

신경회로망을 이용한 카메라 보정기법 개발

Development of Camera Calibration Technique Using Neural-Network

°한성현*, 왕한홍**, 장영희**

*경남대학교 기계설계학과 (TEL : +82-551-49-2624; FAX : +82-551-43-8133; E-mail : shhan@kyungnam.ac.kr)

**경남대학교 기계설계학과 대학원 (TEL : +82-551-49-2617; FAX : +82-551-43-8133)

Abstracts This paper describes the camera calibration based-neural network with a camera modeling that accounts for major sources of camera distortion, namely, radial, decentering, and thin prism distortion. Radial distortion causes an inward or outward displacement of a given image point from its ideal location. Actual optical systems are subject to various degrees of decentering, that is, the optical centers of lens elements are not strictly collinear. Thin prism distortion arises from imperfection in lens design and manufacturing as well as camera assembly. It is our purpose to develop the vision system for the pattern recognition and the automatic test of parts and to apply the line of manufacturing. The performance of proposed camera calibration is illustrated by simulation and experiment.

Keywords Neural Network, Geometric distortion, Recognition, Camera calibration, DSP, SCARA Robot

1. 서론

비전을 이용한 부품 형상인식 및 자동검사 시스템은 vision system 본체, 조명광학부, 시각센서부, display monitor, 데이터 입력장치 등으로 크게 나눌 수 있다. 조명광학부는 안정된 영상 취득을 제공하기 위해 적용분야에 적합한 조명방식 및 장치를 사용하며, 시각센서부는 일반적으로는 흑백 CCD camera를 사용하지만 필요성에 따라 color, 적외선, X-ray, 초음파 등의 센서를 쓰기도 한다. 데이터 입력장치는 keyboard, mouse, track ball 등으로 vision system의 기능을 제어하며, display monitor는 영상(처리전/후), 입력된 데이터, 처리결과 등이 표시된다. vision system 본체는 기능상, 시각센서로부터 디지털 영상을 얻고, 저장, 표시하는 영상취득부, 영상메모리부, 영상표시부, 시각검사 알고리즘을 수행하는 고속영상처리부, 주변제어기기 및 데이터 I/O 등을 제어하는 I/O interface부, 그리고 시스템 제어부 등으로 구성된다. 이러한 구성요소들로 이루어진 비전 시스템이 산업현장에서 성공적으로 적용되기 위해서는 다음과 같은 조건들을 만족해야 한다. 1) 고속성 : 동작속도가 생산 라인의 속도에 맞출 수 있어야 한다. 2) 고신뢰성 : 사용사가 원하는 정확도를 일관되게 유지해야 한다. 3) 유연성(flexibility) : 제품 모델의 변경에 대한 적응성과 생산 방식의 구조차이에 대한 원활한 적응성이 뛰어나야 한다. 4) 저렴성 : 생산비용면에서 타당성이 인정될 수 있어야 한다.

본 연구에서는 이러한 기초 연구를 바탕으로 먼저 부품 검사 및 장착 시스템으로 INTEL 80586 CPU의 SPC8200P 컴퓨터에 영상 취득부, 텍사스 인스트루먼트사의 DSP인 TMS320C30을 이용한 병렬처리부와 실시간으로 중심값과 히스토그램을 추출하는 전용하드웨어를 통한 비전 시스템을 개발하여 SM5 SCARA 로봇 시스템과의 인터페이스를 통한 부품의 자동검사 및 형상인식 시스템 개발을 위한 연구를 수행하고자 한다.

본 연구에서는 16.7 MIPS의 초고속 명령 수행 능력을 가진 텍사스 인스트루먼트사의 DSP인 TMS320C30칩을 CPU로 사용

하는 영상처리 하드웨어를 설계 제작하고, 이를 활용하여 여러 제품들중 특정 제품을 자동인식하고 인식된 제품의 품질을 자동으로 검사하여 품질기준 통과여부를 자동 판단할 수 있는, 이른바 부품의 형상인식 및 자동검사용 비전시스템을 개발하여 실제 생산공정에 적용함으로써 제품의 품질향상, 생산품의 균일화 및 제품 생산성을 향상시킴으로써 생산 원가를 절감하고 기업의 기술 경쟁력을 향상시키고자 하는데 연구의 목적이 있다.

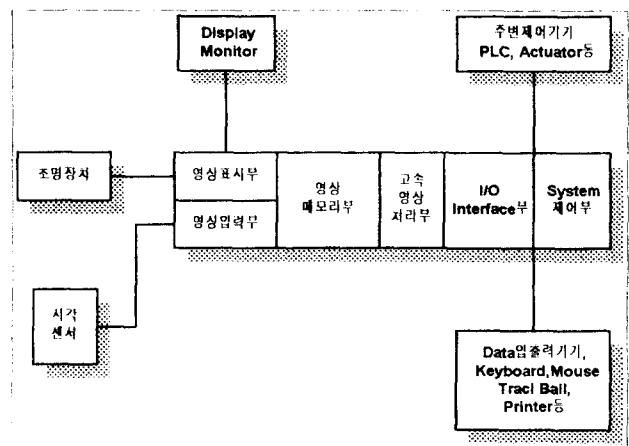


그림 1 로봇 비전 시스템의 일반 구성

Fig. 1 The general configuration of robot vision system.

신경회로망을 이용한 보정기법은 종래의 보정기법에서 볼 수 없었던 병렬형 및 학습 기능면에서 그 우월성을 나타낸다고 할 수 있다. 신경회로망 모델중 가장 많이 적용되는 모델은 다층 신경망이다. 다층 신경망은 다수의 뉴런과 이들간의 연결강도를 나타내는 가중치로 구성된다. 각 뉴런은 외부로부터의 입력에 가중치를 곱하고 이를 모두 합한 다음 시그모이드 활성화함수를 통과시키는 구조를 가지고 있다. 이러한 다층 신경망에서 원하는 기능을 수행하기 위한 적절한 가중치를 구하는 방법으로 오차 역전

키는 구조를 가지고 있다. 이러한 다층 신경망에서 원하는 기능을 수행하기 위한 적절한 가중치를 구하는 방법으로 오차 역전파 알고리즘을 많이 이용한다. 이것은 다층구조에서의 학습법으로써 오차역전파학습이 가장 효율적이라는 것이다. 본 연구에서는 신경회로망을 이용한 카메라 보정기법의 개발에 관한 연구가 수행되었다. 그 진행방법은 우선 3차원 실제계 좌표로부터 2차원 영상 좌표를 구하는 카메라 보정에 대해 시험하고 그 결과를 2단계 보정 방법과 비교하여 그 성능이 검증된다.

2. 신경회로망을 이용한 보정기법

인공적인 신경망 모델은 단순한 신경망, 또는 연결자 모델 또는 병렬 분산 처리 모델 등으로 불리운다. 이 모든 이름은 간단한 계산요소를 밀도있게 연결하므로써 좋은 성능을 얻고자 하는데 그 목적이 있다. 계산요소 또는 신경망 노드의 작동함수는 비선형적이며, 대체로 아날로그식이다. 노드의 출력결정은 내부의 역치(threshold 또는 offset) θ 와 비선형 함수의 유형에 의하여 결정된다. 신경망 모델은 회로망의 모양, 노드의 계산특성, 훈련규칙(training rule) 또는 학습규칙(learning rule)에 의해서도 결정된다. 이러한 규칙은 가중치의 초기화, 성능향상을 위한 계산도중의 가중치 변경 등에 관하여 규정한다.

신경망 모델의 장점은 고도의 병렬처리에 의한 계산성능의 향상에만 있는 것이 아니다. 신경망 모델은 우수한 견실성(robustness)와 오류 내인성(fault tolerance)을 보여준다. 몇 개의 입력 노드가 여러개의 처리 노드에 국소적으로 연결되어 있어 한두개의 오류는 전체 성능에 별 영향을 주지 않고, 전단계의 출력결과에 따라 가중치를 변경할 수 있어서 사물에 따라 적응하여 연속적인 학습이 가능하다. 적응능력은 견실성에도 영향을 주는데, 계산요소의 특성이 미소하게 변화하여도 이를 보정할 수 있다.

이미지상에서 보정을 위해 취득된 샘플점들과 그것들의 상응되는 공간상의 점들과 상호 관계에서 적절한 보정을 하기 위하여 다층 구조 퍼셉트론 형태의 신경망(Multilayer Perceptron type Neural Network)[1]을 고려하였다. 다층 구조 퍼셉트론 형태의 신경망은 한 개의 입력층(input layer)과 출력층(output layer) 그리고 한 개 이상의 은닉층(hidden layer)으로 구성된다. 또한 각 층들은 한 개 이상의 뉴런을 가지고 있으며 각 뉴런들은 인접한 층에 있는 모든 뉴런과 각기 다른 weight를 가지고 서로 연결되어 있다. 각각 뉴런은 상호 연결들을 통해 은닉층에 있는 뉴런들로부터 자신의 입력을 받고, 그 다음층에 있는 뉴런들에 대해 자신의 결과를 전달한다. 입력층의 뉴런을 제외한 이전층의 모든 뉴런들의 출력을 일정하게 가중된 합(a linearly weighted sum)인 신호를 받는다. 다층 퍼셉트론의 성능은 노드 내에서 사용하는 비선형함수에서 나온다.

h 개의 은닉층이 존재한다고 가정하면, 층0과 층($h+1$)은 각각 입력층과 출력층을 나타낸다고 할 수 있다. 그러므로 층 k 에 있는 뉴런 j 의 역할을 다음과 같이 정의된다.

$$\mu_j^{(k)} = h_j \left(\sum_i W_{ij} \cdot v_i^{(k-1)} \right) \quad (1)$$

$$\text{단, } h_j(x) = \begin{cases} \frac{1}{1+e^{-x}} & (\text{뉴런 } j \neq \text{은닉뉴런}) \\ x & (\text{뉴런 } j = \text{은닉뉴런}) \end{cases}$$

그리고 I 는 층 ($k-1$)에 있는 모든 뉴런을 말한다. 식(1)에 있는

$\mu_j^{(k)}$ 에서, k 번째 층에 있는 뉴런들과 j 번째 뉴런의 역할은 ($k-1$)번째 층에 있는 뉴런들과 k 번째 층에 있는 j 번째 뉴런을 연결하는 weight들과 ($k-1$)번째 층에 있는 뉴런들의 역할 함수이다. 비선형 형태의 함수 $h_j(x)$ 는 반드시 sigmoid 함수 ($\frac{1}{1+e^{-x}}$)여야 하는 것은 아니다. x 영역에서 미분가능한 다른 단조로운 함수(monotonic func.)도 $h_j(x)$ 로 사용될 수 있다. 의사 결정 규칙은 최대 출력치를 가지는 출력노드에 대응하는 부류를 선택한다. 이러한 신경망의 의사결정 영역은 다면체가 아닌 부드럽게 변화하는 곡선으로 둘러쌓여 있기 때문에 그러한 신경망의 동작을 분석하는 것은 더욱 복잡한 일이다. 그러나 이러한 신경망은 역전파 알고리즘(back propagation training algorithm)에 의해 훈련된다.

역전파 학습 알고리즘은 이 연구에서 신경망을 학습시키기 위해서 사용하였다. 이 알고리즘은 입력인자들에 대한 목표값이 필요하기 때문에 감독 학습 알고리즘(supervised learning algorithm)으로 분류된다. 이 알고리즘의 기본적인 개념은 서로 연결된 weight의 가장 좋은 계산값들을 얻기 위해, 그리고 주어진 입력에 대해 가능한 목표값에 가까운 신경망의 출력을 구하기 위하여 gradient descent 알고리즘을 이용한 것이다. 그리고 동작함수는 미분가능한 비선형 연속함수이어야 한다. 요망하는 출력치는 입력패턴과 같은 부류의 것에 대응되는 출력노드이던 한자가(0 또는 <0.1)이다. 입력패턴에 대응되는 출력노드이면 고가(1.0 또는 >0.9)이다. 역전파알고리즘을 단계별로 보면, 1단계로 가능치와 옵션치의 초기화하고 2단계로 입력과 요망하는 출력을 제시한다. 3단계로는 실제 출력치를 산출하고, 4단계에서 가중치를 적용시킨다. 그 다음으로 다시 2단계로 보내 이 과정을 반복한다. 이 알고리즘의 핵심적인 부분은 가중치 조절에 필요한 오류를 출력층으로부터 입력층으로 역전파시키는 반복하는 것이다. 이 신경망은 처음에는 작은 무작위적 가중치와 내부 역치를 선택하고 다음에 훈련자료를 반복적으로 제시하므로써 훈련된다. 가중치를 조정하는 방법은 정확한 부류를 지정하는 부차적 정보를 사용하여 시도하고 가중치가 수렴하고 비용 함수의 값이 인정될 수 있을 때까지 반복하는 것이다.

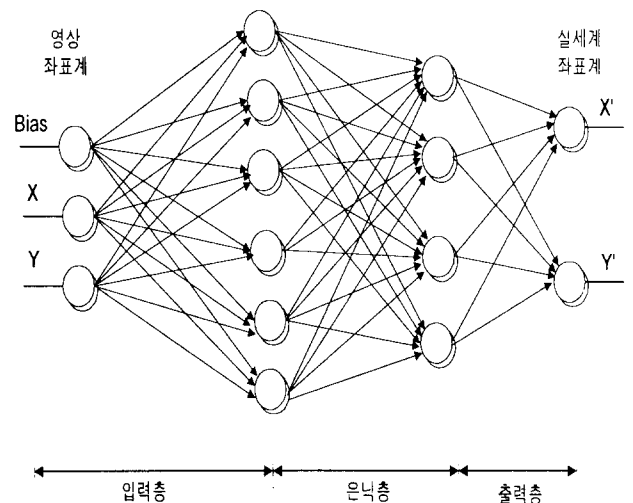


그림 2 사용된 다층구조 신경망
Fig. 2 Example of Multilayer Neural Network

카메라 보정을 위하여 균일한 넓이를 갖는 격자무늬에 대한 영상을 취득한다. 이렇게 취득된 영상으로부터 일정한 간격으로

보정점들을 얻게 되면, 이 점들의 좌표(x,y)가 Fig. 2의 다층구조 퍼셉트론 형태의 신경망의 입력이 된다. 또한 보정점들은 실제공간상의 한 점을 기준으로 일정한 간격을 두고 존재하는 점들이므로 각 보정점들의 실제 위치(x',y')가 목표값이 되어 위의 알고리즘을 수행하게 된다.

3. 실험

제한된 카메라 보정 알고리즘의 타당성을 입증하기 위하여 실험영상에 의한 실험을 수행하였다. SAMSUNG CCD 카메라를 스카라 로봇의 2관절에 연결하여 카메라 보정을 위한 영상을 취득하였으며 IBM PC/AT 586 컴퓨터를 이용하여 카메라 보정 알고리즘을 수행하였다. 보정을 위한 영상은 각기 다른 세 위치에서 취득하고, 신경망을 이용한 카메라보정과 기존의 가장 보편화된 방식인 두단계 방법을 동시에 적용하여 비교분석한다.

두 방식의 객관적 비교를 위해선 보정의 정확도를 관별하는 기준이 설정되어야 한다. 보통 많이 쓰이는 기준은 컴퓨터 영상 좌표로부터 측정된 실세계 좌표의 정확도이다. 그러나 본 연구에서는 실세계 좌표에서 추정된 컴퓨터 영상 좌표의 정확도를 보정의 정확도 평가의 기준으로 설정하였다. 이러한 기준에 의한 정확도는 픽셀의 단위로서 나타나므로 렌즈의 확대율과는 무관하게 카메라 보정의 정확도를 절대적으로 평가할 수 있다. 실세계 좌표로부터 추정된 영상 좌표의 정확도를 규정하기 위해 본 연구에서는 평균 픽셀 단위 오차 평균픽셀오차, AE를 다음과 같이 정의한다.

$$AE = \frac{1}{N} \sum_{i=1}^N [(X'_{ef} - X'_{ef})^2 + (Y'_{ef} - Y'_{ef})^2] \quad (2)$$

여기서 (X'_{ef}, Y'_{ef}) 는 컴퓨터 영상 좌표 (X_{fi}, Y_{fi}) 에 상응하는 실세계 좌표점 (X_{wi}, Y_{wi}, Z_{wi}) 로부터 보정된 변수를 이용하여 계산한 컴퓨터 영상 좌표의 추정치로서 카메라 모델의 관계식에 의해 구해진다.

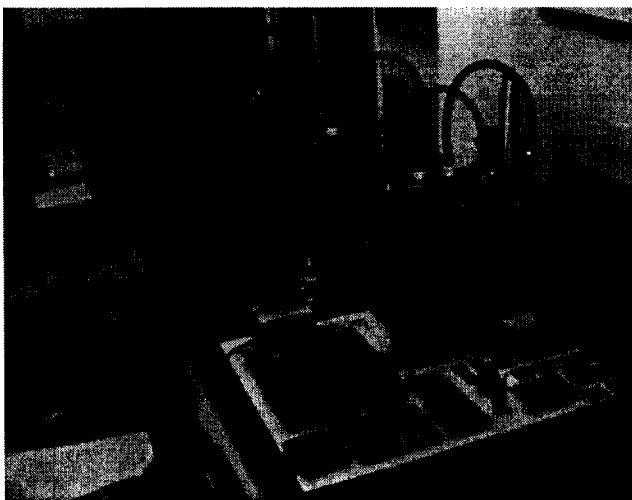
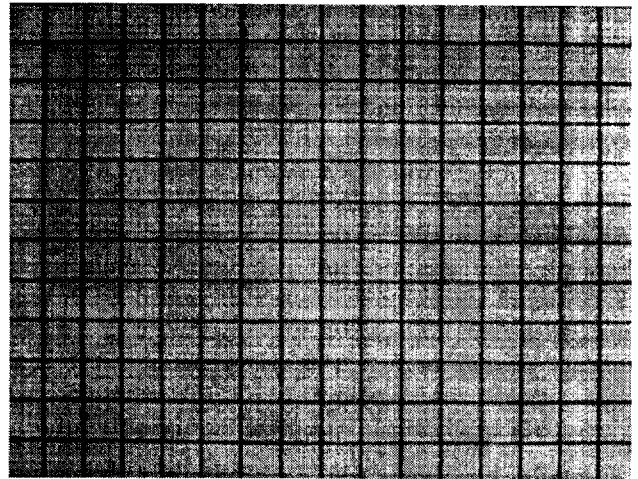


그림 3 실험장치
Fig. 3 Exmental set-up

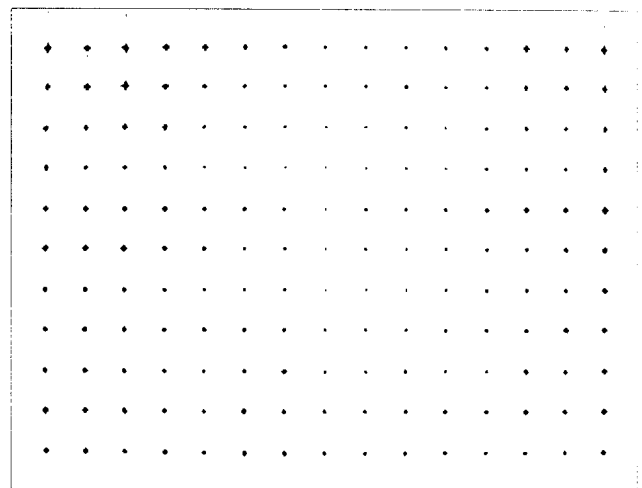
Fig. 4는 카메라 보정 실험을 위하여 취득한 영상이다. 그리고 각 영상으로부터 45개의 심플점을 취득하여 그 중 24개를 학습

데이터로 나머지 8개를 테스트 데이터로 사용하였다.

2단계 방법은 24개 보정점만을 이용하여 최종오차를 구하기 때문에 객관적인 비교를 할 수 없다. 따라서 신경망을 이용한 방법에서도 학습에 사용한 24개 데이터를 테스트하였으며, 마찬가지로 2단계 방법에서도 처음 24개의 보정 데이터로 구하여진 카메라 보정 변수를 이용하여 보정에 사용하지 않는 8개 데이터로 실위치를 예측해 실제 위치와의 오차를 구하였다. 표 1은 두 방법을 실제 영상에서 취득한 데이터로 실험한 결과이다.



(a)



(b)

그림 4 (a)취득된 실영상, (b)보정점 검출결과

Fig. 4 (a) The acquired image from cameram.

(b) The recognition result of calibration point

표 1은 신경망 방법에 의한 보정결과를 나타내며 표 2는 두단계 방법에 의한 보정결과를 나타낸다 표 1과 표 2의 결과를 보면 2단계 방법은 영상을 수직적으로 취득한 경우 결과를 보이지만 Fig. 4와 같이 카메라의 위치가 밀어지거나 비스듬한 경우 오차가 상당히 커짐을 알 수 있다. 특히 8개 테스트 데이터의 경우 중심에서 다소 떨어진 영상의 왜곡이 심한 부분에서 중점을 취득한 경우 2픽셀 이상의 오차를 보여 왜곡에 대한 보상이 원만히 이루어지지 않고 있음을 알 수 있다. 반면에 신경망을 이용한 방법에서는 모든 경우에 있어서 2단계 방법보다 좋은 결과를 내었으며 왜곡이나 카메라의 위치에 관계없이 적절하게 보상이 이루어짐을 알 수 있다.

표 1 신경망 방법에 의한 보정 결과

TABLE 1 Calibration result by neural network method

사용렌즈	렌즈 I (f=8mm)	렌즈 II (f=16mm)
보정점의 수	24	24
AE(픽셀)	0.9597	0.7174
오차합수	0.0043	0.0041
비 고	10번 반복후 결과	10번 반복후 결과

표 2 두단계 방법에 의한 보정 결과

TABLE 1 Calibration result by two-step method

사용렌즈	렌즈 I (f=8mm)	렌즈 II (f=16mm)
보정점의 수	24	24
AE(픽셀)	1.5791	4.6822
오차합수	0.0211	0.2260

4. 결 과

제안된 보정방법은 기존의 방법과 같이 정확한 카메라 모델을 설정하고 그 모델의 변수들을 계산하는 대신에 신경망 모델을 학습에 의해 구성하는 신경망 모델을 이용하는 최초의 시도이다. 기존의 방법들은 렌즈의 왜곡을 근사화하고 왜곡을 구하는데 어려움이 있을 뿐만 아니라 영상의 중점, 불확실성과 같은 변수는 추정이 힘든 문제점이 있으나 제안된 방법은 학습된 신경망 모델이 이러한 모든 요소를 내포하고 있으므로 정확도가 훨씬 높다. 실제 실험 결과에서도 널리 사용되고 있는 2단계 방법과 비교한 결과 더 정확한 보정결과가 나타났다.

논문에서 나타난 보정과정은 단지 3차원 실세계중 2차원 평면상의 점들을 2차원 영상좌표상으로 변환하는 과정만 구성되었다. 향후 제안된 신경회로망을 이용한 보정기법을 적용한 신뢰성 향상에서 대한 연구가 요구된다.

참 고 문 헌

- [1] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning internal representations by error propagation", Parallel Distributed processing, vol. 2, pp. 318-362, MIT Press, Cambridge, 1987
- [2] R. Tsai, "An efficient and accurate camera calibration technique for 3-D machine vision", Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 364-374, 1986
- [3] Y. Yakimovsky and R. Cunningham, "A system for extracting three-dimensional measurements from from a stereo pair of TV cameras", Computer Graphics Image

Processing, vol.7, pp. 195-210, 1978

[4] J. Y. Itoh, A. Miyauchi and S. Ozawa, "Direct measuring method using only simple vision constructed for moving robots", Proc. 7th Int. Conf. on Pattern Recognition, vol. 1, pp.192, 1984

[5] H. A. Martins, J. R. Birk and R. E. Kelly, "Camera models based on data from two calibration planes", Computer Graphics Image Processing vol. 17, pp. 173-180, 1981

[6] I. Sobel, "On calibrating computer controlled cameras for perceiving 3-D scenes", Artificial Intell., vol. 5, pp. 185-188, 1974