

---

# 신경망을 이용한 간단한 카메라교정

전 정 희\*, 김 충 원\*\*

## Simple Camera Calibration Using Neural Networks

Jeong-Hee Jeon, Choong-Won Kim

### 요 약

카메라 교정(Camera calibration)은 알고있는 월드 좌표계(world coordinate system)의 제어점(control points)들에 대하여 카메라의 내부/외부 인자(internal and external parameters)들을 계산하는 과정이다. 정확한 카메라 교정은 정밀한 측정을 위해서 반드시 요구된다.

본 논문에서, 우리는 3D 기하학이나 카메라 광학에 대한 특별한 지식을 요구하지 않는 신경망을 이용하여 간단하면서도 유연한 카메라 교정을 제안한다. 제안한 방법은 내부/외부 인자를 요구하지 않는 응용 분야에 매우 유용하다. 또한 제안한 카메라 교정은 물체가 이미지 평면과 거의 평행할 경우에 발생하는 악조건(ill-condition)문제를 해결할 수 있는 장점을 가졌다. 이러한 악조건은 시각 시스템을 이용하여 제품 검사를 할 경우에 흔히 발생한다.

좀더 정확한 교정을 위해 획득한 이미지는 렌즈의 방사형 왜곡에 따라 두 개의 지역으로 분할하여 교정된다. 그리고 Tsai의 알고리즘을 이용한 결과와 제안한 방법을 이용하여 교정한 결과를 실험을 통해 타당성을 증명한다.

### Abstract

Camera calibration is a procedure which calculates internal and external parameters of a camera with the known world coordinates of the control points. Accurate camera calibration is required for achieving accurate visual measurements.

---

\* 조선대학교 대학원 컴퓨터공학과 박사과정

\*\* 조선대학교 공과대학 컴퓨터공학부 교수

접수일자 : 1999년 11월 19일

In this paper, we propose a simple and flexible camera calibration using neural networks which doesn't require a special knowledge of 3D geometry and camera optics. There are some applications which are not in need of the values of the internal and external parameters. The proposed method is very useful to these applications. Also, the proposed camera calibration has advantage that resolves the ill-condition as object plane is near parallel image plane. The ill-condition is frequently met in product inspection.

For little more accurate calibration, acquired image is divided into two regions according to radial distortion of lens and neural network is applied to each region. Experimental results and comparison with Tsai's algorithm prove the validity of the proposed camera calibration.

## I. 서론

머신비전에서 카메라교정이란 카메라의 초점거리, scale 요소, 영상 중심 그리고 렌즈왜곡(lens distortion)과 같은 내부인자(intrinsic parameter)와 외부인자(extrinsic parameter)라 불리는 어떤 기준 좌표 계에 대한 카메라 좌표 계의 3차원적인 위치와 방향을 결정하는 과정이다[9].

일반적인 카메라교정은 크게 선형(linear), 비선형(nonlinear) 그리고 multi-step 방법으로 나누어진다. 선형방법은 반복적 계산이 요구되지 않기 때문에 속도가 빠르나 다소 정밀도가 떨어지고 비선형 방법은 비교적 정밀한 카메라 교정이 이루어지나 반복적인 최적화 기술을 요구한다. 이와같이 카메라 인자(camera parameter)의 특성에 따라 선형 및 비선형 방법을 이용한 카메라교정이 Tsai의 2단계(two-stage) 방법이다. Tsai의 알고리즘은 월드좌표계와 영상좌표계와의 관계를 동일평면(coplanar)의 경우와 비동일평면(noncoplanar)의 경우로 분류하여 카메라교정을 하였다. 특히, 카메라를 통하여 획득한 영상 내에 포함된 렌즈의 왜곡을 고려하여 교정할 수 있는 카메라 모델을 제시하였다. 그런데 Tsai 알고리즘은 영상평면과 물체평면간의 상대적인 각도가 항상 30° 이상을 유지하여야만 정확한 계산 결과를 산출할 수 있다는 특징이 있다[1].

그러나 물체의 품질검사나 측정 등을 위한 머신비전은 일반적으로 물체평면과 영상평면이 대부분 거의 평행하게 유지된다. Tsai 알고리즘은 정확히 평행한 구조에서는 월드좌표계와 영상좌표계의 관계를 나타내는 회전변환 행렬상의 일부 요소들이 0이 되어 수학적으로 부정이 되고, 거의 평행인 경우에는 카메라인자를 구하기가 매우 어렵게 되는

단점이 있다. 또한 카메라를 통하여 획득되는 영상은 카메라 렌즈, CCD소자, frame grabber 등 여러 종류의 하드웨어를 통과하여 획득된다. 그러므로 근본적으로 내재된 많은 하드웨어의 불완전성으로 인하여 월드좌표계와 영상좌표계와의 관계를 명확하게 규명하기가 어렵게 된다. 또한 렌즈 자체의 불균일성과 렌즈조립시 발생하는 조립정밀도의 저하, 프레임 그레버와 CCD 카메라 동기 신호의 불일치 등 하드웨어의 불완전성도 카메라교정 오차 발생의 큰 원인이 된다[2].

본 논문에서는 이러한 수학적 모델의 한계와 내재된 하드웨어의 불완전성을 극복하고 융통성 있는 카메라교정을 위하여 신경회로망(neural networks)을 이용하여 카메라교정을 한다[3][6][8].

## II. 카메라교정을 위한 신경망 제안

### 1. 다층 퍼셉트론의 구성

카메라 교정을 신경망에 적용하기 위한 전체적인 개념은 다음과 같다. CCD 카메라로 획득한 영상의 포인트들에 대한 좌표를 신경망의 입력노드 값으로 설정한다. 그리고 신경망의 출력은 영상 포인트의 실세계 좌표가 출력되도록 구성을 한다. 그러므로 카메라 교정을 위해 획득한 영상의 각 포인트들에 대한 좌표와 이에 대응한 알고있는 월드좌표를 신경망에 학습시키기 위한 학습자료로 정의한다. 이와같은 학습자료들을 우리가 적용할 신경망의 형태인 다층퍼셉트론(Multi-Layer Perceptron : MLP)에 적용하여 학습시킨다. 학습된 신경망은 카메라로부터 획득한 영상의 영상좌표 ( $X_i, Y_i$ )을 입력하면 이에 대응하는 월드좌표 ( $x_w, y_w$ )가 출력되도록 신경망을 구성한다. 이런 전체적인 개념이 Fig. 1에

있다.

Fig. 1에서와 같이 다층퍼셉트론의 역할은 카메라 교정을 위한 학습과 학습을 통하여 실제 월드 좌표표를 추출할 수 있는 중간역할을 하도록 구성한 것임을 알 수 있다.

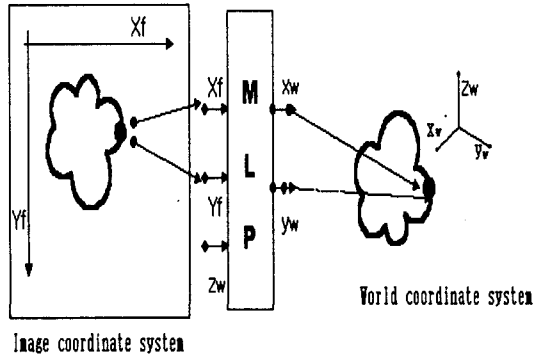


그림 1. 신경망을 이용하여 카메라교정을 하기 위한 도해

Fig. 1. Illustration for camera calibration using neural network

그런데 신경망의 입력층에  $z_w$ 를 추가함으로써 신경망은 동일평면과 비 동일평면 모두 유연하게 적용될 수 있는 장점을 갖추게 된다. 즉 동일평면 일 경우에는  $z_w=0$ 으로 모든 점이 동일 평면에 존재하도록 구성할 수 있으며, 비동일 평면의 경우에는  $z$ 축을 따라 control target을 이동시키면서 교정을 할 수 있는 장점이 발생된다. 이와같은 목적을 달성할 수 있도록 카메라교정에 사용될 다층퍼셉트론의 형태를 구성하면 Fig. 2과 같다.

다층퍼셉트론은 하나의 입력층과 두 개의 은닉층, 그리고 하나의 출력층으로 구성된다. 입력층은 세 개의 노드로 이루어지며 은닉층들은 각 다섯 개의 노드, 그리고 출력층은 두 개의 노드들로 구성된다. 또한 이런 모든 노드들은 완전히 연결되어 있는 형태이다. 이때 입력층과 은닉층간의 활성화 함수는 시그모이드(sigmoid) 함수를 사용하였고 두 번째 은닉층과 출력층 사이에는 선형(linear) 함수를 사용하였다. 이러한 구조를 가진 다층퍼셉트론은 다수의 시행착오를 거쳐서 최종적으로 우수하

다고 생각되는 것을 택한 것이다. 그러나 다른 형태의 성능은 매우 다양하게 변화하였는데, 예로서 동일한 구조라 하더라도 노드간에 완전 연결되지 않는 경우에는 50%이상 성능이 저하되었다.

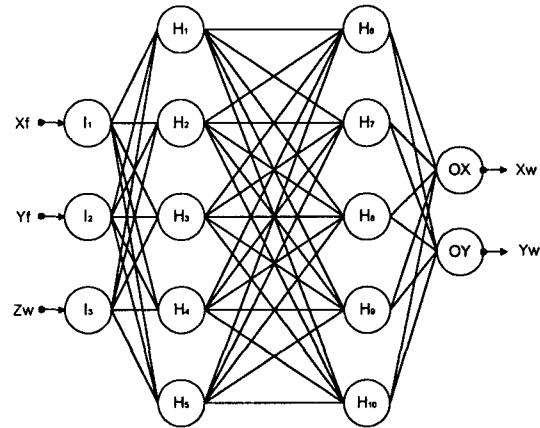


그림 2. 카메라교정을 위한 다층퍼셉트론의 구조  
Fig. 2. Structure of multi-layer perceptron for camera calibration

## 2. 영상의 영역별 분류

본 논문에서는 좀 더 정확히 교정하기 위하여 카메라로 획득한 영상을 렌즈왜곡이 큰 영역과 작은 영역으로 나누고 각 영역에 신경망을 하나씩 두어 카메라교정을 행하는 방법을 제안하고자 한다. 획득한 영상의 모든 좌표와 대응하는 월드좌표를 만일 한 개의 신경망으로 학습시키는 경우 많은 오차가 발생하였다. 그래서 본 논문에서는 렌즈 왜곡이 큰 영역의 점들과 렌즈왜곡이 적은 영역의 점들을 하나의 신경망에 적용할 경우 카메라 교정 오차가 커지는 문제를 줄이고자 영역별로 분류하여 각 영역마다 각각의 신경을 두어 교정할 것을 제안한다. 이렇게 점들을 영역별로 분류하여 각 영역에 따른 신경망을 적용하기 위한 도해가 Fig.3에 있다.

카메라 렌즈의 왜곡에 따라 영상을 영역별로 나누는 방법은 교정할 점들이 영상의 중심으로부터 얼마나 떨어져 있는지를 조사하여 중심에서 가까운 거리에 있는 점들과 멀리 떨어져 있는 점들로 나누고 이를 각 신경망의 학습 데이터로 하여 신

경망을 학습시킨다. 그 이유는 영상의 중심에서 멀어질수록 렌즈의 왜곡이 심해져서 카메라교정 오차가 증가하기 때문이다.

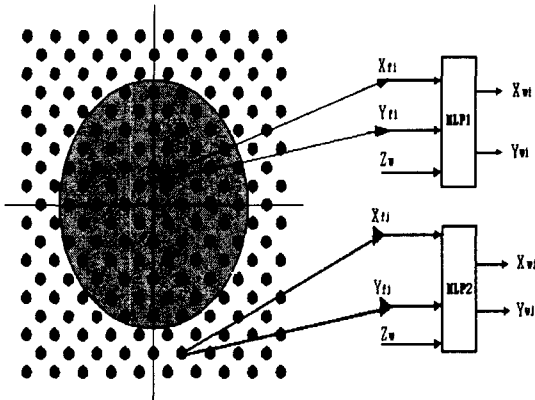


그림 3. 왜곡된 입력영상의 데이터 분류  
Fig. 3. Data classification with distortion of input image

Tsai 알고리즘에서 왜곡계수  $k$ 와 원점에서 임의의 영상좌표  $(X_d, Y_d)$ 까지 거리는 비선형 식으로 다음과 같이 표현된다.

$$\begin{aligned} D_x &= X_d(k_1 r^2 + k_2 r^4 + \dots) \\ D_y &= Y_d(k_1 r^2 + k_2 r^4 + \dots) \end{aligned} \dots\dots\dots (1)$$

$$r = \sqrt{X_d^2 + Y_d^2}$$

위 식에서 왜곡 계수  $k$ 와 거리는 무한대로 발산하므로 임의의 한 점  $(D_x, D_y)$ 에 대하여 렌즈의 중심으로부터 떨어진 거리를 간단히 계산해 보면 아래와 같음을 알 수 있다.

$$\begin{aligned} D &= (D_x^2 + D_y^2)^{1/2} \\ &= r^2 + r^4 + \dots \end{aligned} \dots\dots\dots (2)$$

$$\approx r^2$$

여기에서  $D$ 는  $r^2$ 에 비례한다. 일반적으로 머신비전에서 고려하는 렌즈의 왜곡은 주로 radial distortion만 고려하기 때문에 우리는 radial distortion만 구하여 적용하였다.

그러므로 각 점들은 기준값  $R$ 에 따라서 두 개의 지역으로 분할된다. 기준값  $R$ 은 다음 식으로 구할

수 있음을 Fig. 4 통하여 쉽게 알 수 있다.

$$R = 1/2(X_{fmax} - Cx) \dots\dots\dots (3)$$

Fig. 4는 이미지 중심으로부터 거리에 따른 왜곡의 변화에 따라 점들을 분류한 것에 대한 그림이다.

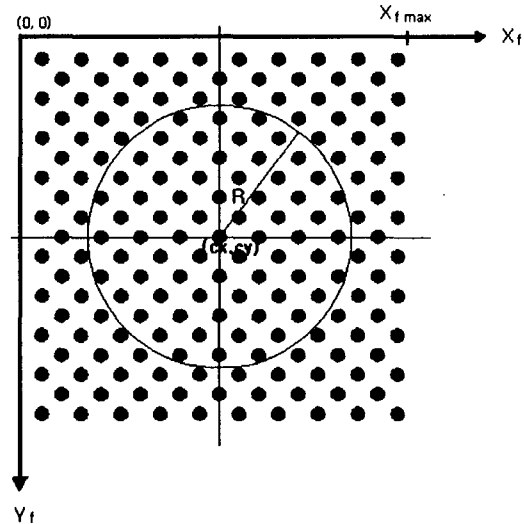


그림 4. 이미지 중심으로부터 거리에 따른 왜곡의 변화  
Fig. 4. Distortion change of distance from image center

즉 원 내부에 존재하는 점들의 집합과 원 외부에 존재하는 점들의 집합을 각각 다른 신경망에 적용한다. 이렇게 영역별로 카메라교정을 위한 데이터 값들이 각각 신경망에 학습되어 서로 다른 가중치(Weight) 값을 갖는 신경망을 구성한다. 이렇게 신경망을 영역별로 적용함으로써 카메라교정 데이터를 거리별로 나누지 않고 학습시켰을 때보다도 훨씬 정확한 출력 값을 얻을 수 있었음을 다음 실험 결과에서 확인하였다.

### III 실험 및 고찰

본 연구에서는 정확한 카메라교정을 위해 control target을 경북대학교 반도체 센터에 의뢰, 제작하여 사용하였다. Target points는 Fig. 3와 같

이 181개의 원으로 이루어 졌으며 모두 정확히 2D 평면에 있다. 각 원의 지름은 2mm이고 점들간의 간격은 6mm이다. 이런 점들 중 100개는 다층퍼셉트론의 학습용으로 사용하였고 나머지 81개의 점들은 카메라교정이 제대로 되었는지를 판단하는 테스트용으로 사용하였다.

카메라교정을 위한 실험은 control target의 영상을 CCD카메라를 거쳐 획득한 후 blob analysis를 하여 control point들의 각각의 좌표 ( $X_f, Y_f$ )를 구한다[4][5][7]. 그리고 알고있는 control point들의 월드좌표 ( $x_w, y_w$ ) 100개와  $z_w$ 을 학습데이터로 하여 신경망을 학습시킨다. 이때  $z_w$ 의 값은 모두 제로로 하였다. 그리고 나머지 81개의 점들에 대한 좌표 값을 이용하여 카메라교정을 위한 성능 평가 자료로 활용했다.

제안한 신경망을 이용하여 간단한 카메라교정의 성능평가를 위하여 카메라교정 오차들의 평균과 최대 및 최소 값은 다음 식으로 구하였다.

$$XE_i = |Xw_i - Ox_i|, YE_i = |Yw_i - Oy_i| \dots\dots\dots (4)$$

- 여기에서,
- ( $Ox_i, Oy_i$ ) : 영상좌표 ( $X_f, Y_f$ )에 대응한 기준점  $i$ 의 다층퍼셉트론 출력값
- ( $x_{wi}, y_{wi}$ ) : 기준점  $i$ 의 월드좌표
- $XE_i$  : 기준점  $i$ 의 x방향 절대오차
- $YE_i$  : 기준점  $i$ 의 y방향 절대오차

카메라교정 실험은 동일평면(coplanar)인 경우에만 적용하였고, 영상 평면과 물체평면인 control target과 이루는 각도가 0°인 경우와 30°인 경우 두 가지를 실험하였다. 이런 실험 결과를 Tsai 알고리즘을 이용한 카메라교정의 결과와 비교하였다. 그러나 Control target과 영상평면의 각도가 0°인 경우(ill-condition), 즉 평행한 경우에는 Tsai 알고리즘을 적용할 수 없기 때문에 카메라교정 결과를 비교할 수 없었으며, 영상을 영역별로 둘로 나누는 경우와 나누지 않는 경우도 비교함으로써 영역별로 나누는 것이 오차가 적게 발생한다는 것을 실험을 통하여 입증하였다. 이에 대한 실험결과가 Table 1과 Table 2에 있다.

표 1. 카메라와 control target과 이루는 각도가 0° 일 때 교정 에러

Table 1. Calibration error with camera and control target in 0°

	영역을 나누지 않는 경우		영역을 둘로 나눈 경우	
	XE	YE	XE	YE
최대	0.251442	0.381690	0.162125	0.103367
최소	0.001167	0.000158	0.000092	0.000114
분산	0.134964	0.144073	0.082195	0.059412
평균	0.099661	0.092952	0.047526	0.041128

	Tsai 알고리즘	
	XE	YE
최대	구할 수 없음	구할 수 없음
최소	-	-
분산	-	-
평균	-	-

표 2. 카메라와 control target과 이루는 각도가 30° 일 때 교정 에러

Table 2. Calibration error with camera and control target in 30°

	영역을 나누지 않는 경우		영역을 둘로 나눈 경우	
	XE	YE	XE	YE
최대	0.726452	0.780254	0.334602	0.239288
최소	0.001848	0.007267	0.002779	0.259833
분산	0.503503	0.269259	0.144453	0.124772
평균	0.231764	0.195621	0.2078620	0.0757202

	Tsai 알고리즘	
	XE	YE
최대	0.649792	0.652612
최소	0.000189	0.000820
분산	0.196771	0.125634
평균	0.085344	0.064060

위 실험결과를 통하여 본 논문에서 제안한 신경망을 이용한 카메라교정 방법이 물체평면과 영상 평면과 이루는 각도에 상당히 유연함을 알 수 있다. 또한 영상을 렌즈의 왜곡에 따라 영역별로 나누어 각각의 신경망에 적용하여 카메라교정 한 경우와 영상의 왜곡률을 고려하지 않고 모든 점들을 하나의 신경망에 적용시킨 경우의 결과를 서로 비교하였다. 비교결과 영역별로 나누어 교정하는 방법이 우수하다는 것을 알 수 있다.

그런데 물체평면과 영상평면의 각도가 30° 기울어진 경우에는 신경망을 이용한 카메라교정 보다 Tsai 알고리즘으로 해를 구한 카메라교정이 더 정확하다는 것을 알 수 있다.

그러나 Tsai 알고리즘은 수학적 모델이 복잡할 뿐만 아니라 계산과정이 복잡하고 카메라교정을 위한 처리시간도 많이 걸리는 단점 때문에 만일 신경망을 이용하여 카메라교정을 하면 각도에 융통성이 있으며, 신경망을 쉽게 구성할 수 있어 많은 장점이 발생됨을 알 수가 있었다. 이런 점들을 고려한다면 본 논문에서 제안한 신경망을 이용한 간단한 카메라교정 기법이 특별한 의미를 갖고 있다고 할 수 있다.

#### IV. 결 론

본 논문에서는 신경망을 이용한 간단한 카메라 교정 기법을 제안하였다. 카메라교정은 동일평면 조건에서 영상평면과 물체평면이 서로 평행한 경우 즉, 0°인 경우와 30° 기울어진 경우에 대하여 영상의 왜곡을 고려하여 교정을 하였다. 카메라를 통해 획득한 영상에서 렌즈의 왜곡이 많은 부분과 왜곡이 작은 부분으로 나누어 각 영역별로 신경망을 적용하여 카메라교정을 하였으며 영상을 영역별로 나누지 않고도 신경망에 적용하여 카메라교정을 하였다. 그리고 각각의 카메라교정 결과를 Tsai 알고리즘과 서로 비교하여 제안한 방법의 성능 및 타당성을 입증하였다. 신경망을 이용한 카메라교정의 장점은 물체평면과 영상평면이 이루는 각도에 관계없이 적용할 수 있어 기존의 알고리즘에 비해 유연하며 복잡한 수식 없이 쉽고 간단하게 구성할 수 있다는 점이다. 또한 카메라에 대한 내부인자와 외부인자의 구체적인 값에 관심이 없는 경우와 컴퓨터 비전 또는 3차원 기하학에 특별한 지식이 없는 경우에도 사용할 수 있다는 장점이 있다. 따라서 물체평면과 영상평면이 대부분 평행하게 적용되는 산업현장에서 물품의 품질검사 등에 유용하게 사용할 수 있을 것이다.

이논문은 과학기술부·한국과학재단 지정 지역협력연구센터인 조선대학교 수송기계부품 공장자동화 연구센터의 연구비 지원에 의해 연구되었음.

#### 참고문헌

- [1] Tsai, R. Y., "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-Shelf TV Camera and Lenses," *IEEE Journal of Robotics and Automation*, VOL. RA-3, NO. 4, pp. 323-344, 1987.
- [2] Zhuang, haqui, "Camera-Aided Robot Calibration," CRC press, pp. 11-61, 1996
- [3] Moon-Kuy Lee, Jung-Hwa Lee, "A 2-D Image Camera Calibration using a Mapping Approximation of Multi-Layer Perceptrons," *Journal of Control Automation and System Engineering*, Vol. 4, No. 4, pp. 487-493, 1998.
- [4] Gonzalez, Wood, "Digital Image Processing," Addison Wesley, 50803, 1992.
- [5] H. R. Myler and A. R. Weeks, "Computer Imaging Recipes C," Prentice Hall, Eaglewood Cliff, U.S.A., 1993.
- [6] L.V. Fausett, Fundamental of neural networks: Architectures, Algorithms, and Applications, Prentice Hall, 1994.
- [7] J. Canny, A Computational Approach to Edge Detection, *IEEE Transactions on Analysis and Machine Intelligence*, VOL. PAMI-8. NO. 6, pp. 679-698, 1986.
- [8] M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis, and Machine Vision, PWS publishing, 2nd . ed., 1998.
- [9] H. Wechsler (ed.), Neural Networks for Human and Machine Perception, Academic Press, 1991.



전 정 희(Jeong-Hee Jeon)  
1993년 2월 조선대학교 컴퓨터  
공학과 (공학사)  
1997년 8월 조선대학교 일반대  
학원 컴퓨터공학과 (공학석사)  
1997년 9월 ~ 현재 조선대학교  
대학원 컴퓨터공학과 박사과정

\*관심분야 : 영상처리, 컴퓨터 비전, 컴퓨터 그래픽스, 웹에서의 3차원복원 및 처리 등



김 충 원(Choong-Won Kim)  
1982년 2월 한양대학교 전자공  
학과(공학사)  
1984년 2월 한양대학교 대학원  
전자공학과(공학석사)  
1989년 2월 한양대학교 대학원  
전자공학과(공학박사)

1989년 5월 ~ 91년 9월 조선대학교 컴퓨터공학과  
전임강사

1991년 10월 ~ 95년 9월 조선대학교 컴퓨터공학과  
조교수

1993년 11월 ~ 94년 7월 일본 구주대학교 전자공  
학과 Post Doc.

1995년 10월 ~ 현재 조선대학교 컴퓨터공학과 부  
교수

\* 관심분야 : 영상처리, 영상처리를 이용한 품질검  
사, 컴퓨터 비전, 컴퓨터 그래픽스 등