

논문 2008-03-20

웹 캠을 이용한 스테레오 영상의 3차원 거리 측정

(3D Distance Measurement of Stereo Images Using Web Cams)

김 승 환, 함 운 철*

(Seung-Hwan Kim, Woon-Chul Ham)

Abstract : In this paper, we propose a three dimensional distance measurement method for a stereo system by using web cams. Using a parallel stereo system, a robot gets two images from each webcam and equalize brightness of both images. And we suggest an image processing method such as labeling, isolating an object from background and finding center of an object. We also propose a method of calculating the focal distance by using least square algorithm based on triangulation and we can reduce calculation error by this method. From experimental results, we show that the proposed method can be effective for 3D distance measurement

Keywords : Stereo Vision, Depth Measurement, Labeling, Image Processing

1. 서론

최근까지 로봇은 주로 로봇의 팔, 다리의 제어 등의 하드웨어 위주로 연구가 되어져 현재는 인간과 흡사한 작동이 가능할 정도로 발전이 되어 있다. 하지만 지능형 로봇에 필수적인 기능인 스스로 외부 환경의 변화나 작업 변경을 인식하고 자율적으로 동작하는 기능은 아직까지 기초적인 수준에 머물러 있다.

외부 정보를 획득하고 자율판단을 하기 위하여 경비로봇, 청소로봇 등의 지능형 로봇에게 초음파 센서나 적외선 센서 혹은 시각장치 등의 센서들이 거의 필수적으로 사용된다. 이들 중 시각장치는 시야에 존재하는 특정 장애물 또는 보행자 등의 움직임을 판단하고 3차원적 위치정보를 검출함으로써 이동 로봇의 경로설정 및 판단의 효율을 증가시킬 수 있다는 점에서 높이 평가되고 있다 [1-2].

시각장치를 이용한 장애물 인식 혹은 거리 추정

등 로봇 비전에 대한 연구는 오래전부터 현재까지 지속되어져 왔다. 로봇 비전의 궁극적 목표는 인간의 시각 시스템의 기능을 로봇에 구현시키는데 있으며, 인간의 시각 시스템과 같은 양안 시차에 의해 얻어진 두 영상으로부터 효과적으로 3차원 입체 정보를 추출해 냄으로써 표적물체와의 거리등의 정보를 찾아내는 것이 기본적인 과제라 할 수 있다 [3-5].

로봇 비전 분야에서 3차원 정보 추출을 위한 기술은 크게 능동형 센서와 수동형 센서를 이용하는 방법이 있다 [6]. 능동형 센서를 이용하는 방법은 실시간 3차원 정보 획득이 가능하고 수동형에 비해 정확한 깊이 정보를 얻을 수 있다는 장점이 있지만, 값이 비싼 하드웨어가 별도로 필요하고, 목표가 이동시 적용이 힘들다. 이에 비해 수동형 센서는 크게 Shape from Shading 방법과 Shape from Stereo 방법으로 나뉜다 [2]. Shape from Shading 방법은 단일 영상 내 음영정보의 점진적인 밝기 변화를 이용하여 3차원으로 형상을 복원하는 방법이고, Shape from Stereo 방법은 두 개의 영상을 획득하여 이들을 정합하고, 좌우측 입력 영상의 이동 물체가 카메라의 중심으로부터 떨어진 거리, 스테레오 시차 및 주시각과 삼각측량법으로 3차원 깊이를 정보를 획득하는 방법으로 정지 영상과 동영상에 모두 적용할 수 있으며 조밀한 깊이 정보를 생성할 수

* 교신저자(Corresponding Author)

논문접수 : 2008. 05. 08., 채택확정 : 2008. 07. 02.

김승환 : 전북대학교 전자공학과

함운철 : 전북대학교 전자공학과

※ 이 연구에 참여한 연구자는 2단계 BK21사업의 지원비를 받았음

있다. 이 방법은 로봇 비전 분야에서 보편적으로 사용된 방법이며, 현재에도 중요한 위치를 차지하고 있다 [7-8].

본 논문에서는 저가형인 스테레오 시스템의 구현을 위하여 주변에서 흔히 사용되는 웹 캠을 이용하여 이를 평행 식으로 구성하여 목표물체의 위치 정보를 획득하였다. 그리고 보다 정확한 거리를 계산하기 위한 웹 캠의 파라미터들과 목표물체의 중심점을 추정하는 방법에 대한 새로운 알고리즘을 제안하고 대상 목표물체를 정의하여 웹 캠을 통해 얻는 좌, 우 영상에서의 물체의 영상 좌표를 이용하여 물체의 궤적을 임계값을 이용하여 영상 쌍을 보정한 후 라벨링 기법으로 물체를 배경과 분리하고 삼각 측량법을 이용하여 목표물체와의 거리 추정을 하였다.

II. 스테레오 비전 알고리즘

사람의 두 눈이 사물을 볼 때 하나의 사물은 각각 다른 시점에서 받아들여지게 되는데, 이 영상 쌍을 적절히 정합(Matching)하여 근소한 차이를 이용하여 거리와 부피감을 만들게 된다. 이렇게 두 눈에서 각각 받아들여진 한 쌍의 그림을 3차원의 관점으로 융합하는 것을 스테레오 비전이라고 한다[9]. 스테레오 영상에는 좌우로 구성된 두 장 이상의 평면 영상이 필요하며, 이들 영상은 한 대의 카메라를 사용하여 획득할 수는 있으나, 일반적으로 동일한 초점거리를 갖는 두 대 이상의 카메라를 이용하여 얻는다. 두 대의 카메라를 구성하는 방법에는 '평행식'과 '교차식' 카메라 시스템이 있다. '평행식'은 거리를 측정하는데 주안점을 두는 방식이고, '교차식'은 3D 입체감을 중시하는 방식으로, 본 논문에서는 거리 측정에 주안점을 두고 있어 평행식 카메라 시스템을 이용하여 거리를 추정하였다.

평행식 카메라 시스템에서 좌, 우 카메라가 광축과 평행하지 않게 배치되면 변위 계산에 있어서 오차가 발생할 확률이 높고 탐색 범위가 증가하게 된다. <그림 1>은 카메라를 최대한 동일한 y, z 좌표 값을 갖게 하고, x 축으로 d 만큼 차이가 개 배치한 배치도이다. 단 여기에서 추정하고자 하는 것은 물체와 두 카메라와의 평면상의 수직 거리를 구하고자 하므로, y 점은 고려하지 않는다.

이때 두 카메라의 초점 거리를 f , 카메라 사이의 거리를 d , 3차원 상의 임의의 점을

$P(x_w, y_w, z_w)$, 점 P 에 위치한 물체가 각 카메라의 영상 면에 투영된 영상 좌표를 각각 (x_1, y_1) , (x_2, y_2) 로 정의하면 삼각법에 의하여 다음과 같은 식을 도출할 수 있다.

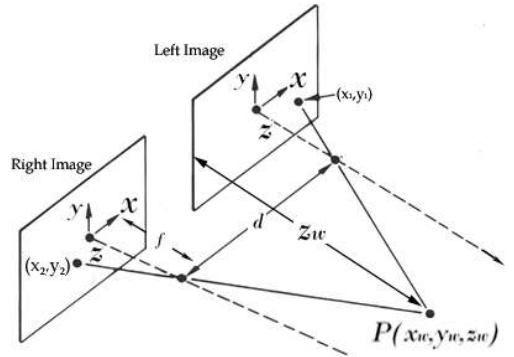


그림 1. 거리측정을 위한 삼각형 기하학
Fig. 1. Triangulation for Distance Measurements

만약 x_1, x_2 에서 각 카메라의 광축까지의 길이의 합을 x_t , 두 영상에서 물체까지의 수직거리를 z_w 라 하며, 각 카메라의 광축에서 물체까지의 수직 거리의 합을 x_d 라 하면 식 (1)을 알 수 있다.

$$\frac{x_t}{f} = \frac{x_d}{z_w - f} \tag{1}$$

다시 카메라의 왼쪽과 오른쪽 광축에서 물체까지의 거리를 x_{d1}, x_{d2} 라 하고, 광축에서 카메라에 각각 투영된 영상의 좌표까지의 거리를 x_{t1}, x_{t2} 라 하고 식 (1)을 각 점과 영상에 적용하면, 식 (2)을 도출할 수 있다.

$$x_{d1} = \frac{x_{t1}}{f}(z_{w1} - f), x_{d2} = \frac{x_{t2}}{f}(z_{w2} - f)$$

(단, $x_{d1} + x_{d2} = d, z_w = z_{w1} = z_{w2}$) (2)

식 (2)에 x_{d2} 대신 $d - x_{d1}$ 를 대입하여 계산하면 식 (3)을 얻을 수 있다.

$$z_w = f + \frac{f d}{x_{t1} + x_{t2}} \quad (3)$$

두 카메라가 z 축을 기준으로 완벽히 평행하게 배치되었을 경우 식 (3)은 정확한 물체와의 거리를 출력한다. 하지만 만약 카메라를 설치할 시 카메라가 기울어 평행이 깨지는 등의 오차가 생길 경우 거리 값의 신뢰도가 떨어지게 된다. 그래서 이에 대한 보정이 필요한데, 이때 각 카메라에 대해 초점거리를 보정하게 되면, 계산이 복잡해지고 삼각법을 적용하기 힘들어진다. 그래서 본 논문에서는 거리 값을 미리 정한 점 n 개를 이용하여 두 웹 캠의 초점거리의 근사치로 동일하게 통일하여 추정하는 방법을 제안한다.

3차원 상의 점 P 의 좌표를 임의로 n 개를 선택하고 다음과 같은 방식으로 초점거리의 참값의 근사치를 도출한다.

식 (1)의 f 를 오차가 포함된 $f + f_e$ 로 치환하면

$$\frac{x_t}{f + f_e} = \frac{x_d}{z_w - (f + f_e)} \quad (4)$$

식 (4)에서 계산의 편의를 위하여 $(z_w - (f + f_e))$ 를 z_t 라 치환하면

$$x_t z_t = x_d (f + f_e) \quad (5)$$

식 (5)를 도출한 후, n 개의 점에 적용을 시키면 다음과 같은 결과를 갖게 된다.

$$\begin{bmatrix} x_{t1} z_{t1} \\ x_{t2} z_{t2} \\ \vdots \\ x_{tn} z_{tn} \end{bmatrix} = \begin{bmatrix} x_{d1} \\ x_{d2} \\ \vdots \\ x_{dn} \end{bmatrix} (f + f_e) \quad (6)$$

식 (6)에서 초점거리 f 의 참값에 최대 근사치 f_t 를 알기 위하여 Least Squares 알고리즘을 사용하여 식 (7)을 도출한다.

$$f_t = \left(\begin{bmatrix} x_{d1} \\ x_{d2} \\ \vdots \\ x_{dn} \end{bmatrix} \right)^T \begin{bmatrix} x_{d1} \\ x_{d2} \\ \vdots \\ x_{d3} \end{bmatrix}^{-1} \begin{bmatrix} x_{d1} \\ x_{d2} \\ \vdots \\ x_{dn} \end{bmatrix} \begin{bmatrix} x_{t1} z_{t1} \\ x_{t2} z_{t2} \\ \vdots \\ x_{tn} z_{tn} \end{bmatrix} \quad (7)$$

III. 물체의 거리측정을 위한 영상 처리

1. 좌우 영상의 밝기 값 보정

카메라를 통해 획득된 영상은 주변 환경 및 카메라의 상태에 따라 밝기의 변화를 가질 수 있다. 스테레오 영상의 경우, 두 영상의 밝기차이에 따라 이진화에 오차가 생겨 목표 물체 이외에 다른 영역도 대상이 되어 오류를 범하게 될 가능성이 있다. 이는 라벨링을 하여 목표 물체의 범위를 구하는데 중요한 요인이 된다. 이런 문제를 해결하기 위하여 본 논문에서는 좌, 우 영상의 밝기 평균 차를 이용하여 다음과 같이 보정을 한다.

먼저 식 (8)과 같이 좌, 우 영상의 밝기 평균 차를 구한다.

$$I_l = \frac{1}{N_x N_y} \sum_{y=0}^{N_y-1} \sum_{x=0}^{N_x-1} (L_l(x,y) - L_r(x,y)) \quad (8)$$

단 여기서 $L_l[x,y], L_r[x,y]$ 는 좌, 우 이미지의 밝기 값,

N_x, N_y 는 가로, 세로 픽셀 수



그림 2. 영상에 임계값 적용

Fig. 2. Image Corrections by Using Thresholds

이제 식 (8)에서 얻은 값을 영상에 적용시킨다.

$$\begin{aligned} L_l(x,y) &= L_l(x,y) + I_l, \\ L_r(x,y) &= L_r(x,y) \end{aligned} \quad (9)$$

식 (9)의 결과로 좌, 우 영상의 밝기 값이 비슷하여지고, 라벨링을 통한 보다 정확한 관심물체 추출을 할 수 있다.

그림 2의 위쪽 두 이미지는 보정 이전의 이미지

이고, 하단의 두 이미지는 보정된 이미지로, 밝기가 비슷해졌음을 알 수 있다.

2. 관심 물체의 영역 추출

카메라로부터 얻어지는 좌우 영상에서 원하는 물체까지의 거리를 알기 위해서는 해당 물체의 중심점을 알아낼 필요가 있으므로 두 영상에서 관심이 있는 물체가 있는 영역을 배경에서 추출해야 한다.

관심 영역을 추출하기 위해 좌우 영상을 3.1에서 얻은 임계값을 적용하여 이진 영상으로 변환한 후, 물체 영역을 라벨링(labeling)을 통해 각각의 물체를 구성하는 개개의 픽셀들을 하나의 영역으로 묶는다. 라벨링이란 인접하여 연결되어 있는 화소들에 특정 번호를 붙이고, 이와는 분리되어 있고 다른 연결된 화소들에는 다른 번호를 붙여서 영역을 구별하기 위한 알고리즘이다. 라벨링된 영상을 이용하면 관심 영역에 대해서만 크기 및 중심점을 계산해 내는 것이 가능하다.

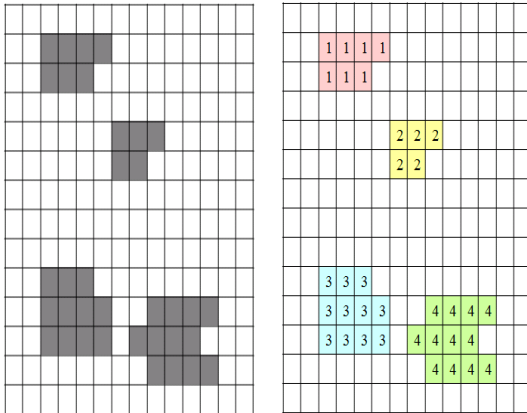


그림 3. 이진영상의 라벨링

Fig. 3. Labeling of Binary Images

본 논문에서 사용된 라벨링 알고리즘은 Glassfire 알고리즘이다. 이 알고리즘은 잔디에서 불이 번져가는 모양과 수행방식이 유사하다고 하여 붙여진 이름으로, 재귀호출(recursive call)을 이용하여 모든 인접화소가 라벨링 될 때까지 현재 관심화소의 주변 인접화소를 차례로 검사하면서 라벨링 하는 방법이다[10]. 라벨링을 수행한 후, 3.3 절에서 Glassfire 알고리즘과 본 논문에서 제안하는 가중치를 적용하는 방식을 통해 목표 물체의 중심점을 추정한다.

3. 관심 물체의 중심점 추정

본 논문에서 중심점을 추출하기 위해서 제안된 알고리즘은, 라벨링 된 각 물체를 이루는 픽셀에 해당 픽셀에 근접한 유효 픽셀의 수를 가산치로 주어 보다 노이즈에 영향을 줄이는 방법이다.

먼저 각 라벨링 된 영상을 Glassfire 알고리즘을 다시 사용하여 기존에 라벨 값들을 <그림 4>와 같이 해당 픽셀 주변의 유효 픽셀의 합으로 갱신한다.

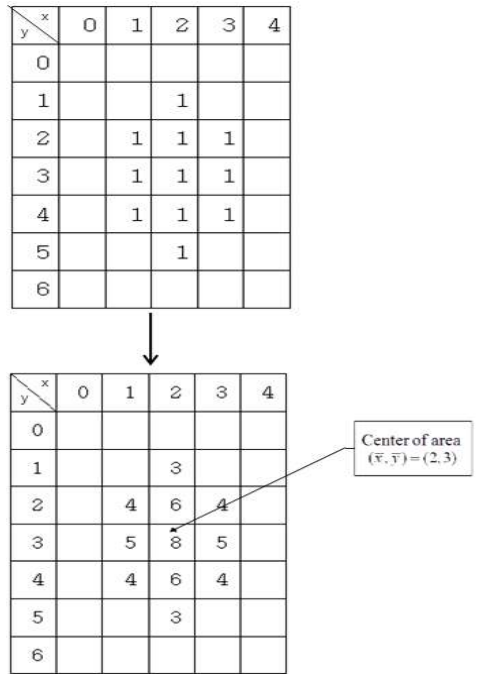


그림 4. 가중치의 적용 예

Fig. 4. Example for Applying Weights



그림 5. 관심영역 추출 및 중심점 계산

Fig. 5. Detect Region of Interest and Calculate the Center Point

그 후, 가산치의 총합을 구하여 식 (10)의 식을 이용하여 목표 물체의 중심점 즉 기하학적 중심을 구할 수 있다.

이제 캠으로 좌, 우 영상을 획득하여 라벨링을 하고 중심점을 <그림 5>처럼 찾은 뒤, 식 (3)을 이용하여 목표 물체와의 거리를 측정한다.

$$\bar{x} = \frac{\sum_{x=0}^{N_x-1} (x \sum_{y=0}^{N_y-1} B[x,y])}{A}, \quad (10)$$

$$\bar{y} = \frac{\sum_{y=0}^{N_y-1} (y \sum_{x=0}^{N_x-1} B[x,y])}{A}$$

단 여기서 $B[x,y]$ 는 가중치 배열,
 A 는 가중치의 총합,
 N_x, N_y 는 가로, 세로 픽셀 수

IV. 깊이 추정 결과

본 실험에서 스테레오 비전 시스템을 이용하여 거리를 측정하기 위해 실험에 사용된 두 웹 캠의 사이 간격은 10cm이다. 그리고 초점거리는 7번의 실험을 통하여 10.5mm로 추정을 하였다. 해상도는 보다 빠른 영상처리를 위하여 320×240로 설정하였고, 카메라의 화소의 크기는 6.35×7.4um 이며, 거리 측정은 실시간으로 하였다. 그리고 목표 물체만의 거리 측정을 위하여 물체는 하나만 있다고 가정하였다.

그림 6은 320×240 해상도의 실험 입력 영상 및 처리 영상이다. 위쪽 두 영상은 실시간으로 입력되는 실험용 좌, 우 입력영상이다. 그리고 밑의 두 영상은 입력 영상의 그레이화, 이진화와 라벨링 처리를 하여 중심점을 찾고, 좌표 및 거리를 추정한 결과이다.

표1은 물체를 이동시켜서 그림6의 처리 영상에서 얻어진 양쪽의 x, y 좌표의 표본 값 5개 이다. 앞에서 언급했듯, 실험에서 추정하고자 하는 것은 물체와 두 카메라와의 평면상에 수직 거리를 구하고자 하는 것이기 때문에 y 점이 모두 동일한 곳에 물체를 두어 측정하였으나, 입력 영상의 처리 결과 약간의 오차가 생겼다. 하지만 오차의 폭이 크기 않았기 때문에 이를 고려하지 않고 결과를 도출했다.

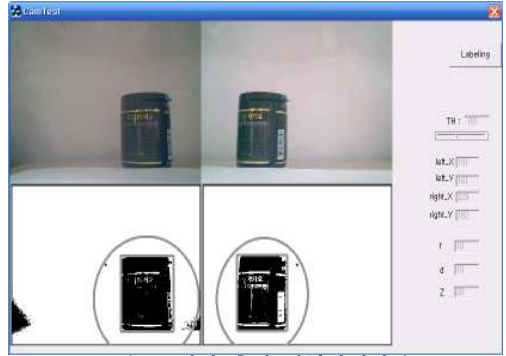


그림 6. 거리 추정 어플리케이션

Fig. 6. Application of Distance Measurements

표 1. 획득 영상의 위치좌표 값 [단위: pixel]
 Table 1. Position Coordinates of Images

횟수	우측		좌측	
	x_{l1}	y_{l1}	x_{r1}	y_{r1}
1	81	120	171	119
2	97	118	172	121
3	160	121	220	120
4	112	120	165	123
5	164	120	182	121

표 2. 물체까지의 실측거리와 추정거리 [단위: cm]
 Table 2. Actual Distances and Estimated Distances between the Object and Web Cams

횟수	실측 거리	추정 거리	오차(%)
1	25	23.64	5.75
2	30	28.37	5.43
3	35	35.46	1.31
4	45	49.49	9.97
5	100	118.2	18.2

표2는 표1에서 얻어진 값들과, 삼각 측량법에 의해 얻어진 추정 거리와, 실제 물체와 웹 캠들 사이의 거리를 1m 범위에서 측정한 결과이다. 근거리의 30번의 실험 가운데, 여기에서 얻어진 오차들의 평균값은 8% 이다. 그리고 오차들 중 크게 차이가 나는 값은, 빛에 따른 오차에 의해 라벨링을 통한 중심점의 추정이 크게 빗나간 경우, 실측 거리

값을 쥘 때 생기는 오차 등으로 발생하는 것으로 추정된다.

그리고 웹 캠을 이용한 원거리 측정은 웹 캠의 낮은 성능과 빛에 강한 영향을 받는 등의 이유로 오차가 굉장히 크기 때문에 이에 대한 보정 알고리즘이 필요하다.

V. 결 론

본 논문에서는 웹 캠을 평행식 스테레오 시스템으로 구성하여 좌, 우 영상을 획득한 후, 두 영상의 밝기 차이를 보정하여 보다 정밀한 라벨링을 구현하였고, 이를 이용하여 목표 물체를 배경과 분리시키고, 중심점을 계산하여 목표 물체와 웹 캠까지의 거리를 구하는 방법을 구현하였다. 그리고 중심점을 구하기 위한 가중치 적용 방법과 삼각측량법을 이용하기 위하여 웹 캠의 초점거리를 구하고, 이에 대한 오차를 보정하는 방법을 제안하였다.

제안한 시스템의 실험 결과, 불확실한 카메라의 초점거리 등과 같은 파라미터를 보정하여 목표 물체까지의 1m 이내 거리 내에서 비교적 정확하게 측정이 가능함을 확인하였다. 또한 라벨링 시 가중치를 적용하면 영상의 노이즈에 영향을 덜 받는 것을 알 수 있었다. 그리고 거리에 대한 오차가 평균 8% 존재함을 확인할 수 있었는데, 이는 웹 캠의 낮은 성능 및 초점거리 오차에, 픽셀 대 각도의 대응 오차, 장비 구성 시 생긴 오차, 빛 등에 따른 것으로 볼 수 있다. 정밀한 거리 측정을 위해서는 이러한 오차들을 줄이는 연구가 필요하다.

향후, 더욱 정밀한 거리 측정을 위하여 이러한 오차들을 줄이는 방안을 연구하고, 웹 캠을 통해 얻은 이미지에 대한 보정방식을 최적화 시키고 차후, 웹 캠을 더 추가하여 다중 시점을 이용한 거리측정에 대한 연구를 진행할 예정이다.

참고문헌

- [1] James Black, Tim Ellis, "Multi-camera image measurement and correspondence", Measurement, vol. 32, pp.61-71, 2002.
- [2] T. Kanade, A. Yoshida, K. Oda, H. Kano, and M. Tanaka, "A stereo machine for video-rate dense depth mapping and its new applications", Proc. of IEEE CVPR '96, pp. 196-202, 1996.
- [3] David Coombs and Christopher Brown, "Real-time binocular smooth pursuit", International Journal of Computer Vision, vol.11, no. 2, pp.147-164, 1993
- [4] D. J. Coombs and C. M. Brown, "Cooperative gaze holding in binocular vision", IEEE Control System Magazine, vol. 11, pp.24-33, June 1991
- [5] Priyan Victor S. Grinberg, Gregg Podnar and M. W. Siegel, "Geometry of binocular imaging", Proc. SPIE Internat. Conf. on Stereoscopic Displays and Virtual Reality Systems, Vol 2177, San Jose, CA, 6-10, pp. 56-65, February 1994
- [6] P. Dias, V. Sequeira, F. Vaz, and J. G. M. Goncalves, "Registration and fusion of intensity and range data for 3D modeling of real world scenes", Proceedings of 3-D Digital Imaging and Modeling, pp. 418-426, 2003
- [7] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", International Journal of Computer Vision, 47 (1-3), pp.7-42, 2002
- [8] Z. Ruo, T. Ping-sing, E.C. James, and S.Mubarak, "Shape from shading : A survey", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 21, no. 8, pp. 690-706, 1999
- [9] U.R. Dhond and J.K. Aggrawal, "Structure from stereo-A review", IEEE trans, System. Man and Cybern., Vol. 19, pp. 1489~510, 1989.
- [10] Ramesh Jain, Rangachar Kasturi, Brian G. Schunck, "Machine Vision", McGRAW-HILL INTERNATIONAL EDITIONS, 1995

저 자 소 개

김 승 환 (Seunghwan Kim)



2007년 2월 전북대학교
전자정보공학부 학사.
2007년 ~ 현재 전북대
학교 전자공학과 석
사 과정.

관심분야: WinCE, 영상처리, 마이크로프로세서.
Email: melodyoflove@nate.com

함 운 철(WoonChul Ham)

1979년 서울대학교 전

자공학과 학사 졸업

1982년 서울대학교 전

자공학과 석사 졸업

1988년 서울대학교 전

자공학과 박사 졸업

1982년 ~ 현재 전북대

학교 전자정보공학부

교수

관심분야: 적응제어, 로봇제어

Email: wcham@chonbuk.ac.kr