

논문 2010-47SP-3-7

행동 탐지 기반의 효율적인 객체 선택 알고리즘

(Efficient Object Selection Algorithm by Detection of Human Activity)

박 왕 배*, 서 용 호*, 두 경 수*, 최 종 수*

(Wang-Bae Park, Yung-Ho Seo, Kyoung-Soo Doo, and Jong-Soo Choi)

요 약

본 논문에서는 행동 탐지 기반으로 사람의 지시행위를 인식하여 지시방향의 객체를 선택하고 이를 추적하는 시스템을 제안한다. 일반적으로 사람은 무엇인가를 지시할 경우, 얼굴 방향을 목표물에 두게 된다. 따라서 얼굴과 손끝을 연결한 직선을 지시방향이라 간주하고, 지시된 객체를 선택한다. 제안된 알고리즘에서는 카메라를 통해 입력된 영상으로부터 움직임 영역을 검출하기 위해 배경 차분을 사용하여 실시간으로 관심 객체의 움직임을 추출한다. 보행 여부의 판단은 주성분(PCA) 분석과 객체의 움직임 변위로 결정되며, 이 때 사람이 정지 상태면, 머리를 기준으로 손에 이르는 벡터 관계를 계산하여 사용자의 지시 방향을 최종적으로 결정한다. 실험결과를 통하여 다시점 카메라를 이용한 다각도의 영상에서 사람의 지시 방향을 정확하게 추정해 냄으로써 제안된 알고리즘의 유효성을 검증하였다.

Abstract

This paper presents an efficient object selection algorithm by analyzing and detecting of human activity. Generally, when people point any something, they will put a face on the target direction. Therefore, the direction of the face and fingers and was ordered to be connected to a straight line. At first, in order to detect the moving objects from the input frames, we extract the interesting objects in real time using background subtraction. And the judgment of movement is determined by Principal Component Analysis and a designated time period. When user is motionless, we estimate the user's indication by estimation in relation to vector from the head to the hand. Through experiments using the multiple views, we confirm that the proposed algorithm can estimate the movement and indication of user more efficiently.

Keywords : Pointing region estimation, Motion analysis, Object tracking, Interactive

I. 서 론

최근 컴퓨터의 하드웨어와 소프트웨어의 눈부신 발달로 인해 컴퓨터 비전 기술을 사용한 HCI(Human Computer Interaction), AR, 보안 감시시스템, 로봇틱스(robotics), 사람의 행동인식(human motion analysis)과

같은 연구가 많이 진행되고 있다^[1]. 일반적으로 사람은 서비스 제공을 받기 위해 관심영역을 표시하고, 이를 상대방에게 알리는 방법으로 손의 움직임을 이용한 지시 동작으로 자신의 의도를 알린다^[2]. 따라서 손의 움직임에 의한 지시 방향의 추정 및 지시 영역의 검출은 서비스 제공자에게는 매우 중요한 일이다. 그러나 컴퓨터가 그러한 지시 방향 및 영역 추정을 수행하기 위해서는 사람의 얼굴, 손 등을 사용한 동작을 실시간으로 센싱(sensing)하고 이를 분석하여 요구에 응하는 과정이 필요하다^[3]. 이러한 지시 영역 추정은 제스처 인식 및 로봇 제어 분야에서 응용되어 연구되고 있다. 행동인식을 기반으로 로봇을 제어하는 연구가 많이 진행되고 있

* 정회원, 중앙대학교 첨단영상대학원
(Dept. of Image Engineering, Graduate School of Advanced Image Science, Multimedia, and film, Chung-Ang University)

※ “이 논문은 2단계 BK21 사업, 서울시 산학연 협력 사업 기술이전지원사업(TR080601), 서울시 산학 클러스터(SFCC)사업(10570) 지원과제임.”

접수일자: 2009년8월17일, 수정완료일: 2010년3월31일

는데 Masanori Nishimura^[4]는 로봇의 제어를 손으로 하였다. 하지만 손을 인식하기 위해서 지정된 화면의 영역에서만 팔을 검출한 후 팔 동작 모델을 사용하여 명령을 주었다.

또한, 유비쿼터스 홈 지능화 시스템 구현을 위해 사용자의 명령에 따라 로봇제어가 가능한 시스템들도 구축되고 있다. E. Sato^[5]는 실험 공간의 천장에 2대, 로봇에 1대 등 총 3대의 카메라를 장착하여 사람이 지시하는 지점으로 로봇이 이동하는 지시이동 방법을 제안하였다. Andreas Butz^[6]는 천장에 라이트와 카메라를 부착하여 사용자가 원하는 책을 입력하면 책이 있는 곳을 라이트로 비추어주는 시스템을 제안하였다. 이 시스템은 책에 고유 마커를 부착하여 사용자가 원하는 책을 검색하였을 경우 라이트가 찾으려고 하는 책을 비추어주는 시스템으로 보다 편리하고 짧은 시간에 원하는 책을 찾아주는 시스템이다. Marie Sester^[7] 등은 자체 제작한 스포트라이트(spotlight) 시스템을 천장에 설치하여 추적된 사람에게 조명을 조사하여주는 시스템을 제안하였다.

본 논문에서는 다시점 영상에서 효율적인 객체 선택 알고리즘을 제안한다. 카메라를 통해 실시간으로 입력되는 영상 시퀀스(image sequence)에서 움직이는 객체를 추출하기 위해 배경 차분하여 관심 객체만을 검출한다. 검출된 관심 객체는 줄 길이 레이블링(run-length labeling)으로 분류한다. 레이블링된 객체의 움직임을 추적하기 위해 칼만 필터(Kalman filter)를 사용한다. 또한 구분한 객체에 주성분 분석(Principal Component Analysis, PCA)알고리즘을 사용하여 사람의 머리 영역을 추정하고, 시간 간격에 따라 사람의 움직임을 이동 상태와 정지 상태로 판별한다. 정지 상태일 때 사람의 머리와 팔을 이용하여 지시자의 지시방향을 판별하고 다른 객체의 유무를 판단한다. 객체가 있을 때에는 지시자의 지시 방향으로부터 객체를 선택할 수 있다. 본 논문의 구성은 II장에서 평면 호모그래피와 배경차분 방법에 대해 기술하고, III장에서는 레이블링 알고리즘과 칼만 필터에 대해 기술하며, IV장에서는 객체의 이동 및 정지와 지시방향을 추정하기 위한 방법에 대하여 기술한다. V장에서는 제안된 방법에 의한 실험 결과를 보여주며, VI장에서는 결론을 맺는다.

II. 전처리 과정

1. 평면 호모그래피

본 논문에서는 실시간 영상 시퀀스로부터 검출된 객체를 실세계좌표인 3차원으로 좌표를 추정하기 위해 객체의 기하학 특성을 고려하여 평면 호모그래피 교정법을 적용하여 각각 객체의 좌표를 추정한다.

두 평면에서 대응하는 모든 점들은 호모그래피(homography)관계로 표현할 수 있다. 즉, 한 대의 카메라로 찍은 평면과 조명이 바라보는 평면에 있는 대응점 x' 와 x 사이에는 식 (1)과 같은 관계가 있다. 이들 관계를 나타내는 변환 H 를 호모그래피라고 한다.

$$\begin{aligned} x' &= Hx \\ H &= \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \end{aligned} \quad (1)$$

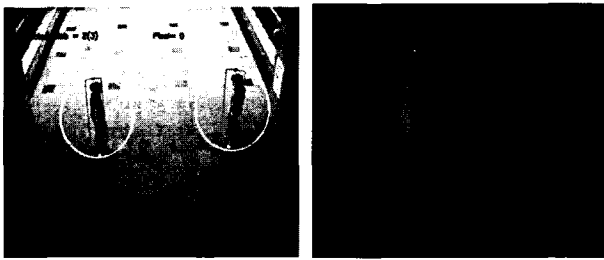
실세계에 존재하는 관심 평면 $Z = 0$ 이라면, 이 평면상의 점 $x = (X \ Y \ 0 \ 1)^T$ 로 표시할 수 있으며, i 번째 카메라에 의해 투영된 점 $x' = (x \ y \ 1)^T$ 로 표현할 수 있다. 카메라 투영행렬 P 가 3차원 공간상의 점 x 를 이미지의 점 x' 로 투영되는 행렬이라고 하면, 실세계 평면에 있는 점 x 와 카메라 투영행렬과의 관계식은 식(2)와 같다.

$$\begin{aligned} x' &= P_i X = K [R | t] \cdot (X \ Y \ 0 \ 1)^T \\ &= K \begin{bmatrix} r_1 & r_2 & r_3 & t \\ r_1 & r_2 & r_3 & t \end{bmatrix} \cdot (X \ Y \ 0 \ 1)^T \\ &= K \begin{bmatrix} r_1 & r_2 & t \\ r_1 & r_2 & t \end{bmatrix} \cdot (X \ Y \ 1)^T \\ &= H_i' (X \ Y \ 1)^T \end{aligned} \quad (2)$$

여기서 P 는 3×4 의 카메라 투영 행렬이며, K 는 3×3 의 카메라 내부 매개변수 행렬이고, $\begin{bmatrix} r_1 & r_2 & t \\ r_1 & r_2 & t \end{bmatrix}$ 는 카메라 외부 매개변수 중 회전행렬의 첫째와 둘째 열과 이동 벡터에 해당한다.

2. 배경차분

고정된 카메라에서 획득한 영상 시퀀스에서 원하는 객체를 추출하기 위해 사용하는 방법으로서 RGB칼라 영상으로 취득된 배경 영상과 현재 영상을 회색영상으로 변환, 입력영상(Input Frame, IF)과 배경영상(Background Frame, BF)의 회색영역(gray scale)을 화소 단위로 차분하는 것이다. 배경 영상과 입력 영상을 차분하여 전경영상(Foreground Fame, FF)을 구하는 식



(a) 입력 영상 (b) 객체 추출

그림 1. 객체 검출
Fig. 1. Object detection.

은 식 (3)과 같다.

$$FF(x, y) = \begin{cases} 1, & \text{if } |IF(x, y) - BF(x, y)| > \text{Threshold} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

영상차분에 의해 검출된 현재 영역을 이진화(binanzation)하여 관심 객체만을 검출하기 위해서 침식(erosion)과 팽창(dilation)연산을 수행하여 객체 영역에 포함되는 화소와 포함되지 않는 화소를 구분한다. 이 방법은 다른 배경차분 방법보다 간단하고 빠르며 고정된 카메라에서 사용할 경우 효율적이다.

III. 객체 구분 및 추적

1. 줄 길이 레이블링(Run-length labeling)

본 논문에서는 실시간 처리를 위해서 배경차분을 통해 추출된 객체에 대해서 줄 길이 레이블링을 적용하여 객체를 구분한다. 기존의 Grassfire 레이블링 방법은 한 영역의 시작 픽셀로부터 영역의 끝 픽셀까지 한 번에 검색을 하지만, 줄 길이 레이블링은 영상의 한 줄씩 검색을 해 나가는 방법으로 한 줄의 검색이 수행된 후 다음 줄의 검색을 수행하면서 각각의 줄에 물체로서 판별이 되는 영역에 순차적인 숫자를 매겨준다. 이렇게 판별된 영역에 대해 상위 줄의 판별 영역과 비교를 통하여 인접한 영역이 있을 때 최종적으로 같은 영역으로 판별을 한다. 이 알고리즘은 면적스캔 카메라에서는 물론 라인스캔 카메라에서도 이용할 수 있다. 이 경우 대용량의 데이터를 라인 단위로 실시간 처리할 수 있어, 스캔이 끝나는 순간에 레이블된 영상뿐만 아니라, 물체들에 대한 정보 추출도 같이 끝나기 때문에 빠르고 정확한 결과를 얻을 수 있는 효과적인 알고리즘이다.

2. 주성분 분석

(Principal Component Analysis, PCA)

주성분 분석이라는 통계학 기법을 이용하면 움직이는 객체가 구성하고 있는 정점(vertex)들의 배치 형태를 분석하고 그 형태에 대한 자연스러운 좌표축을 결정할 수 있다. 주성분 분석을 이용하면, x, y좌표들로 된 정점 위치들의 배열처럼 서로 연관된 자료들의 집합이 서로 연관되지 않은 성분들로 분리될 수 있는 하나의 좌표공간을 찾을 수 있다. 데이터의 기본 주성분은 데이터가 가장 많이 변하는 방향을 의미한다.

N개의 정점들 P_1, P_2, \dots, P_i 로 이루어진 임의의 집합이 있다고 하자. 우선 다음 식 (4)을 이용해서 평균 위치 m 을 구한다.

$$m = \frac{1}{N} \sum_{i=1}^N P_i \quad (4)$$

그리고 다음 식 (5)를 이용해서 2x2 행렬 C를 만든다. 이 행렬을 공분산 행렬(covariance matrix)이라고 부른다.

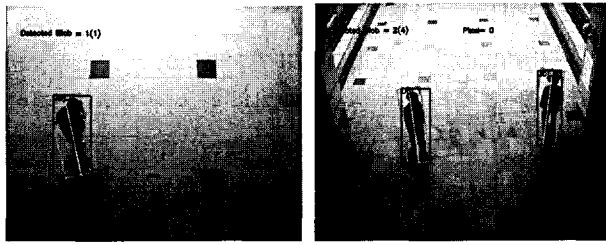
$$C = \frac{1}{N} \sum_{i=1}^N (P_i - m)(P_i - m)^T \quad (5)$$

공분산 행렬은 식 (6)과 같은 3개의 서로 다른 성분들을 가진 대칭 행렬(symmetry matrix)이다.

$$\begin{aligned} C_{11} &= \frac{1}{N} \sum_{i=1}^N (x_i - m_x)^2 \\ C_{12} &= C_{21} = \frac{1}{N} \sum_{i=1}^N (x_i - m_x)(y_i - m_y) \\ C_{22} &= \frac{1}{N} \sum_{i=1}^N (y_i - m_y)^2 \end{aligned} \quad (6)$$

공분산 행렬의 성분들은 각 x, y좌표들의 쌍(pair)들 사이의 상관관계를 의미한다. 어떤 공분산 행렬의 성분이 0 이라는 것은, 그 성분을 계산하는데 쓰인 두 좌표들 사이에 상관관계가 없다는 의미이다. C가 대각행렬(diagonal matrix)이면, 두 좌표들이 완전히 무관하다는 뜻으로, 이는 점들이 각 축에 대해 균등하게 분포하고 있음을 의미한다.

정점들을 적절히 변환하면 공분산 행렬이 대각행렬이 될 수 있다. 그러한 변환을 수행하는 행렬을 구하는 것이 일차적인 목표이다. 변환행렬 A를 각 점 $\{P_i\}$ 에 적용했을 때, 변환된 점들에 대한 공분산 행렬 C'는 다음



(a) 단일 물체의 경계상자 (b) 복수 물체의 경계상자

그림 2. 2차원 객체의 최소 경계상자
Fig. 2. 2D object's minimum bounding rectangle.

과 같이 주어진다.

$$\begin{aligned}
 C' &= \frac{1}{N} \sum_{i=1}^N (AP_i - Am)(AP_i - Am)^T \\
 &= \frac{1}{N} \sum_{i=1}^N A(P_i - m)(P_i - m)^T A^T \\
 &= ACA^T
 \end{aligned} \tag{7}$$

여기에서 A 는 직교행렬이어야 하며, A 의 전치행렬이 행렬 C 를 대각화해야 한다는 것을 의미한다. C 는 대칭 행렬이므로, C 의 고유벡터(eigen-vector)들은 직교이다. 행렬이 C 의 고유벡터들이 행렬은 A 가 갖춰야 할 조건들은 모두 만족한다. 그러한 행렬은 정점 좌표들이 서로 연관되지 않는 공간으로 정점들을 변환시킨다.

C 의 고유벡터를 R, S 라고 하면, 방향 R 은 항상 주축을 의미한다. 주축(principal axis)은 공분산 행렬의 가장 큰 고유 값에 해당한다. S 는 각각 두 번째로 큰 고유 값을 의미한다. 즉, λ_1, λ_2 이 순서대로 R, S 라고 하면, 크기 관계는 $|\lambda_1| \geq |\lambda_2|$ 이다. 이런 주성분 분석을 통해서 그림 2와 같이 2D객체의 최소 경계상자(MBR)를 구한다.

3. 객체 추적

각각 구분하여 검출된 객체에 대해서 칼만 필터를 적용하면 모든 영상을 프로세싱 할 필요 없이 다음 영상에서 사람이 어느 곳으로 이동할지 예측할 수 있기 때문에 시간을 줄일 수 있다. 단, 본 논문에서 객체 추적 실험은 매 영상마다 사람이 아주 빠르게 비선형으로 이동하지 않는다는 조건하에서 이루어 졌다.

칼만 필터는 이산 신호 처리에서 이전 상태를 기반으로 이후 상태를 추정하기 위해 일반적으로 사용된다. 식 (8), 식 (9)은 각각 시스템 방정식과 관측 방정식을 보여주고 있다. x 는 해당 시간의 알고자 하는 변수들의 상태 벡터를 의미하며, z 는 해당 시간의 변수들의 관측

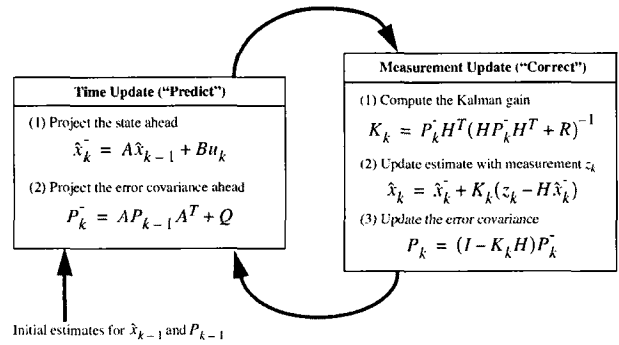
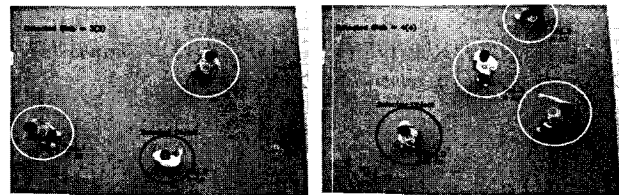
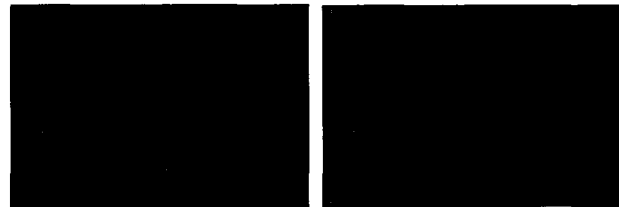


그림 3. 칼만 필터 수행 흐름도
Fig. 3. Diagram of the operation of the kalman filter.



(a) 입력영상



(b) 이동 궤적

그림 4. 칼만 필터를 적용한 객체 이동경로
Fig. 4. Kalman filter applied to move the path of the object.

벡터를 의미하고, A 는 이전 상태와 현재 상태의 관계를 정의해주는 상태 전이행렬을 의미한다. H 는 이전 상태와 현재 관측치와의 관계를 정의해 주는 관측행렬이고, w 와 v 는 각각 시스템 오차, 관측 오차를 의미한다. w 와 v 는 평균값이 0이며, 분산이 각각 Q, R 을 갖는 정규 확률 분포를 따른다.

수식 (10)은 2차원 영상에서 칼만 필터를 적용하기 위한 상태 전이행렬 A 와 관측 행렬 H 를 표현한 것이다.

$$x_{k+1} = Ax_k + w_k \quad p(w) = N(0, Q) \tag{8}$$

$$z_{k+1} = Hx_k + v_k \quad p(v) = N(0, R) \tag{9}$$

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \tag{10}$$

그림 3은 칼만 필터가 실제 수행되는 흐름도를 보여 주고 있다. 예측 부분에서 Bu_k 은 선택적 추가 입력으로서, 본 논문에서는 이 값을 0으로 설정하였다. P_k 와 P'_k 는 각각 이전 상태 벡터와 현재 상태벡터에 대한 연역적(a priori) 추정 오차 분산과 귀납적(a posteriori)추정 오차 분산을 의미한다.

칼만 필터는 시스템의 상태 벡터를 추정하는 예측 과정과 관측치로부터 칼만 필터를 갱신하는 교정과정으로 구성된다. 그림 4는 추출된 객체의 이동경로를 보여준다.

IV. 지시 영역 추정

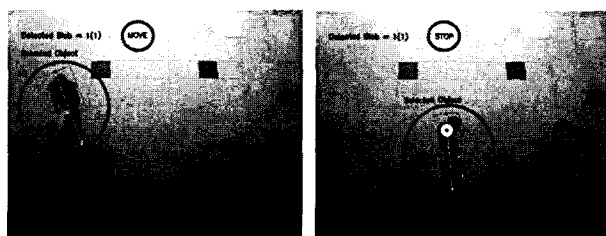
1. 머리 추정

영상에서 검출된 전신은 앞서 III장에서 설명한 주성분 분석에 의해 알아낸 고유벡터 R, S 로 최소 경계 상자(Minimum Boundary Rectangle, MBR)를 생성하여 머리 위치를 추정하는데 사용한다.

우선, R, S 를 기반으로 최소 경계 상자를 만들려면, 그 방향들에 따른 정점 집합의 최소, 최대 범위들을 추정한다. 정점 집합의 최소, 최대 범위들을 구하면, 그것으로부터 최소 경계 상자의 4점들을 결정한다.

정점 집합의 최소, 최대 범위를 알아내는 방법은 다음과 같다. 각 정점 위치 P_i 와 단위 길이 벡터 R, S 의 내적(inner-product)을 얻고, 그 값들의 최소, 최대값을 추정한다. 최소 경계 상자의 크기는 R, S 방향들에 대한 최소, 최대 내적들 사이의 차이이다. 최소 경계 상자의 중심 Q 는 상자를 반으로 나누는 두 점들, 즉 서로 마주보는 두 점들의 중간에 있는 평면의 교점이다. 사람 모델의 장축의 길이를 알 수 있기 때문에 인체 비율을 적용하여 사람의 머리위치를 추정 한다.

사람이 지시하는 행동을 취하려면 누군가를 본 후,



(a) 이동상태 (b) 정지상태

그림 5. 인체 비율을 적용하여 추정된 머리
Fig. 5. Estimated by applying the percentage of body head.

손으로 가리키는 행동을 한다. 본 논문에서는 칼만 필터를 사용하여 이동하고 있는 객체의 좌표를 실시간으로 알 수 있기 때문에 객체의 이동 상태와 정지 상태를 판별할 수 있다. 정지 상태에서 지시를 한다는 전제조건하에서 선택된 객체가 이동하고 있을 때 영상에 'MOVE'라고 출력하였고, 정지 상태에는 'STOP'이 출력되도록 설정하였다. 정지 상태가 되었을 때 행동을 시작하게 되고, 주성분 분석을 통해 알고 있는 인체 비율의 정보를 통해 머리의 위치를 추정한다.

2. 손 검출

지시하는 영역을 계산하려면 손의 방향을 결정하는 과정이 필요하다. 손의 방향을 결정하기에 앞서 영상에서 손이 위치하는 영역을 추출한다. 사람이 정지한 상태에서 지시 영역을 추정하는 전제 조건에 따라 움직임이 없는 상태를 결정하는 과정이 필요하다. 본 논문에서는 프레임간 차분을 사용하여 사람의 움직임 여부를 판단한다. 정지 상태임을 판단한 후 손의 위치를 판단하고 머리와 손의 관계에 따라 방향을 추정한다.

사람의 움직임 여부를 결정하기 위하여 본 논문에서는 칼만 필터를 사용한다. 칼만 필터 내부의 표준 편차에 의해 사람이 이동 상태인가 정지 상태인가를 판별할 수 있다. 또한, 시간의 간격을 결정해야 할 필요가 있으므로 시간 간격은 중심 좌표변화의 블록 크기로 결정한다. 정지 상태 후 프레임을 저장하여 다시 현재 프레임과 프레임간 차분을 수행하였다. 임계치에 따라 블록은



(a) 차분영상 (b) 이진영상



(c) 차분할 프레임 (d) 차분영상

그림 6. 프레임간 차분을 통해 검출한 팔 영역
Fig. 6. Arm area detected by differential frame.

정지 상태 혹은 이동 상태로 결정하게 된다. 정지 상태 프레임에 대해 지시를 수행한 프레임을 제거함으로써 손 영역을 검출할 수 있다. 그림 6은 팔을 검출하기 위해 프레임간 차분을 수행하고, 사람이 정지한 프레임을 저장하여 손을 들어 지시하는 행동에 따라서 팔 영역만을 검출한 결과를 보여주고 있다.

3. 지시 영역 추정

다시점 카메라에 대하여 사용자의 머리와 손 끝 영역을 검출하였다. 지시를 수행하기 위해 획득한 머리와 손의 정보를 사용하여 직선의 방정식을 얻어낸다. 선택된 객체의 머리좌표와 손 영역의 좌표 값을 사용하여 지시 직선을 추정한다. 하지만, 선택하고자 하는 반대 방향에 또 다른 부객체가 있을 경우에는 객체선택의 혼동이 발생한다. 따라서 주 객체의 장측에서 손을 뻗은 방향의 객체만 선택 되도록 하였다.

V. 실험결과

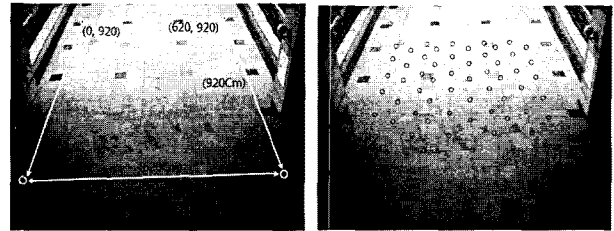
본 논문은 다시점 카메라를 사용하여 사용자의 머리와 손을 검출하고, 2차원 공간상에서 사용자가 지시하는 객체를 선택하는 방법을 제안하였다. 제안하는 2차원 위치추정객체가 실제로 정확하게 위치를 추정하고, 다수의 객체 중 한 객체를 선택한다. 선택된 객체는 다른 객체에게 지시행동을 하기 위한 머리와 손을 추정하여, 지시하였을 때 지시된 객체가 올바르게 선택되었는가에 목적이 있다.

평면 호모그래피를 통해 얻어낸 영상좌표와 호모그래피를 통해 공간 좌표를 추출하여 각각의 포인트를 정해 두 좌표계에서 발생할 수 있는 오차율을 측정하여 정확도를 분석한다. 다시점 카메라를 사용하여 카메라의 위치에 따라 화면이 달라지기 때문에 본 논문의 알고리즘을 다양한 각도에서 비교 및 분석한다.

1. 평면 호모그래피 추정

실시간으로 들어오는 고정된 배경 영상(640×480)에서 코너점을 취득한 후, 사전 정보를 이용하여 4개점의 영상좌표를 추출한다.

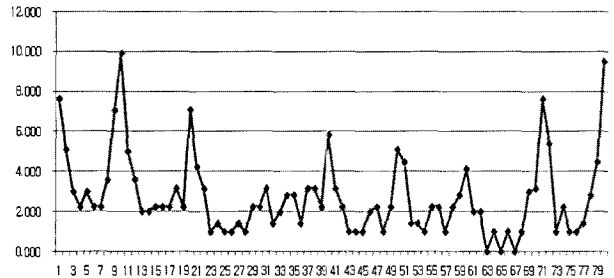
다음으로 공간상의 좌표를 구하기 위해 앞에서 추출한 4점들의 공간상의 좌표를 설정한다. x축의 길이는 0.62m이고, y축의 길이는 0.92m이다. 이러한 방법으로 호모그래피를 구한 후, 총 60~80개의 좌표를 각 지점



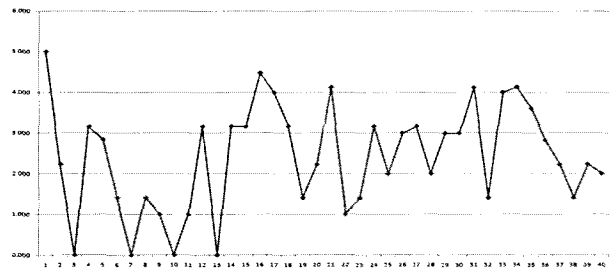
(a) 배경영상 좌표 (b) 실세계 취득 좌표

그림 7. 호모그래피 좌표 추정

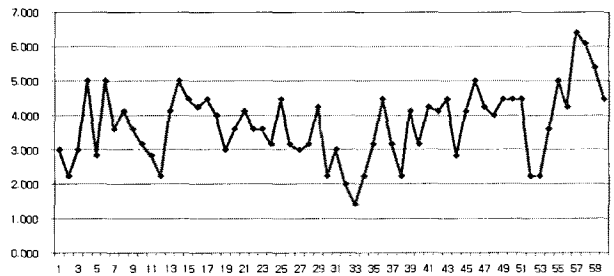
Fig. 7. Homography estimated coordinates.



(a) 30도



(b) 45도



(c) 60도

그림 8. 호모그래피 오차 값

Fig. 8. Homography error value.

에 찍어 결과를 산출하였다.

그림 8은 위 실험 결과의 오차를 나타낸 표이다. 위 실험과 같은 방식으로 카메라 각도 30도, 45도, 60도에서 실험을 해본 결과 카메라 특성상 발생하는 왜곡(distortion)때문에 영상의 주점에서 멀리 떨어진 외곽 부분의 오차값이 크게 발생하였다. 평균 오차 값은 30도에서 2.73cm, 45도에서 2.42cm, 60도에서는 3.72cm로

추정되었으며, 이러한 오차값은 응용 시스템에 적용할 수 있는 허용오차범위 안에 포함된다고 볼 수 있다.

2. 지시 영역 추정

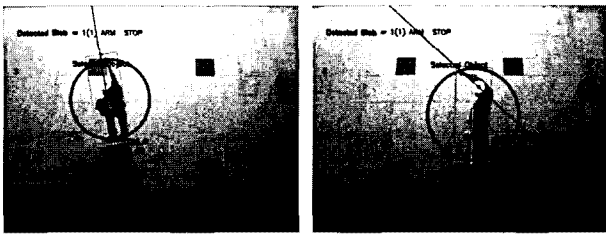
그림 9는 사용자의 머리와 손끝을 추정하는 실험을 보여주고 있다.

선택된 객체의 인체 움직임으로부터 지시하는 방향의 객체를 인식한 후, 지시된 객체로 선택하는 실험을 하였다. 선택된 주 객체는 현재 선택되어 있는 객체로서, 지시가능한 객체이며, 부 객체는 아직 선택되지 않

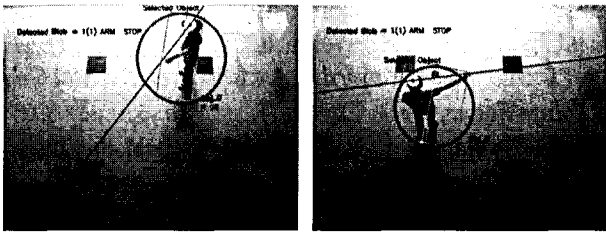
은, 주 객체 외의 모든 객체를 의미한다. 후보객체는 주 객체가 지시한 객체를 의미한다. 본 실험은 카메라의 각도를 30도, 45도, 60도의 변화를 주어 실험을 진행하였다.

그림 10, 11, 12에서의 실험결과는 주 객체가 지시하는 방향에 따라 객체가 검출되었을 때, 주목 객체가 변경되어 새로운 객체를 선택한 결과를 각도변화에 따라 보여주고 있다.

그림 13은 카메라 각도에 따라 주 객체가 후보객체를 지시하였을 때, 지시하는 주 객체의 직선과 후보객체의

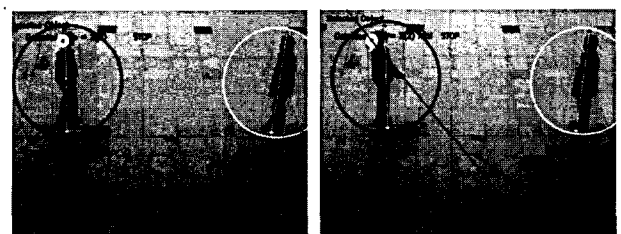


(a) (b)

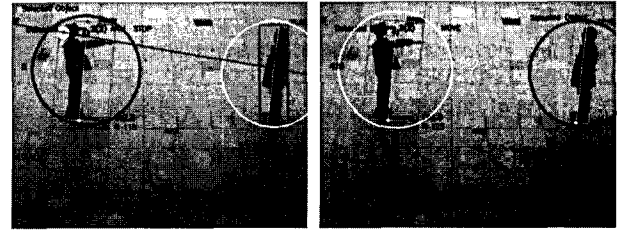


(c) (d)

그림 9. 지시영역 추정 결과
Fig. 9. Result of pointing region estimation.

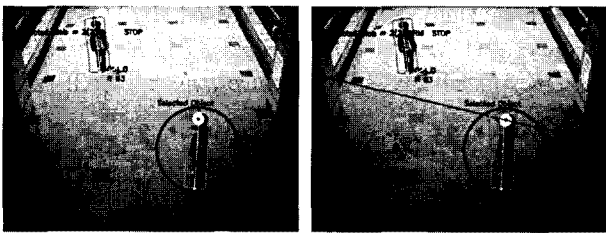


(a) (b)

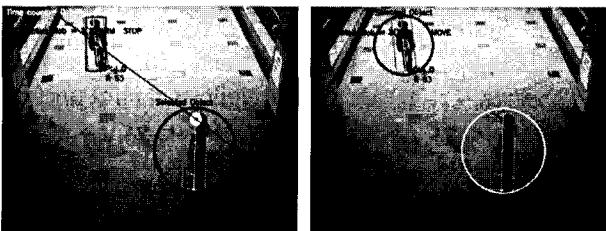


(c) (d)

그림 11. 카메라 각도 45도
Fig. 11. Camera angle 45.

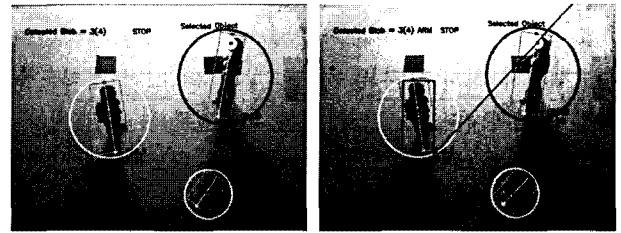


(a) (b)

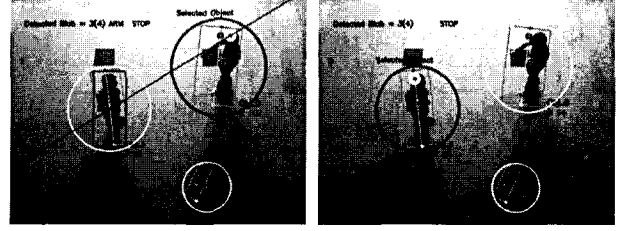


(c) (d)

그림 10. 카메라 각도 30도
Fig. 10. Camera angle 30.



(a) (b)



(c) (d)

그림 12. 카메라 각도 60도
Fig. 12. Camera angle 60.

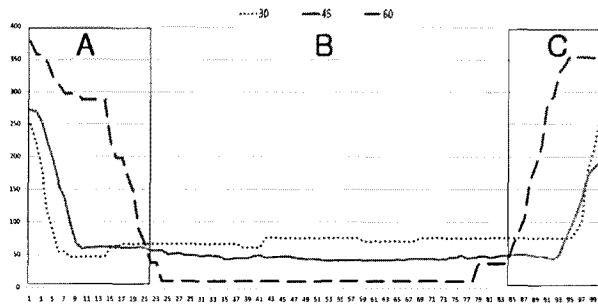


그림 13. 각도 변화 따른 추정 오차(cm)

Fig. 13. An estimated error by camera angle.

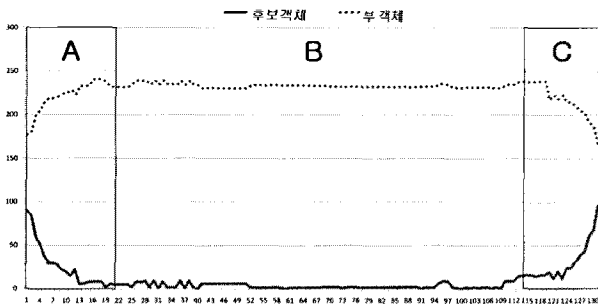


그림 14. 객체 선택 추정 결과

Fig. 14. Result of object selection estimation.

중심 값과의 거리를 측정된 결과이다. A영역은 주 객체가 후보객체를 지시하기 위해 행동을 취하고, B영역은 후보객체를 지시하고 있을 때의 선택결과 이다. C영역은 지시 행동을 하지 않았을 때의 결과를 보여주고 있다. 중심에 가까울수록 추정 오차는 5cm를 넘지 않는 결과로부터 60도에서 취득한 영상이 가장 좋은 결과를 보여주고 있다.

그림 14는 다수의 객체가 있을 때 주객체가 지시하는 후보객체와 지시 하지 않는 부 객체와의 지시 정확도를 측정된 결과이다. 두 결과 값은 후보객체와 부 객체 사이의 거리에 따라 주객체가 선택하고 있는 후보객체와 부 객체의 거리 차이를 보여 주고 있다.

따라서 주객체 주변에 여러 객체가 존재할 경우, 주객체가 지시하고 있는 후보 객체의 지시가 명확히 추정될 수 있음을 보여주고 있다.

VI. 결 론

본 논문에서는 다시점 카메라에서 인체 움직임 분석 기술을 사용하여 지시 동작을 추정하고, 이를 토대로 다른 객체를 효율적으로 선택하는 알고리즘을 제안하였다. 또한 실시간으로 처리하는 영상에서 다수의 객체를

검출하였고, 사람의 지시 행동을 인식하여 지시방향의 객체를 검출하고 추적하였다. 다시점 카메라를 사용하여 다양한 각도에 따른 객체 검출과 사람의 행동(지시)에 대한 실험 결과, 제안된 방법의 정확성을 확인할 수 있었다. 또한 다양한 실험 데이터를 통해서 다시점 카메라를 사용해 객체의 전체적, 부분적 특징 표현이 가능하였다.

향후 연구에서는 제안된 알고리즘을 응용하여 실시간 상호작용 응용시스템 및 실제 스포트라이트 조명을 사용하여 선택된 객체에 조명을 비추고 다양한 사람의 행동인식에 관한 연구로 진행할 예정이다.

참 고 문 헌

- [1] L. Wang, W. Hu, and T. Tan, "Recent developments in human motion analysis", *National Laboratory of Pattern Recognition*, Vol. 36, pp. 585-601, 2003.
- [2] A. Pentland, "Looking at people: Sensing for Ubiquitous and Wearable Computing", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, pp. 107-119, 2000.
- [3] R. Cipolla, P. A. Hadfields and N. J. Hollinghurst, "Uncalibrated stereo vision with pointing for man-machine interface", *Proc. of IAPR workshop on Machine Vision Application*, pp. 163-166, 1994.
- [4] M. Nishimura, A. Nishikawa, K. Koara, and F. Miyazaki, "A Multi-Modal Interface for Recognizing Gestures Expressed by Cyclically Repeated Motion of the Hand", *IAPR Workshop on Machine Vision Applications*, pp. 219-222, 2000.
- [5] E. Sato, A. Nakajima, T. Yamaguchi, and F. Harashima, "Humatronics(1) - natural interaction between human and networked robot using human motion recognition", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2794-2799, 2005.
- [6] A. Butz, M. Schneider, and M. Spassova, "Searchlight - A Lightweight Search Function for Pervasive Environments", *Proceedings of Int. conf. on pervasive computing*, 2004.
- [7] R. Hartly and A. Zisserman, "Multiple view geometry in computer vision 2 edition", Cambridge University Press, pp. 184-243, 2000.
- [8] M. Piccardi, "Background Subtraction Techniques: a reviews", *Int. Conf. on Systems*,

Man, and Cybernetics, Vol. 4, pp. 3099-3104, 2004.

- [9] 조상현, 최홍문, "솔더페이스트의 고속, 고정밀 검사를 위한 이차원/삼차원 복합 광학계 및 알고리즘 구현", *대한전자공학회 논문지* 제41권, SP편 제3호, pp. 139-146, 2004.

— 저 자 소 개 —



박 왕 배(정회원)
 2007년 2월 남서울대학교 멀티미디어학과 졸업(공학사)
 2009년 2월 중앙대학교 첨단영상대학원 영상학과 졸업(공학석사).

<주관심분야 : 콘텐츠, 멀티미디어>

서 응 호(정회원)
 대한전자공학회 논문지 제46권 SP편 제2호 참조

두 경 수(정회원)
 대한전자공학회 논문지 제45권 SP편 제6호 참조

최 종 수(평생회원)-교신저자
 대한전자공학회 논문지 제46권 SP편 제2호 참조