

논문 2008-45SP-4-4

다중 어안 카메라를 이용한 움직이는 물체 검출 연구

(A Study on Detecting Moving Objects using Multiple Fisheye Cameras)

배 광 혁*, 서 재 규*, 박 강 령**, 김 재 희***

(Kwanghyuk Bae, Jae Kyu Suhr, Kang Ryoung Park, and Jaihie Kim)

요 약

기존의 보안·감시 시스템은 화각이 좁은 일반 렌즈를 주로 사용하여 천정이 낮고 실내가 넓은 환경에 적용하기가 어려웠다. 이를 해결하기 위해 단순히 카메라의 수를 늘리는 방법은 비용의 증가와 설치의 어려움 등의 문제가 있다. 따라서 본 논문에서는 화각이 180도인 어안 카메라를 다수 설치한 사용자 감시 시스템을 제안하였다. 단일 어안 카메라에서 물체 간의 교차에 의한 가림현상이 발생되는 문제를 해결하기 위해서 카메라 간의 상동관계를 다중 어안 카메라 시스템에 적용하였다. 17×14 m의 공간의 2.5 m 높이에 설치된 4대의 어안 카메라에서 5명이 서로 교차하면서 움직이도록 하여 수행한 결과, 단일 어안 카메라에서 최대 46.1% 낮은 검출율을 보인 반면 제안된 시스템에서 83.0%로 향상된 성능을 보였다.

Abstract

Since vision-based surveillance system uses a conventional camera which has a narrow field of view, it is difficult to apply it into the environment whose the ceiling is low and the monitoring area is wide. To overcome this problem, the method of increasing the number of camera causes the increase of the cost and the difficulties of camera set-up. For these problems, we propose a new surveillance system based on multiple fisheye cameras which have 180 degree field of view. The proposed method handles occlusions using the homography relation between the multiple fisheye cameras. In the experiment, four fisheye cameras were set up within the area of 17×14 m at height of 2.5 m and five people wandered and crossed with one another within this area. The detection rates of the proposed system was 83.0% while that of a single camera was 46.1%.

Keywords : Surveillance, Multiple Fisheye cameras, Homography

I. 서 론

최근 실내에서 범죄 감시와 이동경로 추적, 특정지역의 이동량 조사 등을 위한 카메라 기반의 영상감시 시

* 학생회원, ** 평생회원, 연세대학교 전기전자공학과, 생체인식연구센터

(School of Electrical and Electronic Engineering, Yonsei University, Biometrics Engineering Research Center)

** 정회원, 동국대학교 전자공학과, 생체인식연구센터

(Dept. of Electronics Engineering, Dongguk University, Biometrics Engineering Research Center)

※ 본 연구는 2002년도 정부(과학기술부)의 재원으로 한국과학재단 지정 생체인식연구센터의 지원을 받아 이루어졌습니다. (R112002105070020(2008))

접수일자: 2008년2월29일, 수정완료일: 2008년5월7일

스템의 설치가 증대되고 있다. 특히, 지하철, 주차장, 백화점 등과 같이 천정이 낮고 실내가 넓은 공간에서 카메라가 설치되고 있다. 이러한 장소에서 단일 카메라를 이용한 영상감시 시스템은 카메라가 획득할 수 있는 시각적 범위가 한정되어 있어 광범위한 지역에서 사람의 움직임이나 행동패턴의 추적이 쉽지 않다. 또한 단일 카메라를 사용할 경우 움직이는 물체간의 교차에 의한 가림(occlusion) 현상이 발생하여 물체를 분리해 내는 작업이 어렵게 된다.

보다 넓은 공간을 이동하는 사람이나 물체를 감시하기 위해 화각이 넓은 전방향(omnidirectional) 카메라를 이용하여 움직이는 물체를 추적하거나^[1~3] 여러 대의 카메라를 넓은 범위에 걸쳐 설치하는 방법이 활발하게

연구되고 있다^[4~9]. 단일 카메라에서 움직임을 모니터링하는 감시자의 역할이 다중 카메라로 확대될 경우 여러 영상을 한 번에 모니터링 해야 하는 부담이 증가하며, 한 카메라에서 움직이는 물체가 다른 카메라에서 어떻게 연관이 되는지 알기가 힘들다는 단점이 있다. 이 같은 문제점을 해결하기 위하여 최근 다중 카메라를 이용한 영상감시시스템에서는 카메라 사이의 관계를 알아내어 동일한 물체의 움직임을 추적하는 연구와 카메라 사이의 정보 교환에 의한 가림 현상의 해결에 관한 연구들이 진행되고 있다^[4~9]. 하지만 기존 방법들^[4~9]에서는 화각이 좁은 일반 렌즈를 사용하여 천정이 낮고 넓은 실내 환경에서 적용하기가 어렵다.

본 논문에서는 천정이 낮고 넓은 실내 환경에서 적용이 가능한 다중 어안 카메라를 이용한 영상감시 시스템을 제안하였다. 본 논문은 구성은 다음과 같다. II장에서는 다중 카메라를 이용한 영상감시 시스템에서 렌즈의 선택에 관하여 설명한다. III장에서는 본 논문에서 제안된 다중 어안 카메라 시스템에서 카메라들 사이의 상동관계에 의한 움직이는 물체 검출 방법을 사전 처리부와 실시간 처리부로 나누어 설명한다. IV장에서는 천정이 낮고 넓은 실내 공간에서 제안된 시스템을 실험 및 평가하였다. 마지막으로 V장에서는 결론과 향후 연구과제에 대해서 논의한다.

II. 다중 카메라를 이용한 영상감시 시스템

다중 카메라를 사용하여 가림현상을 해결하기 위해서는 그림 1과 같이 각 카메라의 화각을 겹치게 배치하여 공통된 영역이 존재하도록 해야 한다. 천정이 낮고 넓은 실내 환경에서 설치된 일반 렌즈[그림 1에서 화각 ①과 화각②, 그림 2(a),(b)]를 사용할 경우에 그림에서 보이는 것처럼 공통된 부분이 좁아 넓은 영역을 감시할 수가 없다. 광각 렌즈[그림 1에서 화각③, 그림 2(c)]의 경우에 카메라를 기울이면 그림과 같이 공통영역이 넓어지게 된다. 하지만 화각을 겹치게 배치하기 위해 카메라 방향의 조절이 필요하다. 어안렌즈[그림 1에서 화각④, 그림 2(d)]의 경우 카메라의 광축을 바닥면에 수직으로 설치하면 수평면상의 전 방향에서 감시가 가능하게 되어, 카메라 사이에 공통영역이 항상 존재하게 된다. 또한 어안 카메라를 물체보다 높이 설치할 경우 감시 영역에서 물체가 화각에서 벗어나게 되는 사각영역이 없게 된다.

기존에 다중 카메라를 사용한 영상감시 시스템에서

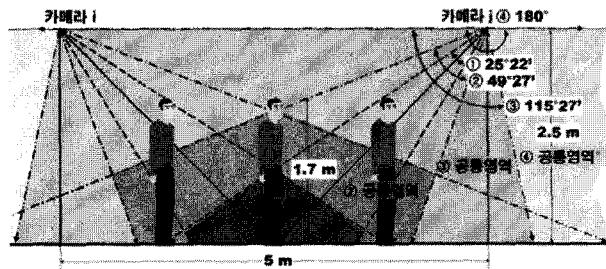
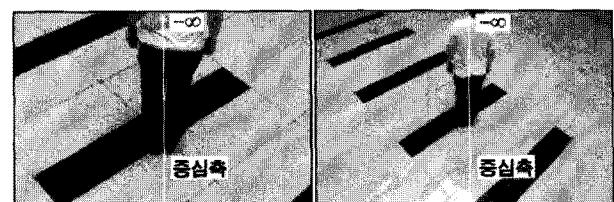


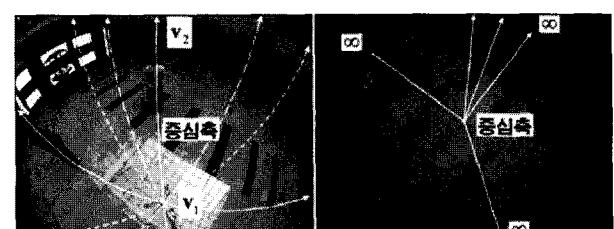
그림 1. 천정이 낮은 실내 환경에서 렌즈의 화각에 따른 공통 영역의 비교 [① 일반 렌즈 (화각: 33° × 25°), ② 일반 렌즈 (화각: 65° × 49°), ③ 광각 렌즈 (화각: 154° × 115°), ④ 어안렌즈 (화각: 180° × 180°)]

Fig. 1. Comparison of the field of views (FOVs) in indoor environment [① general lens (FOV: 33° × 25°), ② general lens (FOV: 65° × 49°), ③ wide angle lens (FOV: 154° × 115°), ④ fisheye lens (FOV: 180° × 180°)]



(a)

(b)



(c)

(d)

그림 2. 실제 영상에서의 화각과 중심축 비교 (a) 일반 렌즈 ①, (b) 일반 렌즈 ②, (c) 광각 렌즈 ③, (d) 어안렌즈 ④

Fig. 2. Comparison of the FOVs and principal axis (a) general lens ①, (b) general lens ②, (c) wide angle lens ③, (d) fisheye lens ④.

카메라 사이의 대응관계를 찾는 방법으로는 영역기반의 방법과 점기반의 방법이 있다^[5]. 영역기반의 방법은 한 카메라에서 움직이는 물체의 영역과 다른 카메라에서의 물체 영역간의 유사도를 영상 그레이 값, 색정보 등을 기반으로 찾는 방법이다. 따라서 넓은 영역을 감시하기 위해 카메라사이의 거리를 멀게 설치하거나 같은 사물을 다른 방향에서 바라보게 설치하게 되면 카메라 사이의 공통 영역에 대한 유사도가 떨어지게 된다는 단점을

갖는다. 점기반 방법은 삼차원 공간에서 카메라 사이의 위치관계를 미리 파악한 후 극상선(epipolar line) 제약 조건을 이용한다. 하지만 실제 사용에서 설치된 카메라 간의 정확한 관계를 얻어내는 것이 어렵게 된다.

이러한 단점을 극복하기 위하여 최근에는 카메라 사이의 상동관계(homography)를 이용한 방법들이 연구되고 있다^[5~9]. 상동관계에 의한 영상감시 방법은 화각 내의 공통된 영역에 존재하는 평면과 물체의 중심축을 이용한다. 기존에 어안 렌즈가 아닌 화각이 좁은 일반 렌즈를 사용한 시스템^[5~8]에서는 움직이는 물체의 중심축이 형성하는 두 소실점(vanishing points)이 영상에서 그림 2(a),(b)와 같이 y축상의 $\pm\infty$ 에 존재하여 중심축은 y축과 평행한 직선에서 찾게 된다.

광각 렌즈[그림 1에서 화각③, 그림 2(c)]의 경우에 그림과 같이 카메라를 기울이면 공통영역이 넓어지게 된다. 하지만 렌즈의 화각이 커질수록 반지름 방향의 왜곡이 커지고, 광각 렌즈를 그림 1과 같이 기울여 설치할 경우에 그림 2(c)에서 보이는 것처럼, 물체의 중심축은 두 소실점을 지나는 곡선이 되다. 이러한 이유로 광각 렌즈를 사용할 경우에 설치할 때마다 소실점을 다시 구해야 한다는 문제점이 있다.

반면 어안 렌즈의 경우에 바닥면과 수직으로 설치하면 물체의 중심축이 광축과 평행하게 되어 하나의 소실점은 반지름 방향의 $+\infty$ 에 존재하고 다른 하나는 영상의 주점(principal point)에 위치한다. 따라서 반지름 방향의 직선위에서 움직이는 물체를 검출하게 된다. 따라서 천정이 낮고 넓은 실내 환경에서 가림현상에 강인한 감시 시스템으로는 넓은 공통영역을 갖고 상동관계를 쉽게 적용하여 가림현상을 해결할 수 있는 다중 어안 카메라 시스템이 적합하다.

III. 다중 어안 카메라를 이용한 움직이는 물체 검출 방법

제안된 다중 어안 카메라를 이용한 영상감시 시스템은 광축이 바닥면과 수직하도록 천정에 설치된 $N(>2)$ 개의 어안 카메라로 구성되고 이 시스템의 전체 흐름도는 그림 3과 같다. 전체 시스템은 카메라 설치 후에 카메라 사이의 관계를 구하는 사전 처리부와 단일 어안 카메라에서 움직이는 물체를 검출한 후에 전체 카메라의 정보를 융합하여 가려진 물체를 구별하고 색인하는 실시간 처리부로 구성된다.

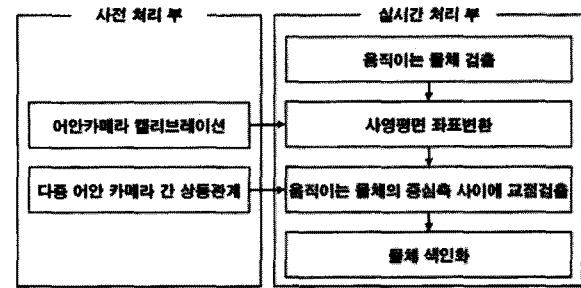


그림 3. 전체 시스템 흐름도

Fig. 3. Block diagram of the overall system.

1. 사전 처리부

다중 어안 카메라 설치 후 어안 카메라의 내부변수와 각 카메라 사이의 상동관계를 미리 구한다.

가. 어안 카메라의 캘리브레이션

어안 카메라 영상에서 반지름 방향의 왜곡모델과 카메라 내부변수(초점거리, 주점, 종횡비)를 추정하여 입력된 영상의 좌표와 원근 투영(perspective projection) 좌표 사이의 변환관계를 알아낸다. 원근 투영 시의 반지름 방향 길이 r_u 는 식(1)과 같이 그림 4의 초점거리 f 와 $\tan\theta$ 함수의 곱 형태로 표현된다^[10].

$$r_u = f \tan \theta \quad (1)$$

어안 렌즈의 구면수차에 의해서 왜곡 된 반지름 방향 길이(r_d)는 식(2)와 같이 무한급수로 표현이 가능하다^[10].

$$r_d(\theta) = k_1 \theta + k_2 \theta^3 + k_3 \theta^5 + k_4 \theta^7 + k_5 \theta^9 + \dots \quad (2)$$

본 논문에서는 식 (2)의 왜곡함수 r_d 에 대해 9차방정식을 사용하였고, Brandt의 방법을 이용하여 어안 카메라의 왜곡변수(k_1, \dots, k_5)와 내부 변수(초점거리, 주점, 종횡비)를 추정하였다^[10]. 입력받은 영상의 좌표 x 에서 r_d 와 ϕ 를 구하고 식 (2)에 의해 θ 를 구하면 원근 투영 좌표는 식(3)과 같이 표현된다.

$$\hat{x} = \begin{pmatrix} x_u \\ y_u \end{pmatrix} = f \tan \theta \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad (3)$$

원근 투영된 좌표 \hat{x} 는 계산편의를 위해 사영평면(projective plane)의 좌표 $\tilde{x} = (x_u, y_u, 1)^T$ 를 사용한다. 그림 4(b)는 캘리브레이션을 위한 평면 패턴 영상과 그 영상을 추정된 왜곡변수로 보정한 결과를 보여준다.

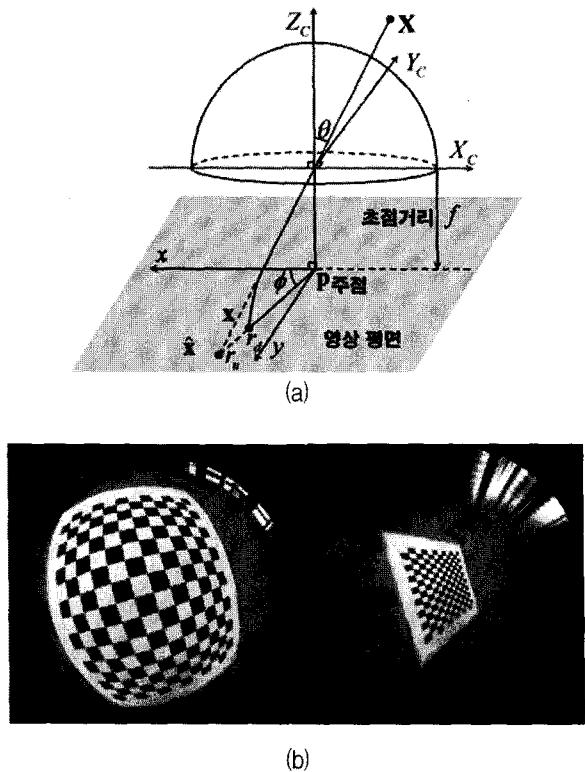


그림 4. (a) 어안 카메라 모델, (b) 캘리브레이션을 위한 패턴 평면 영상과 왜곡 보정 영상
Fig. 4. (a) Fisheye camera model, (b) calibration pattern plane image and its undistortion image.

나. 다중 카메라 간 상동관계

일반적으로 편홀카메라의 두 영상 간의 관계는 하나의 행렬로 표현할 수 있다. 만일 그림 6(a)와 같이 두 영상에 실세계의 동일평면이 촬영되면 이 행렬은 상동관계를 나타내게 된다. 그림 6(a)에서 잠시 영역에 존재하는 바닥면의 한 점 \$Y\$가 두 어안 카메라(카메라 \$i\$, 카메라 \$j\$)로 투영된 점을 각각 \$\tilde{y}_m^{(i)}\$, \$\tilde{y}_m^{(j)}\$이라고 하고, 식 (3)에 의해서 반지름 방향의 왜곡을 보정한 좌표를 사영 평면에서 \$\tilde{y}_m^{(i)}\$, \$\tilde{y}_m^{(j)}\$이라 하면, 두 점 사이의 관계는 식 (4)과 같게 된다.

$$\tilde{y}_m^{(i)} = \mathbf{H}_{ij} \tilde{y}_m^{(j)} \quad (m=1,2,\dots,k) \quad (4)$$

여기서, 행렬 \$\mathbf{H}_{ij}\$를 상동관계라고 하고, \$3 \times 3\$의 행렬이 된다. 이 행렬은 바닥위의 공통된 4개의 점 이상이 주어지면 Direct Linear Transformation (DLT)에 의해 다음과 같이 구할 수 있다^[13]. 바닥면 위에 \$k\$ 개의 대응쌍 \$\tilde{y}_m^{(i)} = (\tilde{x}_m^{(i)}, \tilde{y}_m^{(i)}, 1)^T\$, \$\tilde{y}_m^{(j)} = (\tilde{x}_m^{(j)}, \tilde{y}_m^{(j)}, 1)^T\$이 주어졌을 때, 식 (4)의 상동관계에서 두 벡터 \$\tilde{y}_m^{(i)}\$와 \$\mathbf{H}_{ij} \tilde{y}_m^{(j)}\$는 동일해야

하므로 식 (5)로 표현된다.

$$\tilde{y}_m^{(i)} \times \mathbf{H}_{ij} \tilde{y}_m^{(j)} = 0 \quad (5)$$

여기서 \$\mathbf{H}_{ij}\$의 \$s\$ 번째 행을 \$\mathbf{h}^{st}\$라고 하면, \$\mathbf{H}_{ij}\$는 식(6)과 같이 표현되고 이를 이용해서 식(5)는 식(7)이 된다.

$$\mathbf{H}_{ij} = \begin{pmatrix} \mathbf{h}^{1T} \\ \mathbf{h}^{2T} \\ \mathbf{h}^{3T} \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \quad (6)$$

$$\begin{pmatrix} \tilde{y}_m^{(i)} \mathbf{h}^{3T} \tilde{y}_m^{(j)} - \mathbf{h}^{2T} \tilde{y}_m^{(j)} \\ \mathbf{h}^{1T} \tilde{y}_m^{(j)} - \tilde{x}_m^{(i)} \mathbf{h}^{3T} \tilde{y}_m^{(j)} \\ \tilde{x}_m^{(i)} \mathbf{h}^{2T} \tilde{y}_m^{(j)} - \tilde{y}_m^{(i)} \mathbf{h}^{1T} \tilde{y}_m^{(j)} \end{pmatrix} = \mathbf{0} \quad (7)$$

여기서 \$\mathbf{h}^{st} \tilde{y}_m^{(j)} = \tilde{y}_m^{(j)T} \mathbf{h}^s \quad (s=1,\dots,3)\$으로,

$$\begin{bmatrix} \mathbf{0}^T & \tilde{y}_m^{(j)T} & \tilde{y}_m^{(i)} \tilde{y}_m^{(j)T} \\ \tilde{y}_m^{(j)T} & \mathbf{0}^T & -\tilde{x}_m^{(i)} \tilde{y}_m^{(j)T} \\ -\tilde{y}_m^{(i)} \tilde{y}_m^{(j)T} & \tilde{x}_m^{(i)} \tilde{y}_m^{(j)T} & \mathbf{0}^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0} \quad (8)$$

가 되고 세 번째 행은 첫 번째와 두 번째 행으로 구할 수 있어 두 행만이 독립이므로 \$\mathbf{h}\$가 미지수인 식 (9)와 같은 선형 방정식으로 표현된다.

$$\mathbf{A}\mathbf{h} = \begin{bmatrix} \mathbf{0}^T & \tilde{y}_m^{(j)T} & \tilde{y}_m^{(i)} \tilde{y}_m^{(j)T} \\ \tilde{y}_m^{(j)T} & \mathbf{0}^T & -\tilde{x}_m^{(i)} \tilde{y}_m^{(j)T} \\ -\tilde{y}_m^{(i)} \tilde{y}_m^{(j)T} & \tilde{x}_m^{(i)} \tilde{y}_m^{(j)T} & \mathbf{0}^T \end{bmatrix} \begin{pmatrix} \mathbf{h}^1 \\ \mathbf{h}^2 \\ \mathbf{h}^3 \end{pmatrix} = \mathbf{0} \quad (9)$$

여기서 \$\mathbf{A}\$는 \$2 \times 9\$ 행렬이다. \$k > 4\$가 주어질 때, \$k \times 9\$ 행렬 \$\mathbf{A}\$를 SVD(singular value decomposition)하여 \$\mathbf{h}\$를 구할 수 있다. 즉, \$\mathbf{A} = \mathbf{UDV}^T\$로 분리되고, 여기서 \$\mathbf{D}\$는 양수의 대각행렬이고 대각 요소를 크기에 따라 정렬했을 때, 미지수 \$\mathbf{h}\$는 \$\mathbf{V}\$의 마지막 열이 된다. 따라서 상동관계 행렬 \$\mathbf{H}_{ij}\$는 \$\mathbf{h}\$로부터 구할 수 있다. 다수의 카메라 중 하나를 기준 카메라로 설정하고 이를 1이라고 하였을 때, 바닥위의 대응점들을 사용하여 기준 카메라와 나머지 카메라들 간의 상동관계 \$(\mathbf{H}_{12}, \mathbf{H}_{13}, \mathbf{H}_{14}, \dots)\$를 구할 수 있다.

2. 실시간 처리부

실시간 처리부에서는 단일 어안 카메라에서 움직이는 물체를 검출한 후에 전체 카메라의 정보를 융합하여 가려진 물체를 구별하고 색인하여 움직이는 물체를 검출한다.

가. 움직이는 물체 검출

움직이는 물체를 분리하기 위해서 어안 카메라 입력 영상에 가우시안 혼합 모델(Gaussian Mixture Model, GMM)을 적용하여 배경영상을 구하고 이를 통해 전경 영역(foreground region)을 추출하는 방법을 사용하였다^[11]. 하지만 일반적으로 이렇게 추출된 전경 영상은 그림자 영역이나 순간적 밝기 변화영역 등 원치 않는 부분을 포함하게 된다. 이러한 전경 추출 오차는 움직이는 물체의 중심과 바닥면에 교차점을 찾는데 영향을 주어 전체 성능에 영향을 줄 수 있다. 이를 제거하기 위해 KaewTraKulPong이 제안한 방법^[12]을 사용하여 그림자 영역과 순간적 밝기 변화 영역을 찾아 제거하였다. 예를 들어, 각 픽셀의 컬러 값 변화를 채도(chromaticity)와 밝기(illumination) 정보로 분리하고 GMM에 의한 배경영상과 비교할 때 채도는 유사하지만 밝기 값이 더 낮다면 그림자로, 밝기 값이 더 높다면 순간적 밝기 변화에 의한 부분으로 판단한다. 그리고 만약 GMM에 의한 배경영상과 비교에서 채도의 차이가 크다면 움직이는 물체로 판단한다. 그림 3(a)는 각각의 카메라에서 배경영상을 분리한 후 그림자를 제거하여 움직이는 물체 후보를 검출한 결과이다. 이 그림에서 검은 색, 회색, 흰색은 각각 배경 영역, 그림자 혹은 순간적 밝기 변화 영역, 전경 영역을 나타낸다.

그림 5(a)와 같이 전경 영역을 추출한 뒤, 물체가 나타날 수 있는 영역 내에서 일정 크기 이상의 면적을 갖는 전경 영역들을 선택하여 움직이는 물체 후보로 지정

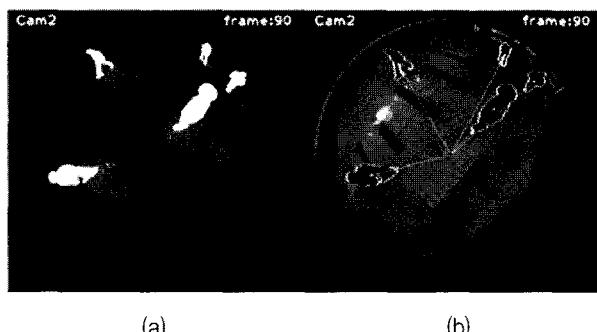


그림 5. (a) GMM과 그림자 및 순간적인 밝기 영역 제거를 통해 얻어진 전경 영역 추출 결과, (b) 움직이는 물체 후보의 중심점과 물체 후보의 중심축이 바닥면과 만나는 교차점 검출 결과

Fig. 5. (a) Foreground extraction results by using GMM and shadow and highlight elimination, (b) Detection of the center of moving object candidate and the intersection between the principal axis of moving object and the ground plane.

한다. 그 후, 후보 물체의 중심점과 소실점을 연결하여 중심축을 알아낸다. 중심축 상에서 물체 후보 영역에 속하면서 소실점에 가장 근접한 점을 움직이는 중심축과 바닥면이 교차하는 점으로 선택한다. 어안 카메라의 광축이 바닥면과 수직으로 설치되면 하나의 소실점은 주점과 같고 다른 하나는 반지름 방향의 $+\infty$ 에 존재한다. 본 논문에서는 서론에서 언급한 것과 같이 전방향을 감시하기 위해 광축이 바닥면과 수직으로 설치되었다고 가정하였다. 그림 5(b)는 각 카메라에서 움직이는 물체 후보의 중심점과 바닥면이 교차하는 점을 보여준다.

나. 다중 어안 카메라의 상동관계에 의한 움직이는 물체검출

그림 6(b)와 같이 어안 카메라 i 에서 움직이는 물체의 중심점 $(\tilde{\mathbf{x}}_1^{(i)}, \tilde{\mathbf{x}}_2^{(i)}, \dots, \tilde{\mathbf{x}}_k^{(i)})$ 과 주점 $\tilde{\mathbf{p}}^{(i)}$ 이 주어졌을 때, 다음과 같이 상동관계 \mathbf{H}_{ij} 를 이용하면 물체의 중심점은 $\tilde{\mathbf{x}}_m^{(i,j)} = \mathbf{H}_{ij} \tilde{\mathbf{x}}_m^{(i)}$ ($m=1, 2, \dots, k$)에 대해서 카메라 1의 좌표 $(\tilde{\mathbf{x}}_1^{(1,i)}, \tilde{\mathbf{x}}_2^{(1,i)}, \dots, \tilde{\mathbf{x}}_k^{(1,i)})$ 로 변환되고 주점은 $\tilde{\mathbf{p}}^{(1,i)} = \mathbf{H}_{ij} \tilde{\mathbf{p}}^{(i)}$ 가 된다 (k 는 물체 후보의 개수). 물체의 중심축은 물체의 중심점 $\tilde{\mathbf{x}}_m^{(i,j)}$ 과 각 카메라의 주점 $\tilde{\mathbf{p}}^{(1,i)}$ 을 연결시킨 선분위에 존재한다. i 번째 카메라의 m 번 물체 후보의 중심축 $\mathbf{l}_m^{(i)}$ 은 식 (10)와 같이 계산된다.

$$\mathbf{l}_m^{(i)} = \tilde{\mathbf{p}}^{(1,i)} \times \tilde{\mathbf{x}}_m^{(1,i)} \quad (10)$$

그림 6(b)에서 보여지는 어안 카메라 i 와 j 에서 검출된 움직이는 물체 후보의 중심축 $\mathbf{l}_m^{(i)}$ 과 $\mathbf{l}_n^{(j)}$ 의 교점 $\omega_{m,n}^{(i,j)}$ 은 식 (11)과 같이 계산된다.

$$\omega_{m,n}^{(i,j)} = \mathbf{l}_m^{(i)} \times \mathbf{l}_n^{(j)} \quad (11)$$

카메라 수와 움직이는 물체 수가 늘어날수록 교차점 수는 증가한다. 그림 7의 ④, ⑤, ⑥, ⑦번 교차점과 같이, 이들 교차점 중에는 동일 물체의 중심축이 아닌 서로 다른 물체의 중심축 간의 교차점도 존재한다. 이러한 잘못 교차된 점들을 제거하기 위해 아래와 같은 방법을 제안하였다. 중심축들의 교점 $\omega_{m,n}^{(i,j)}$ 은 i 번째 카메라에서 후보의 중심점 $\tilde{\mathbf{x}}_m^{(i)}$ 과 주점 $\tilde{\mathbf{p}}_m^{(i)}$ 사이에 존재하고 j 번째 카메라에서 후보의 중심점 $\tilde{\mathbf{x}}_n^{(j)}$ 과 주점 $\tilde{\mathbf{p}}_n^{(j)}$ 사이에 존재해야 하므로 다음 조건을 모두 동시에 만족해야 한다.

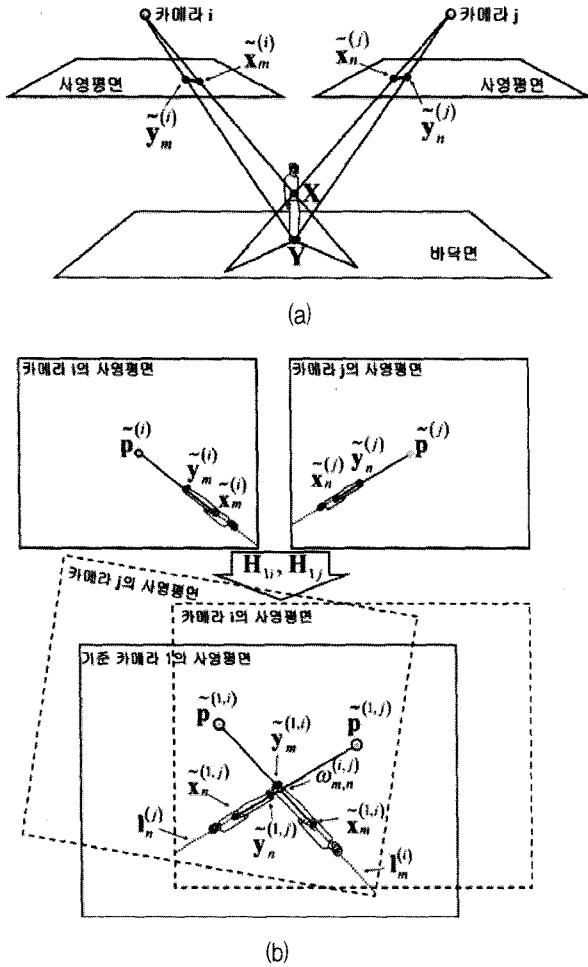


그림 6. (a) 두 어안 카메라(i, j)의 원근 투영 좌표 변환에 의한 사영평면과 공통 평면인 바닥면, (b) 상동관계에 의한 움직이는 물체 검출 방법
Fig. 6. (a) Projective planes in perspective projection converted from two fisheye cameras(i, j) and a common ground plane, (b) Detection method of moving objects by using homographies.

$$\begin{aligned} \|\tilde{x}_m^{(i,j)} - \tilde{p}^{(i,j)}\| &\geq \|\omega_{m,n}^{(i,j)} - \tilde{p}^{(i,j)}\| \\ \|\tilde{x}_n^{(i,j)} - \tilde{p}^{(i,j)}\| &\geq \|\omega_{m,n}^{(i,j)} - \tilde{p}^{(i,j)}\| \end{aligned} \quad (12)$$

$$\begin{aligned} (\tilde{x}_m^{(i,j)} - \tilde{p}^{(i,j)}) \cdot (\omega_{m,n}^{(i,j)} - \tilde{p}^{(i,j)}) &\geq 0 \\ (\tilde{x}_n^{(i,j)} - \tilde{p}^{(i,j)}) \cdot (\omega_{m,n}^{(i,j)} - \tilde{p}^{(i,j)}) &\geq 0 \end{aligned} \quad (13)$$

그림 7에서 교차점 ④, ⑤, ⑥, ⑦은 위의 조건을 만족하지 못한다. 교차점 ④는 카메라로부터 중심점으로 향하는 방향이 같아 식 (13)을 만족하지만 중심점까지 거리 내에 위치하지 않아 식 (12)을 만족하지 못한다. 교차점 ⑥는 중심점까지 거리 내에 위치하여 식 (12)를 만족하지만 반대 방향에 위치하여 식(13)을 만족하지 못한다. 교차점 ⑤, ⑦는 반대 방향에 위치하여 중심점

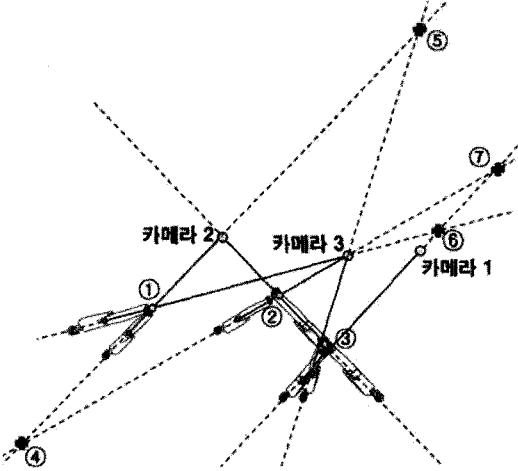


그림 7. 동일한 물체의 교차점 조건

Fig. 7. Constraint for crossing points of the same object.

까지 거리 내에 위치하지 않아 식 (13)과 식 (12)를 동시에 만족하지 못해 교차점에서 제외된다.

식 (12)을 만족하는 교차점들 중에서 동일한 물체에 대응되는 중심축 쌍의 교차점을 찾기 위해 식 (14)과 같은 대응 거리 $D_{m,n}^{(i,j)}$ 를 이용하였다^[5].

$$D_{m,n}^{(i,j)} = \|\tilde{y}_m^{(i,j)} - \omega_{m,n}^{(i,j)}\| + \|\tilde{y}_n^{(i,j)} - \omega_{m,n}^{(i,j)}\| \quad (14)$$

여기서 $\tilde{y}_m^{(i,j)}$ 와 $\tilde{y}_n^{(i,j)}$ 는 2장의 '가'에서 찾은 바닥면과 교차하는 점 $y_m^{(i)}$ 와 $y_n^{(j)}$ 를 상동관계 H_{ii} 와 H_{jj} 에 의해서 변화된 좌표를 의미한다. 동일한 물체에 대응되는 중심축 $I_m^{(i)}$ 과 $I_n^{(j)}$ 사이의 교차점 $\omega_{m,n}^{(i,j)}$ 은 중심축이 바닥면과 접하는 점 $\tilde{y}_m^{(i,j)}$, $\tilde{y}_n^{(i,j)}$ 와 일치해야 한다. 하지만 물체 검출 오차 때문에 정확히 일치하지는 않는 다. 따라서 동일한 물체에 대응되는 중심축 쌍의 교차점은 대응 거리 $D_{m,n}^{(i,j)}$ 가 최소가 되는 교차점 $\omega_{m,n}^{(i,j)}$ 이 된다. 카메라 사이의 대응되는 모든 교차점에 대해서 대응 거리 $D_{m,n}^{(i,j)}$ 가 최소가 되는 교차점 $\omega_{m,n}^{(i,j)}$ 을 찾고 기준 카메라 1을 시작으로 순서적으로 색인화(index)하여 같은 교차점에 대해 같은 이름을 붙인다. 같은 교차점에 대한 평균 위치가 움직이는 물체의 최종 위치가 된다. 다시 상동관계 H_{ii} 와 H_{jj} 를 이용하여 입력영상에 검출된 물체의 이름을 불인다. 이러한 과정을 거치면, 그림 7의 교차점 ③에 있는 사람은 카메라 2에서 가림에 의해 보이지 않아 카메라 3에서의 중심축의 교차점에 의해서는 검출되지 않지만 카메라 1을 추가하므로 교차점 ③을 찾고 이 위치에 있는 물체를 검출할 수 있다.

IV. 실험

본 논문에서는 실험을 위하여 그림 8과 같이 $17m \times 14m$ 실내 환경에서 네 대의 어안 카메라를 바닥으로부터 2.5 m 높이에 수직으로 설치하였다. 어안 렌즈는 Fujinon 사의 YV2.2x1.4A-2^[14]를 사용하였고 이를 네트워크 카메라에 연결하여 영상을 획득하였다. 입력 영상의 크기는 1032×779 pixels이고, 초당 3.5 frames 속도로 카메라 사이의 동기를 맞추어 동영상을 취득하여 실험하였다. 실험영상은 천정을 통해 태양광이 입사되는 실내 환경에서 취득되었다. 사전 처리부는 입력 영상 크기에서 수행하였고, 카메라 사이의 상동관계는 바닥면 위의 타일을 이용하여 10개의 좌표를 가지고 DLT 방법^[13]으로 구했다. 기준 카메라 영상에 각 카메라의 대응 좌표를 상동관계와 어안렌즈 내부변수의 해석에서 다시 투영했을 때, 기준 카메라 영상 좌표와 각 카메라의 대응 좌표 사이의 평균 오차는 평균 2.18 픽셀이었다. 그

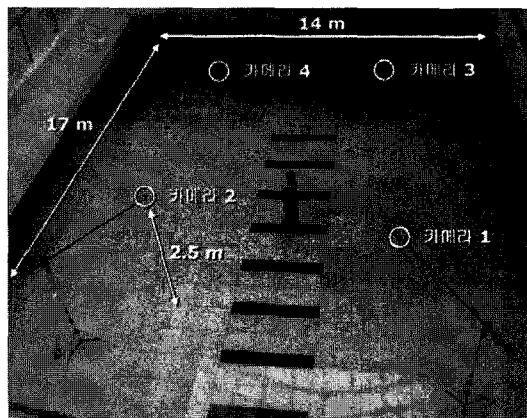


그림 8. 실험 환경

Fig. 8. Experimental environment.

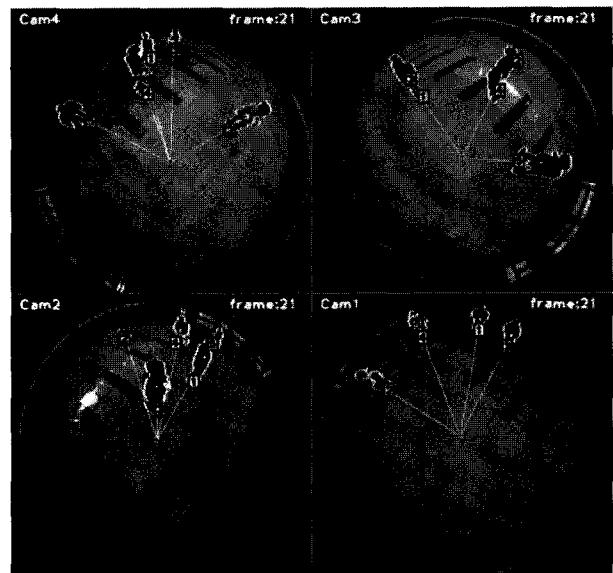
표 1. 개별 카메라를 사용한 방법과 제안된 방법의 검출율 비교

Table 1. Comparison of detection rates.

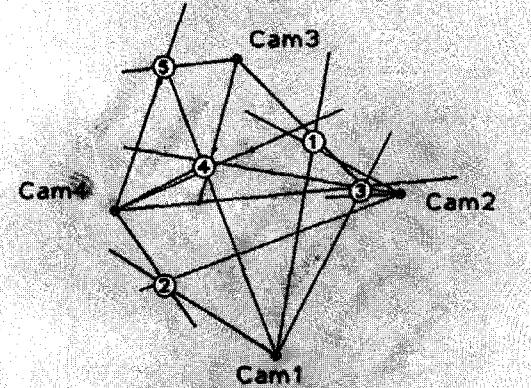
사 람 수	실험환경		검출율 (%)					
	측정 시간	프레 임수	단일 카메라 성능				제안 방법	
			카메라 1	카메라 2	카메라 3	카메라 4	성능	
2	4분 33초	955	87.3	82.9	77.3	81.3	98.3	
3	7분 44초	1624	69.8	72.0	69.8	70.9	95.9	
4	6분 16초	1314	58.1	54.7	50.8	55.2	87.1	
5	7분 08초	1498	44.4	45.5	46.1	40.3	83.0	

림 9(b)는 상동관계에 의한 각 카메라의 사영평면을 카메라 1에 정합한 결과와 카메라 위치를 보여준다.

실시간 처리부는 516×390 pixels으로 다운 샘플링된 영상에 적용 되었다. 성능평가를 위해서 검출율은 다음



(a)



(b)

그림 9. 실험 결과 예시 (a) 가림현상이 있을 때 움직이는 물체 후보 검출과 후보 결정 후 어안 카메라 사이의 연관 관계를 표시한 영상, (b) (a)에서 상동관계를 사용하여 기준 카메리의 사영평면에 각 카메라의 위치와 중심축을 표기하고 움직이는 물체를 검출한 결과

Fig. 9. Example of experimental result (a) Detection of moving object candidates and labelling of moving objects with occlusion, (b) Principal axes and camera positions on projective plane of camera 1 with the detection of moving objects.

과 같이 식 (15)에 정의하였다. 검출율 측정을 위해 정해진 인원이 감시 영역을 서로 교차하면서 돌아다니도록 하였다.

$$\text{검출율} = \frac{\text{검출된 인원이 실제 인원과 같은 프레임 수}}{\text{전체 프레임 수}} \times 100 \quad (15)$$

그림 9는 실험 결과를 보여준다. 그림 9(a)는 각 어안 카메라의 영상에서 움직이는 물체를 검출한 결과를 보여준다. 그림 9(b)는 각 카메라에서 검출된 움직이는 물체들을 기준 카메라에서 바라본 가상평면으로 변환하고 각 중심축들의 교차점을 추출하여 최종적으로 얻어진 움직이는 물체들의 위치를 나타낸다.

그림 9(b)에서 각 카메라에서 가려진 물체는 각 카메라의 주점과 직접적으로 연결된 선분이 존재하지 않는다. 그림 9의 물체들 사이의 상호간 교차에 의해 구별이 되지 않았던 후보가(카메라 1의 5번, 카메라 2의 5번, 카메라 3의 3번과 2번, 카메라 4의 1번) 제안된 방법을 수행하여 해결되었다.

표 1은 움직이는 인원수를 변화시키면서 실험하였을 때 단일 카메라를 사용한 경우와 제안한 시스템을 물체 검출율로 비교하였다. 인원수를 증가할수록 단일 카메라에서 움직이는 물체의 가림현상이 빈번히 발생하여 검출율이 급격히 낮아졌지만 제안된 방법의 경우 검출율이 높게 유지되는 것을 발견할 수 있었다.

2명이 움직였을 때 단일 카메라에서 최대 87.3%이나 제안된 방법은 98.3%로 향상되었으며, 5명이 움직였을 경우 46.1%에서 83.0%로 성능이 향상되었다. 다중 어안 카메라를 사용하여도 성능이 100%가 안 나오는 이유는 두 물체 사이가 너무 가까운 경우이다. 이는 연속적인 추적방법과 결합하면 해결될 수 있을 것이다.

V. 결 론

본 연구에서는 천정이 낮고 넓은 실내공간에서 움직이는 물체의 위치 및 개수를 파악할 수 있는 다중 어안 카메라 기반 감시 시스템을 제안하였다. 제안된 방법에서는 일반적인 화각을 가지는 렌즈가 아닌 전 방향의 화각을 갖는 어안렌즈를 사용하여 천장이 낮고 넓은 실내에서 보다 넓은 영역을 감시하는 것을 가능하게 하였으며, 다중 어안 카메라와 각 카메라들 간의 상동관계를 사용하고 움직이는 물체와 관련이 없는 교차점 제거 방법을 제안함으로써 움직이는 물체 간에 나타나는 가림현상 문제를 효과적 풀 수 있었다.

향후 검출된 물체들의 위치를 연속적으로 추적하는 알고리즘과 결합하여 성능을 향상시키는 연구를 수행하고자 한다.

참 고 문 현

- [1] 배광혁, 박강령, 김재희, “전 방향 카메라 영상에서 사람의 얼굴 위치검출 방법”, 대한전자공학회 학제 종합학술대회, 30권, 1호, 2007.
- [2] K. Bae, K. R. Park, and J. Kim, “3D Head Position Estimation using a Single Omnidirectional Camera for Non-intrusive Iris Recognition”, MRCS 2006, LNCS, Vol. 4105, pp. 167-174, 2006.
- [3] K. Bae, K. R. Park, and J. Kim, “Tracking Head Positions in Three Dimensional Space by Using a Single Omnidirectional Camera”, Optical Engineering, Vol. 47, No. 4, 2008.
- [4] W. Hu, T. Tan, L. Wang, and S. Maybank, “A survey on visual surveillance of object motion and behaviors”, IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, Vol. 34, no. 3, pp. 334 - 352, 2004.
- [5] W. Hu, M. Hu, X. Zhou, T. Tan, J. Lou, and S. Maybank, “Principal axis-based correspondence between multiple cameras for people tracking”, IEEE Trans. Pattern Recognition and Machine Intelligence, Vol. 28, no. 4, pp. 663 - 670, 2006.
- [6] L. Lee, R. Romano, and G. Stein, “Monitoring Activities from Multiple Video Streams: Establishing a Common Coordinate Frame”, IEEE Trans. Pattern Recognition and Machine Intelligence, Vol. 22, no. 8, pp. 758-767, 2000.
- [7] K. Kim and L. S. Davis, “Multi-camera Tracking and Segmentation of Occluded People on Ground Plane Using Search-Guided Particle Filtering”, ECCV 2006, Part III, LNCS, Vol. 3953, pp. 98-109, 2006.
- [8] W. Du and J. Piatet, “Muti-camera People Tracking by Collaborative Particle Filters and Principal Axis-Baed Integration”, ACCV 2007, Part I, LNCS, Vol. 4843, pp. 365-374, 2007.
- [9] E. Garcia and L. Altamirano, “Decision Level Multiple Camera Fusion Using Dezert-Smarandache Theory”, CAIP 2007, LNCS, Vol. 4673, pp. 117-124, 2007.
- [10] S. S. Brandt, “A Generic Camera Model and Calibration Method for Conventional, Wide-Angle, and Fish-Eye Lenses”, IEEE Trans. Pattern Recognition and Machine Intelligence, Vol. 28, no. 8, pp. 1335-1340, 2006.

- [11] T. Horprasert, D. Harwood, and L.S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in Proc. IEEE ICCV'99 FRAME-RATE Workshop, 1999.
- [12] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection," in Proc. European Workshop on Advanced Video Based Surveillance Systems (AVBS01), 2001.
- [13] R. Hartley and A. Zisserman, *Multiple View Geometry in computer vision*, Cambridge University Press, 2003.
- [14] <http://www.fujinon.com/> (accessed on 2008.02.26)

저자 소개



배 광 혁(학생회원)
2001년 건국대학교 전자공학과
학사 졸업.
2003년 연세대학교 전기전자공
학과 석사 졸업.
2008년 연세대학교 전기전자공
학과 박사 과정.

<주관심분야 : 지능형감시시스템, 생체인식, 컴퓨터비전>



박 강 령(정회원)
1994년 연세대학교 전자공학과
학사 졸업.
1996년 연세대학교 전자공학과
석사 졸업.
2000년 연세대학교 전기·컴퓨터
공학과 박사 졸업.
2000년 2월~2003년 2월 LG전자 기술원 Digital
Vision Group 선임연구원.
2003년 3월~2008년 2월 상명대학교 소프트웨어
대학 디지털 미디어 학부 조교수.
2008년 3월~현재 동국대학교 전자공학과
조교수.
2003년 8월~현재 생체인식연구센터 총괄과제
책임자.
<주관심분야 : Biometric 영상 처리, 패턴인식, 컴퓨터비전>



서 재 규(학생회원)
2005년 인하대학교 전자공학과
학사 졸업.
2007년 연세대학교 전기전자
공학과 석사 졸업.
2008년 연세대학교 전기전자
공학과 박사 과정.

<주관심분야 : 지능형감시시스템, 생체인식, 컴퓨터비전>



김 재 희(평생회원)-교신저자
1979년 연세대학교 전자공학과
학사 졸업.
1982년 Case Western Reserve
University 전기공학과
석사 졸업.
1984년 Case Western Reserve
University 전기공학과
박사 졸업.
2008년 연세대학교 전기전자공학부 교수.
2008년 (과학기술부/과학재단 지정 ERC)
생체인식연구센터 소장.
2008년 한국바이오인식포럼 의장.
2008년 한국공학한림원 정회원.
<주관심분야 : 생체인식, 패턴인식, 컴퓨터 비전,
영상인식>