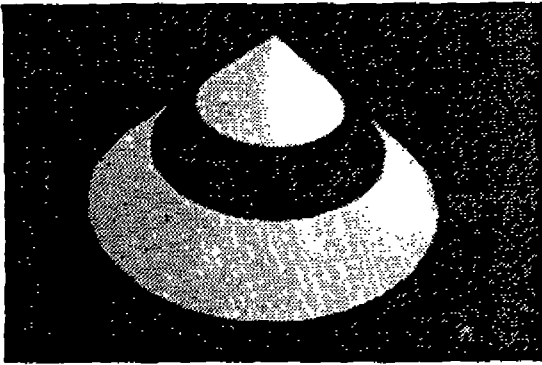


Fig. 10 Calculated depth distribution of cone image

6. 결 론

본 연구에서는 3차원 거리 정보를 추출하기 위하여 자기조직 신경회로망을 이용한 대응점 결정 알고리즘을 개발하고 여러 대상물에 대하여 성능을 실험하였다. 자기조직 신경회로망은 축성 스테레오 시스템의 특성상 앞 영상이 가지고 있는 데이터의 리던던시를 활용할 수

있음에 기초를 두고 있는데 이로 인해서 잡음에 둔감한 대응 결과를 얻을 수 있었다. 여러가지 대상물에 대하여 이 알고리즘의 대응성능을 평가하였는데 단순한 대상물에 대해서는 2, 3회의 입력으로 신경 회로망이 대응점에 수렴하였으며 복잡한 대상물의 경우에는 10회 정도의 입력으로 대응점에 수렴함을 알 수 있었다. 구해진 대응 결과로부터 거리 정보를 추출하고 그 정밀도를 해석하였는데 거리정보에 대한 에러가 카메라와 대상물 사이의 거리에 비해 2~3% 이내에 있음을 확인하였다.

또한 본 연구에서는 에지의 추출을 위해 3X3 윈도우를 사용하는 알고리즘을 제시하고 그 성능을 실험하였는데 그 결과 매우 효과적으로 에지를 찾아낼 수 있었다. 실제의 카메라 시스템은 이상적이 아니므로 이것을 보정하기 위한 매개변수의 추정은 3차원 거리계산에서 매우 중요한 역할을 하게 되는데 본 연구에서 사용된 카메라 보정 방법은 이러한 매개변수를 찾는 데 매우 효과적이었다.

전체 작업 시간에 대한 결과를 Table 1에 나타내었다. 에지를 찾기 위한 컨볼루션에 대부분의 시간이 소요되므로 이 시간을 줄이는 것이 실제 시스템을 구축하는데 중요하리라 판단된다.

Table 1. Processing time

Task	Processing time
Edge Extraction	6.0 sec/image
Scene radiance	0.3 sec
Matching process	(0.5 - 1.5) sec/iteration
Depth calculation	0.5 sec
System overhead	1 sec
Total	13 - 18 sec

참고문헌

1. H. P. Moravec, 1977, "Towards automatic visual obstacle avoidance" Proc. 5th int AL, p. 584
2. D. Marr, Hildreth, 1978, "Theory of edge detection", proc. Royal Soc. London, Vol. B207, pp. 223~229

3. J. F. Canny, 1985, "A computational approach to edge detection", IEEE PAMI, Vol. 8, No. 6, pp.679~698
4. H. H. Baker and T. O. Binford, 1981, "Depth from edge and intensity based stereo", in Proc. 7th Joint Conf. AI Vancouver, Canada, pp.631~636
5. Y. Ohta and T. Kanade, 1985, "Stereo by intra- and inter-scanline search" IEEE Trans. PAMI Vol. 7, No. 2, pp.139~154
6. Nicolas Alvertos, 1989, "Camera geometries for image matching in 3D machine vision", IEEE, PSMI, Vol. 11, No. 9, pp.897~915
7. Eric Krotkov, knud Henriksen, 1990, "Stereo ranging with camera", IEEE, PAMI, vol. 12, NO. 12, pp.1200~1205
8. David vermon, 1990, "Using camera motion to estimate range for robotic parts manipulation", IEEE RA. Vol. 6, No. 5, 1990, pp.509~521
9. D. Marr, T. Poggio, 1979, "A computational theory of human stereo vision", Proc. Royal. Soc. London Vol. B204, pp.301~328
10. Grimson, 1981, "A computational experimnents with a feature based stereo algorithm", IEEE, PAMI, Vol. 7, pp.17~34
11. Mayhew, Frisby, 1981, "Psychological and computational studies bowards a theory of human stereopsis", AI, Vol. 17, pp.349~385
12. Peter Burt, Bela Julesz, 1980, "A disparity gradient limits for binocular fusion", SCIENCE, Vol. 208, pp.615~617
13. Nasser M. Nasrabadi, 1992, "Hopfield Network for Stereo Vision Correspondence", IEEE, Tran, Neural Networks. Vol. 3, No. 1, pp.5~13
14. E. De Micheli, 1989, "Localization and Noise in edge detection", IEEE, PAMI, Vol. 11, No. 10, pp.1106~1117
15. Roger Y. Tsai, 1986, "An efficient camera calibration techniques for 3D Machine vision", IEEE, Proc. Computer vision and Pattern recognition, pp.364~374
16. Jhon hertz, Anders Krough, Richard G. palmer, "Introduction to the theory of neural computation", Addison-Wesley Publishing company