

컴퓨터 視覺 裝置의 開發

(Development of a Simple Computer Vision System)

朴 東 哲*, 石 玟 秀**
(Dong Chul Park and Min Soo Suk)

要 約

시각 능력에 의한 작업 대상의 인식 기능을 갖는 로봇 시스템에 필요한 image digitizer와 기본적인 몇 가지 소프트웨어 기법의 개발에 관한 연구 결과를 보고한다.

Image digitizer는 CROMEMCO SYSTEM III 마이크로컴퓨터와 C. C. T. V. 카메라를 이용하여 개발하였으며, 소프트웨어 기법들은 3 차원 물체의 인식에 중점을 두었는데, 이 기법들의 유용성을 간단한 물체인 직육면체에의 응용을 통해 보였다.

Abstract

To give the recognition capability of task objects by computer vision to a sensor-based robot system, an image digitizer and some basic software techniques were developed and reported here.

The image digitizer was developed with the CROMEMCO SYSTEM III microcomputer and C.C.T.V. camera to convert the analog valued scene into digitized image which could be processed by a digital computer.

Basic software techniques for the computer vision system were aimed at the recognition of 3-dimensional objects. Experiments with these techniques were carried out using the image of a cubicle which could be considered as typical simple 3-dimensional object.

I. 序 論

최근 산업용으로 활용되고 있는 로봇트는 유연성(flexibility)과 다기능성(versatility)을 특징으로 하며, 작업 조건의 변화가 심하고 인간이 작업하기 어려운

환경에서 많이 이용되고 있다. 종래의 로봇트는 정밀하게 작성된 프로그램과 간단한 감지기(sensor) 들의 도움으로 어느 정도 지능적인 작업을 수행하고는 있으나 임의적인 작업 조건의 변화에는 충분히 대처할 수 없다. 이러한 로봇트에 시각에 의한 작업 대상 인식기능과 고도의 지능을 부여함으로써 개념적으로 간단히 작성된 프로그램만으로도 복잡하고 정밀한 작업을 수행해 낼 수 있을 것이다.

이때, 목적물의 인식을 위해 목적물에 대한 영상이 필요하게 되는데, 일반적으로 생각되는 영상은 디지털 컴퓨터로 처리할 수 없는 아날로그 영상이다. 이러한 이유로 본 연구에서는 시각 장치의 입력이 되는 C.C.

* 正會員, 金星社 中央研究所
(Gold Star Co. Central Research Laboratory)

** 正會員, 韓國科學技術院 電氣 및 電子工學科
(Dept. of Electrical Science, KAIST)

接受日字: 1982年 7月 6日

(※ 이 論文은 1981年度 韓國科學財團 受託研究費에 依하여 研究된 것임.)

T.V. 카메라의 출력 신호를 계수화하여 컴퓨터로 처리할 수 있게 하는 image digitizer를 개발하였고, 이 image digitizer에서 얻어진 계수화된 영상을 컴퓨터로 처리하여 목적물의 크기, 위치, 방향 등을 결정하는데 필요한 컴퓨터 시각장치(vision subsystem)의 알고리즘을 개발하였다.

본 연구에서는 일반적으로 사용되고 있는 간단한 물체의 2차원적인 해결 방법^[1, 2]을 사용하지 않고, 복잡하기는 하지만 3차원적인 방법을 사용하였는데, 이는 앞으로의 계속적인 연구에서 발생될 수 있는 현실적인 문제 해결을 위한 기초가 될 수 있기 때문이다.

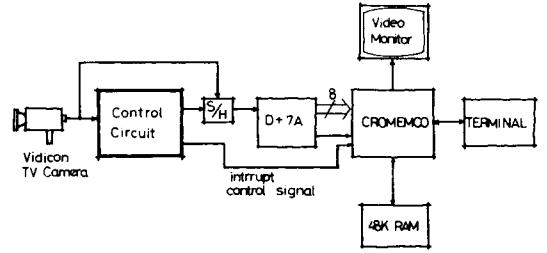


그림 1. 비디오 제어 시스템의 구조
Fig. 1. Block diagram of a video interface system.

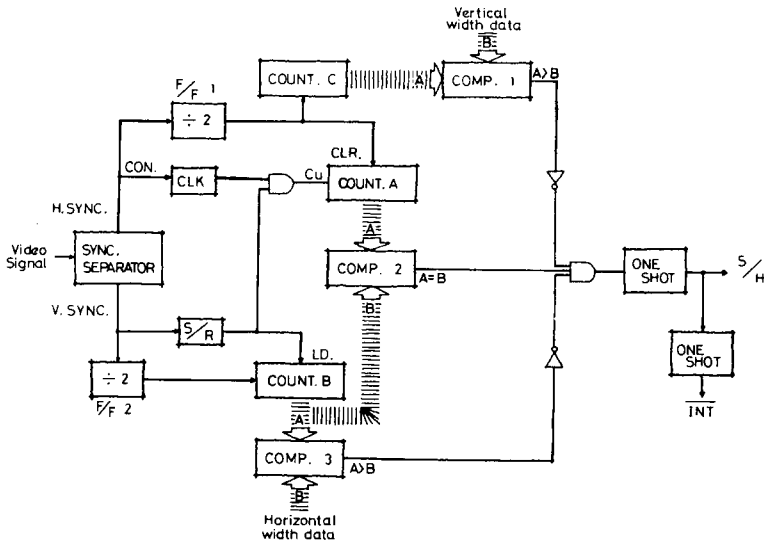


그림 2. 제어회로의 구조
Fig. 2. Block diagram of control circuit.

본 논문의 구성은 다음과 같다. 제 II 장에서 디지털 영상의 데이터를 얻기 위한 image digitizer의 개발에 관해 쓰였으며, 제 III 장에서는 컴퓨터 시각 장치의 기본적 알고리즘에 관해 언급되었으며, 제 IV 장에서는 이들 알고리즘에 대한 간단한 응용과 그 결과를 통해 유용성을 보였고, 제 V 장에서는 결론과 앞으로의 연구 방향을 제시하였다.

II. 시각 입력 장치의 개발

영상처리(image processing)에 있어 디지털 영상처리(digital image processing)의 비중이 점차 증가하는 경향을 보이고 있는데, 디지털 영상처리에서는 image enhancement, noise reduction, data compression,

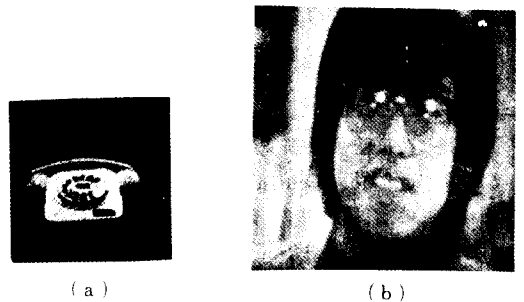


그림 3. Digitize된 image (a) 100×100 (b) 200×200
Fig. 3. Digitized images displayed on video monitor.

pattern recognition, tomography 등의 분야에서 급속한 발전을 보여왔다.^[4, 6] 그런데, 아나로그 입력 영상은 디지털 형태로 변환되어야만 디지털 영상처리가 가능하기 때문에, 본 연구의 첫번째 단계로 image digitizer의 개발이 필요하게 되었다.

이에 CROMEMCO SYSTEM III 마이크로컴퓨터와 Vidicon C.C.T.V. 카메라를 이용하는 image digitizer를 개발하였는데, 전체적인 시스템의 구조가 그림 1, 2에서 각각 보여지고 있다.

개발된 image digitizer의 제원은 다음과 같다.

- 1) NTSC(National Television Standard Committee) 방식의 T.V. 신호가 입력신호로 사용되었다.
- 2) CROMEMCO SYSTEM III 마이크로컴퓨터에 있는 D+7 A 보오드를 A/D 변환기로 사용하였다.
- 3) Digitized image 크기는 100×100이나 200×200을 외부 단자에 의해 임의로 선택할 수 있다.
- 4) 변환 시간은 100×100 영상의 경우에 약 3.33초, 200×200 영상의 경우에는 약 6.67초가 소요된다.
- 5) 화소의 해상도(resolution)는 256계층이다.

그림 3에서 비디오 모니터에 디스플레이된 16계층의 해상도를 갖는 image가 보여진다.

III. 컴퓨터 시각 장치의 기본적 알고리즘

본 연구의 중요한 목적중의 하나가 컴퓨터 시각 장치에 유용한 기본적인 알고리즘의 개발이다. 다음에 이들 알고리즘이 일련의 과정으로 기술된다.

1. 경계 추출

계수화한 화상에서 목적물의 중요한 정보를 잃지 않고 목적물을 가장 간단히 표현할 수 있는 방법의 하나가 바로 경계점들을 찾는 것이다. 여기에서 홍 순호의 기법^[3]과 같이 복잡하나 정확한 알고리즘이 사용될 수도 있겠으나, 본 연구에 사용되는 입력 영상에 잡음이 비교적 적기 때문에 처리 시간을 줄이기 위해 Sobel 기법^[4]을 사용했다.

Sobel 기법은 다음의 식에 의해 정의될 수 있는데, 명암이 급격히 변화하는 곳에서 큰 값을 갖는 물리적 의미를 지닌다.

A	D	F
B	(i, j)	G
C	E	H

$$\begin{aligned}
 X &= (I(A) + 2I(B) + I(C)) \\
 &\quad - (I(F) + 2I(G) + I(H)) \\
 Y &= (I(A) + 2I(D) + I(F)) \\
 &\quad - (I(C) + 2I(E) + I(H)) \\
 G(i, j) &= \sqrt{X^2 + Y^2} \\
 I(x) &: \text{좌표 } x \text{에서의 grey level 값}
 \end{aligned}$$

이렇게 하여 얻어진 값들을 적당한 값으로 thresholding하여 경계점들을 얻는다. 원래의 영상과 찾은 영상 점들의 영상이 그림 8 (a), (b)에서 각각 보여진다.

2. 직선 특징의 추출^[5]

앞에서 얻어진 경계점들은 아직도 정보량이 많아서 128×128의 영상의 경우 수백개나 되고, 아직 그 자체만으로는 어떤 구조적인 특징을 가지고 있지 않다. 따라서 이들은 각 경계점들 사이의 상관 관계에 의해 한 그룹으로 모아지거나 제거되어야 한다. 이러한 일들이 MST(minimal spanning tree)알고리즘에 의해 구조적으로 같은 경계점들끼리 모아졌으며, 이렇게 모아진 점들이 Ramer의 근사화 알고리즘에^[5]의해 직선으로 표시되었다. 이 결과를 그림 8 (c)에서 보았다.

순수하게 위의 두 알고리즘을 거친 결과는 그림 8 (c)에서 보듯이 물체의 꼭지점 부근에서 왜곡된 결과가 나타난다. 이것은 MST 알고리즘 자체에 의한 것으로 긴 선분 두개 사이에 끼인 작은 선분의 형태로 나타나는데, 이 작은 선분을 없애고 긴 선분 두개를 연장시켜 만나는 점을 새로운 꼭지점으로 하면 어느 정도 완전한 결과를 얻게 된다.

3. 특징 추출

목적물을 인식하게 하는데는 여러 가지 특징들이 있겠으나, 본 연구에서는 3 차원 물체의 인식이 주목적이기 때문에 평행선, 선분의 길이 등이 중요한 특징이 되어 앞에서 얻어진 선분들의 길이, 방향, 영상 평면상의 위치 등을 이용해 꼭지점들을 비교적 정확하게 그림 7에 대응시킬 수 있었다. 예를들어 그림 7의 꼭지점 ⑤는 꼭지점들 중에서 Y좌표가 큰 것들 2개중 X좌표가 가장 큰 것이며, 꼭지점 ⑦은 Y좌표가 작은 것들 2개중 X좌표가 큰 쪽의 것이며, 꼭지점 ⑥은 ⑤와 ⑦을 뺀 나머지 중에서 X좌표가 가장 큰 것이다. 이렇게 하여 꼭지점 각각을 그림 7의 각 꼭지점에 일대일 대응시킬 수 있었다. 이 결과가 그림 8 (d)에서 보여진다.

4. 퍼스펙티브 변환

본 연구의 궁극적인 목표가 3 차원 물체의 인식이기 때문에, 2 차원의 영상 데이터를 3 차원 좌표축상의 정보로 변환시키는 것이 필요하다. 이러한 변환이 투영(perspective) 변환인데, 이것은 3 차원 공간의 한 점이 영상 평면의 한점으로 대응되므로 이때의 변환관계식을 알면 역으로 영상 평면의 한점은 3 차원 공간의 점으로 표현될 수 있다는 것이다. 그러나, 이때 3 차원 공간의 점들이 2 차원 좌표인 영상 평면에 표현

될 때 many-to-one mapping이 되므로 영상 평면 위의 한 점이 3차원 공간의 한 점으로 변환되려면 또 다른 정보를 필요로 하게 된다. 즉, 육면체의 물체가 책상위에 놓여 있고, 책상면의 Z축 좌표를 0이라고 하면, 육면체의 꼭지점 중에서 책상면과 닿고 있는 네 꼭지점의 Z축 좌표는 0이 될 것이다. 이것을 이용하면 육면체의 밑바닥에 있는 네 꼭지점은 2차원 좌표로 쉽게 변환될 수 있을 것이다.

이제 그림 4에서 보이는 바와 같이 투사선(projection line)과 Z축 사이의 각, θ 와 투사선과 x축이 이루는 각, ϕ 를 이용해서 3차원 공간과 2차원 영상 평면 사이의 변환 관계식을 구해 보자.

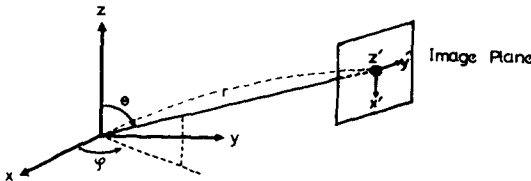


그림 4. 퍼스펙티브 변환의 요소
Fig. 4. Elements of a perspective transformation.

$$R = R_1(\phi) \cdot R_2(\theta)$$

$$= \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos \theta \cos \phi & -\sin \phi & \cos \phi \sin \theta \\ \sin \phi \cos \theta & \cos \phi & \sin \phi \sin \theta \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

이렇게 하여 영상 평면위의 좌표, (X_i, Y_i, Z_i) 와 실제 좌표축상의 좌표, (X_o, Y_o, Z_o) 와의 변환 관계가 다음 식으로 주어지게 된다.

$$(X_i, Y_i, Z_i) = (X_o, Y_o, Z_o) \times$$

$$\begin{bmatrix} \cos \phi \cos \theta & -\sin \phi & \cos \phi \sin \theta \\ \sin \phi \cos \theta & \cos \phi & \sin \phi \sin \theta \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

$$X_i = \cos \phi \cos \theta X_o + \sin \phi \cos \theta Y_o$$

$$Y_i = -\sin \phi X_o + \cos Y_o$$

따라서 위의 식으로부터 영상 평면상의 좌표로 목적물의 실제 좌표축상의 좌표를 역으로 구해보면

$$X_o = \frac{1}{\cos \theta} [\cos \phi \cdot X_i - \cos \theta \sin \phi \cdot Y_i]$$

$$Y_o = \frac{1}{\cos \theta} [\sin \phi \cdot X_i + \cos \theta \cos \phi \cdot Y_i],$$

$$\theta \approx 90^\circ$$

으로 되는데, 물론 이 식에 카메라의 비균일성(nonuniformity), 배율 등이 스케일 필터로 고려되어야 할 것이다.

IV. 간단한 응용 예제와 실험 결과

앞장에서 개발된 몇 가지 기법은 시각 기능을 지니

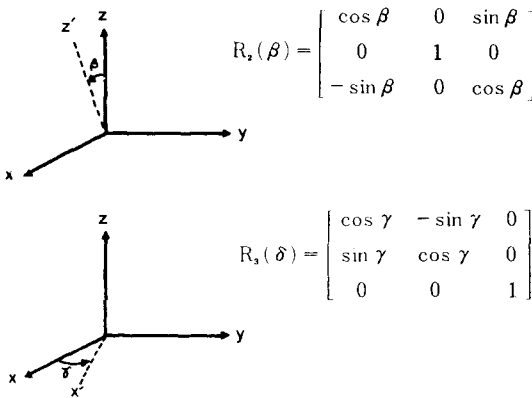


그림 5. 회전 행렬식
Fig. 5. Rotation matrices.

먼저 Y'축을 기준으로 Z축을 반시계 방향으로 θ 만큼 회전시켜 Z축을 Z'축에 평행하도록 하면 이 변환은 $R_1(\theta)$ 로 표현된다. 다시 변환된 Z축을 기준으로 하여 X축을 반시계 방향으로 ϕ 만큼 회전시키면 X축, Y축은 X', Y'에 평행하게 되며 이 변환이 $R_2(\phi)$ 로 표현 될 수 있다. 이제 이 두개의 변환을 연결시키면 다음의 식으로 표현된다.

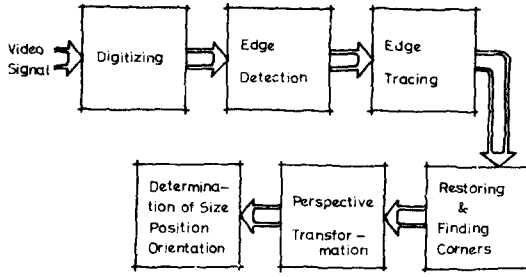


그림 6. 간단한 컴퓨터 시각 장치의 블럭도
 Fig. 6. Block diagram of a simple computer vision system.

는 로봇트 시스템에 기본적으로 사용될 수 있는 것들인 바, 이 장에서는 실제적인 유용성을 보이기 위하여 3 차원 물체의 방향, 크기, 위치 등을 인식하는 간단한 컴퓨터 시각 장치에 응용해 보인다.

그림 6 에서 간단한 컴퓨터 시각 장치의 한 블럭다 이아그램이 보여지고 있다. 그림의 첫번째 단계에서는 본 연구에서 개발된 image digitizer를 이용하여 카메라에 포착된 이미지가 계수화되며, 다음단에서 계수화된 이미지에서 Sobel 기법에 의해 목적물의 에지 포인트가 추출된다. 이들 에지 포인트는 다시 다음 단에서 MST알고리즘과 근사화 기법에 의해 선분들로 연결된다. 다음단에서는 이들 선분의 정보를 이용하여 물체의 꼭지점들을 구해내며, 그 다음단에서 이 꼭지점들이 실제 좌표축상의 좌표로 변환된다. 마지막 단계에서는 앞에서의 정보를 이용하여 물체의 위치, 크기, 방향 등을 결정한다.

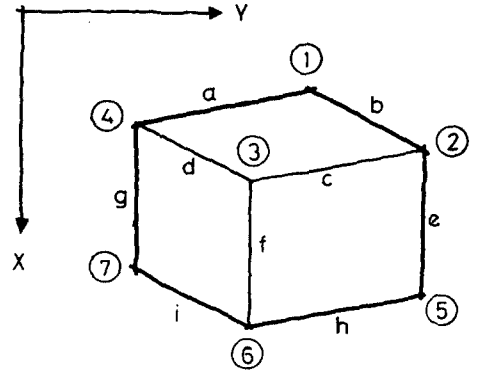


그림 7. 영상 평면위에서의 각 부분의 명칭
 Fig. 7. Naming of corners and lines in image plane.

본 연구에서는 육면체의 물체가 모델로서 사용되었는데, 이는 복잡한 3 차원 물체의 인식을 목표로하는 계속적 연구의 첫번째 단계로 시도되었기 때문이다.

표 1 에서 실제의 값과 인식된 값의 결과가 주어져 있다. 여기서 목적물의 X위치, Y위치, 방향은 다음의 식으로 정의된다.

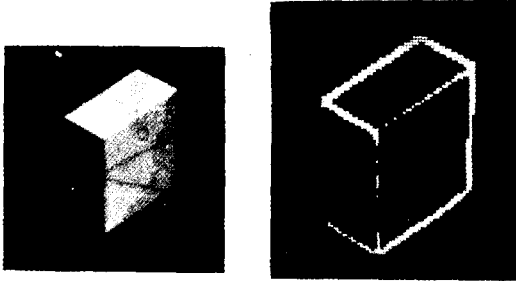
• X 위치 = $\frac{1}{2} (\textcircled{5} (x) + \textcircled{7} (x))$

• Y 위치 = $\frac{1}{2} (\textcircled{5} (y) + \textcircled{7} (y))$

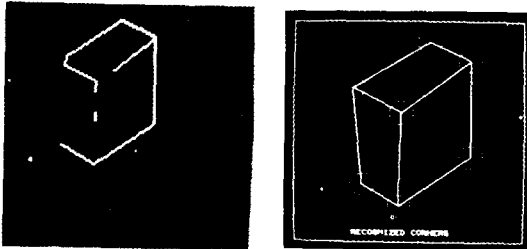
• 방 향 = Angle { max (h length, i length) }

표 1. 결 과
 Table 1. Results.

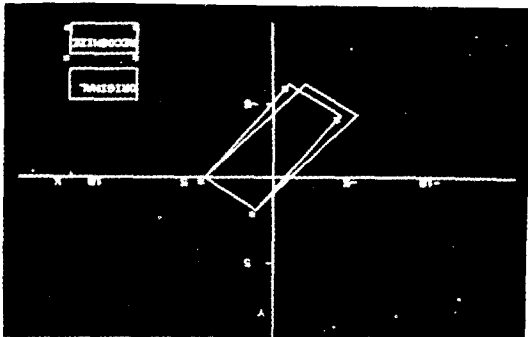
모 델	크기(가로×세로×높이)			중심의위치(X, Y)			방 향		
	입 력	출 력	오 차	입 력	출 력	오 차	입 력	출 력	오 차
1	3.5 ×	3.75 ×	0.25 ×	(-1.6, 4)	(-1.7, 3.75)	(0.1, 0.25)	135°	138.5°	3.5°
	7.5 ×	6.94 ×	0.56 ×						
	3.5	3.47	0.03						
2	9 ×	8.78 ×	0.22 ×	(0.42, 1.62)	(0.4, -1.8)	(0.02, 0.18)	52.2°	56.3°	4.1°
	4 ×	3.72 ×	0.28 ×						
	7	6.92 ×	0.08 ×						
3	34 ×	2.65 ×	0.75 ×	(0.34, 7.66)	(0.15, -7.6)	(0.15, 0.06)	29.61°	24.87°	4.74°
	3.6 ×	3.48 ×	0.12 ×						
	7.0	6.7	0.3						
4	3.5 ×	2.95 ×	0.55 ×	(3.75, 1.75)	(3.93, 1.7)	(0.18, 0.05)	0°	3.25°	3.25°
	7.5 ×	7.82 ×	0.32 ×						
	3.5	3.77	0.27						



(a) Digitized image. (b) Edge output.



(c) Traced lines. (d) Recognized corners.



(e) Perspective transformed position.

그림 8.
Fig. 8.

여기서 ⑤, ⑦, h, i는 영상 평면상의 목적물의 각 변과 꼭지점을 규정시킨 그림 7의 각각에 해당되는 실제 좌표축에서의 값이다.

그림 8 (e)에서 보이듯이 실제의 위치와 인식된 위치와의 차이는 vidicon 카메라의 비직선성에 대한 불완전한 보상과 에지 포인트가 실제 에지 포인트 부근에서 두껍게 생기는데, 이들 에지 포인트에서 선분을 추적할 때 생기는 오차에 주로 기인한 것이다.

그림 8의 일련의 그림에서 이 인식 과정에 대한 결과가 보여지고 있다.

V. 結 論

아날로그 값을 가지는 화상으로부터 디지털 이미지 프로세싱을 가능하게 하는 image digitizer를 C.C.T.V. 카메라와 CROMEMCO SYSTEM III 마이크로컴퓨터를 이용해 개발하였고, 시각 기능을 가지는 로봇 시스템에 유용한 몇 가지 기본적인 소프트웨어 기법을 개발하여 간단한 응용을 수행하였다. 이들 소프트웨어 기법은 3 차원 물체의 인식에 중점을 두었으며, 응용의 결과로 실제의 유용성을 입증하였다.

앞으로의 계속적인 연구에서는 목적물의 제한을 없앨 수 있게끔 곡선의 추적을 이용한 기법에 대한 연구가 수행되어야 하겠다.

參 考 文 獻

- [1] S. Kashioka et al., "A transistor wire-bonding system utilizing multiple local pattern matching techniques," *IEEE Trans. S.M.C.* vol. SMC-6, no. 8, Aug. 1976.
- [2] M.L. Baird, "SIGHT-I : A computer vision system for automated IC chip manufacture," *IEEE Trans. S.M.C.* vol. SMC-8, no. 2, Feb. 1978.
- [3] S.H. Hong, *A New Edge Extraction Techniques for Noisy Images*. M.S. Dissertation, KAIST, Feb. 1981.
- [4] W.K. Pratt, *Digital Image Processing* John Wiley & Sons, Inc., pp. 479-492, 1978.
- [5] O.Y. Song, *Linear Feature Detection from Complex Scene Imagery*. M.S. Dissertation, KAIST, Feb. 1982.
- [6] E.L. Hall, *Computer Image Processing and Recognition*. Academic Press, Inc., pp. 76-88, 1979.