복수 등장인물을 대상으로 한 제스처 인식

Gesture Recognition in Multiple People Environment

홍석주, Seokju Hong, Nurul Arif Setiawan, 김송국, Songgook Kim, 김장운, Jangwoon Kim,
이칠우, Chilwoo Lee
전남대학교 컴퓨터정보통신공학과

요약 지금까지 진행된 제스처 인식 연구는 한 사람을 대상으로 정적인 환경을 가정하여 이루어져 왔다. 본 논문에서는 복수의 등장인물이 존재하는 환경에서 대화 상대를 선택하여 제스처를 인식하는 방법에 대해 기술한다. 먼저 복수의 인물이 존재하는 환경에서 배경영역을 제외한 행위자의 영역을 추출한다. 그 후 각각의 행위자를 트래킹하면서 카메라와 가장 가까이 있는 행위자를 대화 상대자로 선택한다. 대화상대자가 선택되면 스테레오 카메라에서 입력된 영상에서 추출된 실루엣 이미지를 이용하여 얼굴과 두 손을 특징 영역으로 하여 매 프레임마다 Kalman filter 를 사용하여 각 영역을 트래킹한다. 트래킹되는 특징 영역의 2 차원 좌표 값을 모델 제스처의 2 차원 좌표 값과 비교하여 가장 높은 유사값을 갖는 모델 제스처를 입력 제스처로 인식하게 된다. 본 논문에서 사용한 방법은 복수의 등장인물이 있는 경우 효과적으로 행위 대상자를 선택하여 제스처를 인식할 수 있다. 또한 제스처 인식에 있어서 단순한 큐 매칭을 사용함으로써 계산이 복잡하지 않은 장점이 있다. 실험을 통해 본 논문에서 제시한 방법을 적용함으로써 복수의 인물이 등장하는 환경에서 제스처 인식이 가능함을 입증한다.

핵심어: HCI, Gesture Recognition, Multiple People

1. 서론

최근 정보통신기술의 발전으로 기존의 특정 산업 분야에서만 사용되던 로봇이 홈 서비스 분야에서도 중요한 역할을 할 것이라고 보고 그에 대한 연구가 활발히 진행되고 있다. 이러한 연구활동의 결과로 2006년 국내에서 개최된 로보월드 2006에서는 세계 최초 연예인 로봇 에버투-뮤즈가 가수로 데뷔했으며 에버원은 개막식에서 사회를 보았고 휴보는 사람과 함께 테이프 커팅을 하는 모습을 보여주었다. 또한 기존의 제조업 로봇을 비롯해 서비스용 로봇, 군사용 로봇, 재난극복용 로봇, 교육용 로봇 등 인간과 밀접한 관련이 있는 분야의 로봇이 많이 소개되었다. 여기에서 로봇이 더 이상 일부 사용자에 의해서 사용되는 것이 아닌 인간과 밀접한 일반적인 서비스 분야에서 다양한 서비스를 제공하는 로봇으로 바뀌어 가고 있다는 것을 알 수 있다. 로봇이 일반적인 서비스 분야에서 널리 사용되기 위해서는 여러 사람이 있는 환경에서 각각의 사람을 인지하고 서비스 제공 상대를 선택하여 선택된 사람에게 서비스를 제공하기 위한 제스처 인식 기술이 필요하다.

지금까지 복수의 등장인물을 대상으로 한 제스처 인식 연구는 이루어지지 않았다. 단지 복수의 인물을 대상으로 한 추적에 관한 연구와 한 사람을 대상으로 한 제스처 인식 연구가 분리되어 이루어져 있을 뿐이다. 각각의 최근 연구동향을 살펴보면 다음과 같다.

먼저 기존의 비전 기반 복수인물 추적에 관한 연구 동향을 살펴보면 처리 알고리즘의 내용에 따라 확정적(deterministic)인 방법과 확률적(stochastic)인 방법으로 세분될 수 있다. 확정적인 영상추적(deterministic tracking)의 기본 원리는 입력 영상과 미리 확정된 모델 영상의 매칭을 반복적으로 수행하여 그 연속된 궤적을 추적해 가는 방식이다[1-4]. 모델 선정에 사용되는 중요 특징으로는 사람의 피부색, 외관(appearance) 정보, 개인 식별 정보, 신체 골격정보 같은 요소들이 많이 사용된다. 하지만 이러한 방법의 단점은 순간적으로 빠른 움직임이나 불연속성의 움직임이 발생할 때 추적에 실패하는 경우가 있다. 확률적인 영상추적(stochastic tracking)의 기본 원리는 영상으로부터 구해진 각종 특징 정보들을 확률적인 방법에 의해 처리하여 가장 가능성이 높은 궤적 또는 형상을 추적해가는 방식이다[5-7]. 주로 외적 환경정보와 사람의 움직임 정보를 동시에 이용하는 방법이 널리 쓰인다. 하지만 이러한 방법은 수학적인 계산이 필요하므로 사람의 수에 제한이 있고 사람의 위치를 추적하는데 많은 시간이 걸린다는 단점이 있다.

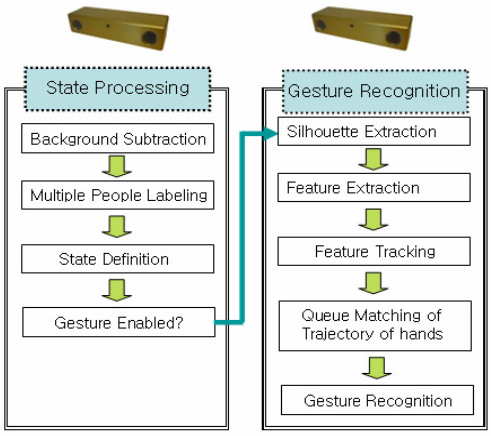

그림 1 시스템 구성도

기존의 비전 기반 제스처 인식 방법을 살펴보면 다음과 같다. 2차원 정보에 기반한 스킨 모델 방식은 단순히 스킨 정보를 추출하면 되므로 계산이 단순하다는 장점이 있지만 복잡한 배경이나 변화하는 조명에서는 힘들다는 단점이 있다[8]. 에지 정보에 의한 방식은 인체의 중심점으로부터 양손의 거리정보만 이용하므로 실시간 인식 시스템에 적합하지만 다른 제스처와 혼동하여 인식할 수 있는 단점이 있다[9]. 3차원 기반 제스처 인식 방법은 정교한 움직임 추출이 가능하지만 계산이 복잡하고 데이터 베이스 구축에 어려움이 따른다[10]. 또한 기존의 방식들은 공통적으로 복수의 등장인물이 아닌 행위자를 한 명으로 간주하고 연구가 진행되었다.

위에 설명한 방식으로 복수의 인물이 등장하는 경우의 추적방식과 제스처 인식 방식을 나누어 처리하고 각 결과를 합친다면 복수의 등장인물이 등장하는 환경에서 제스처를 인식하는 방법이 가능하다. 본 논문에서 실험한 시스템은 그림 1과 같이 복수의 인물을 추적하는 세그먼테이션 부분과 행위자를 선택하여 제스처를 인식하는 인식 부분으로 나누어져 있다. 그림 1과 같이 상태 처리 부분에서는 먼저 입력 영상에서 행위자가 존재하는 영역을 제외한 배경 영역을 제거한다. 그 후에 각각의 행위자의 상태를 인식한 후 거리가 가깝거나 특정 제스처를 취하는 행위자를 제스처 인식 대상자로 확정한다[11]. 제스처 인식 대상자가 결정되면 두 번째 스테레오 카메라에서 해당 대상자의 실루엣 이미지를 사용하여 행위자의 특징 점을 추출한다. 추출된 특징 점 중에서 양손에 해당하는 특징 점의 궤적을 추적하기 위하여 매 프레임마다 Kalman filter를 사용한다. 추출된 양손의 움직임의 궤적은 큐 매칭 방법을 이용하여 가장 비슷한 값을 가진 모델 제스처를 행위자의 제스처로 인식한다. 전체적인 시스템의 구성도는 그림 1과 같다.

2. 제스처의 정의 및 상황인식

이 장에서는 제스처 인식 시스템에서 인식하고자 하는 제스처를 정의한다. 다음으로 세그먼테이션 과정을 통해 복수의 등장인물을 추출하고 각 등장인물의 상태를 정의한다. 여러 등장인물 중 거리가 가장 가까운 대상자를 행위자로 결정하여 제스처 인식을 효율적으로 하는 상황인식기술에 대해서 설명한다.

2.1 제스처의 정의

인간은 일상생활에서 상대방과 대화하기 위하여 수많은 제스처를 사용한다. 제스처에 주로 사용되는 것으로는 신체의 움직임, 그 중에서 양손과 머리의 움직임이 많이 사용된다. 따라서 대부분의 제스처는 양손과 머리의 움직임을 사용하여 정의할 수 있다.

인간이 사용하는 수많은 제스처를 시스템이 이해하기는 힘들며 정의하기도 힘들다. 그래서 본 논문에서는 먼저 인식하기 위한 제스처를 정의하기 위해 다음과 같은 시나리오를 만들었다.

"철수와 영희는 일요일 오후 학교 숙제를 하기 위해 박물관에 갔다. 학교 숙제를 하기 위해 고구려 관련 소개 코너로 가야 하는데 박물관이 워낙 커서 찾기가 힘들었다. 저 멀리 안내 로봇이 보이길래 그쪽으로 다가갔다. 근처에 2-3명이 있었지만 로봇은 철수와 영희가 다가오자 방향을 돌렸다(대화상대 선택). 철수는 로봇에게 가까이 오라고 손짓했다(come here). 로봇이 가까이 다가오자 철수는 반가운 마음에 악수를 청했다(shake hands). 고구려 사 관련 코너가 어디에 있는지 어떻게 가야 하는지 로봇에게 물어보았다. 로봇은 2층 중간에 있다고 상세하게 알려주었다. 옆에서 지켜보던 영희가 다가와(대화상대 변경) 로봇에게 감사의 표시를 하고(heart) 헤어졌다(bye bye).

이와 같은 시나리오를 사용하여 다음과 같은 5가지의 제스처를 정하고 각 제스처의 특징을 표 1에 정의하였다.

표 1

제스처	정의
come here	양손을 아래에서 위로 흔드는 동작
stop	양손을 어깨 높이에서 멈추는 동작
shake hands	오른손을 내밀고 위/아래로 흔드는 동작
heart	양손을 좌우로 벌려 위로 올리는 동작
bye bye	오른손을 들어 좌우로 흔드는 동작

2.2 행위자의 추출

일반적인 환경에서 획득된 이미지에는 수많은 물체들이 포함되어 있고 시간에 따라 조명 변화가 생기는 등 많은 잡음을 가지고 있다. 또한 복수의 등장인물이 영상 내로 진입하는 경우 등장인물끼리 서로 겹쳐서 서로 가리기 때문에 각 등장인물을 제대로 추출하는 것 또한 문제이다.

본 논문에서는 [11]에 제시된 알고리즘을 사용하여 스테레오 카메라부터 입력된 영상에서 복수의 등장인물을 추출하였다. 여기에서 사용된 알고리즘은 개선된 HLS 공간상에서 가우시안 잡음을 사용하므로 입력된 영상에서 실시간으로 행위자를 추출하기에 적합하다. [11]에 제시된 알고리즘을 사용하면 입력된 영상에서 각 등장인물의 실루엣 이미지를 얻을 수 있으며 그림 2를 통해 확인할 수 있다.

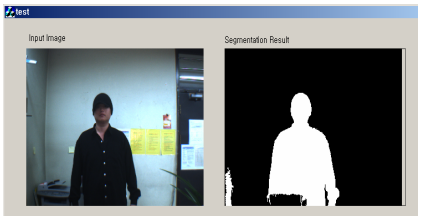

그림 2 세그먼테이션 결과

2.3 상황인식 및 대화자의 선택

일상생활에서 상대방과 대화할 때 보통 특정 제스처를 취하거나 상대방에게 다가간 후 대화를 시작한다. 본 논문에서는 이러한 사실을 이용하여 제스처 인식 대상자를 확정하였다. 또한 입력 영상에 최대 3-4명의 등장인물이 존재한다고 가정하고 각 행위자에게 다음과 같은 5가지의 상태를 정의하였다.

↲표 2

상태	정의
Inactive	행위자가 움직이지 않음
Active	행위자가 움직이고 있음
Walking	행위자가 걸어가고 있음
Running	행위자가 뛰어다니고 있음
Gesture Recognition Enabled	행위자가 거리가 가까운 경우 True 가 됨

2.2장에서 각 등장인물의 실루엣 영상에서 라벨링된 각 행위자에게는 표2와 같은 상태 값이 정의되어 있다. 각각의 행위자 중 카메라와의 거리가 가까워지는 행위자는 [Gesture Recognition Enabled] 상태가 [True]가 되면서 해당 행위자를 제스처 인식 대상자로 결정한다.

3. 특징점 추출 및 트래킹

2장에서 설명한 방식으로 제스처 인식 대상자를 결정하게 되면 인식하기 위한 전 단계로 인체의 주요 특징점을 추출한다. 주요 특징점 중에서 머리와 양손의 특징점이 제스처를 인식하기 위한 것으로 각 점을 Kalman filter를 사용하여 지속적으로 트래킹한다.

3.1 특징점 추출

본 논문에서는 세그먼테이션의 결과 이미지인 실루엣 영상을 사용하여 인체의 주요 특징점을 추출하였다. 특징점을 추출하기 위해서 먼저 원본 영상에서 [11]에 제시되어 있는 알고리즘을 적용하여 그림 2와 같은 실루엣 영상을 추출한다. 추출된 실루엣 영상은 배경과 행위자의 경계 사이에 노이즈가 심하게 발생하므로 식(1)과 같은 팽창 영산을 적용하여 노이즈를 없앤다.

$$A \oplus B = \{z \mid [(\hat{B})_z \cap A] \subseteq A\} \quad (1)$$

노이즈를 없앤 영상에서 행위자가 차지하는 영역은 흰색으로 보인다는 점을 이용하여 라벨링을 한 후 경계 점을 추출한다. 추출된 경계 점을 서로 이어주면 그림 3과 같이 행위자의 제스처에 따른 외곽선을 볼 수 있다.

제스처를 인식하기 위해서 인체의 주요 특징점을 행위자의 좌우 최하단점(LB, RB), 양손(LH, RH), 머리 부분(HP), 인체의 중심점(CP)의 총 6가지의 주요 특징점을 정의하였다. 정의된 특징점을 추출하기 위해 사용한 방법은 다음과 같다. 그림 4에서 행위자의 [LB] 위치를 탐지하기 위해서 외곽선에서 추출된 점의 위치를 검색한다. 각 위치 값을 검색하면서 X값이 가장 낮고 Y 값이 가장 높은 점을 [LB]로 정의한다. 다음으로 행위자의 [RB]위치를 검색하기 위해서 외곽선의 각 점에서 Y 값이 가장 높은 점을 검색한 후 추출된 점들 중에서 X값이 가장 높은 것을 [RB]로 정의한다. 사람의 머리에 해당하는 [HP]의 위치는 외곽선의 각 점에서 [LB], [RB] 위치 사이에 있는 점들 중에서 Y값이 가장 낮은 것으로 정의된다. [CP]의 위치는 [HP]의 X축 선과 [LB], [RB]가 이루는 선의 중점으로 정의된다. 양 손의 위치를 나타내는 [LH], [RH]은 X값이 가장 낮은 것과 가장 높은 것으로 정의된다.

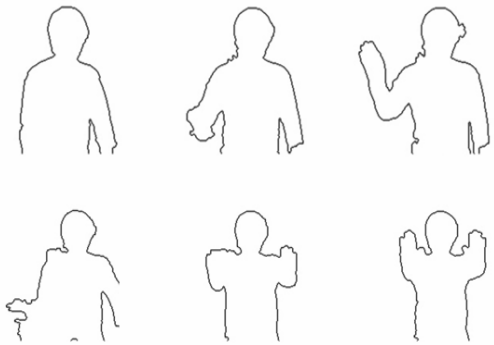

그림 3 행위자의 외곽선 추출

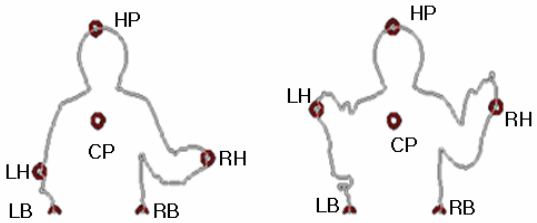

그림 4 인체의 주요 특징점 추출

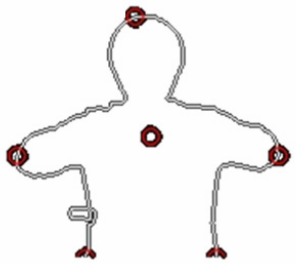

그림 5 양손 위치의 잘못된 추출

3.2 특징점 트래킹

하지만 위와 같은 방법으로 특징점을 추출하면 그림 5와 같이 손의 영역이 행위자와 몸통 부분과 겹치면 잘못된 부분을 손의 위치로 추출한다. 이러한 문제를 해결하기 위해서는 양손의 위치가 급격하게 변할 때 기존의 위치를 사용하여 양손의 위치를 추정해야 한다. 본 논문에서는 양손의 위치를 지속적으로 트래킹하기 위하여 Kalman filter를 적용하였다.

Kalman filter는 이동 중인 물체 추적에 주로 사용되는 알고리즘으로 순환적 선형구조를 가지고 있고 알고리즘이 단순하며 시간 영역에서의 직접적인 처리가 가능하므로 양 손의 점을 추적하기에 적합한 방법이다. 본 논문에서는 다양한 방식 중에서 이산형 Kalman filter를 적용하였다.

Kalman filter 방정식은 time update과정과 measurement update 과정으로 나뉜다. Time update 과정에서는 다음 단계를 위해 이전 평가를 얻어 현재의 상태와 에러 공분산을 예측한다. Measurement update 과정에서는 새로운 측정값을 이전 추정에 포함시켜 개선된 다음 평가 값을 얻는다.

Time update 과정은 예측 방정식이라 볼 수 있고 measurement update 과정은 정정 방정식이라 볼 수 있다.

Time update 과정을 식으로 나타내면 다음과 같다.

$$\hat{x}_{k+1} = A_k\hat{x}_k + Bu_k \quad (2)$$

$$P_{k+1}^{-} = A_k P_k A_k^T + Q_k \quad (3)$$

여기에서 $\hat{x}_k$ 는 시간 k에서의 측정 이후 추정 상태를, P는 에러 공분산을 나타낸다. Measurement update 과정을 식으로 나타내면 다음과 같다.

$$K_k = P_k^{-} H_k^T (H_k P_k^{-} H_k^T + R_k)^{-1} \quad (4)$$

$$\lim_{R_k \to 0} K_k = H_k^{-1} \quad (5)$$

$$\hat{x}_k = \hat{x}_k^{-} + K(z_k - H_k\hat{x}_k^{-}) \quad (6)$$

$$P_k = (I - K_k H_k)P_k^{-} \quad (7)$$

여기에서 K_k 는 Kalman gain으로 상태값과 에러 공분산을 갱신하는데 사용된다. 그림 6에서 보는 바와 같이 Kalman filter는 time-measurement 과정 즉 예측과 정정의 과정을 계속 반복한다. 예측 단계에서는 식(2,3)를 이용하여

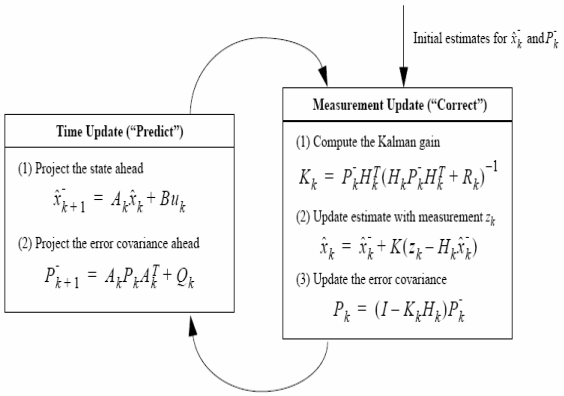

그림 6 Kalman filter 순서도

상태값과 에러 공분산을 예측하고 정정 단계에서는 식(4)를 이용하여 먼저 Kalman gain을 계산하고 실제 측정값을 이용하여 식(6)에 의해 상태 값을 갱신한다. 그리고 식(7)을 이용하여 에러 공분산을 수정한다.

지금까지의 과정을 순서도로 나타내면 그림 6과 같다. 그림 6에서 초기값과 에러 공분산이 정해지면 예측-정정단계를 지속적으로 수행하여 실제 측정값과 예측 값의 에러를 줄이는 것을 확인할 수 있다.

Kalman filter를 적용하여 머리와 양손의 위치를 추적하기 위한 과정을 요약하면 다음과 같다. 인체의 주요 특징점 중에서 머리와 양 손의 실제 위치를 Kalman filter의 측정값으로 정의한다. 매 프레임 머리와 양 손의 실제 위치를 탐지하면서 Kalman filter의 상태 값과 에러 공분산을 수정한다. 만약 탐지된 머리와 양 손의 실제 위치가 급격하게 바뀌면 Kalman filter에서 추정한 값을 머리와 양 손의 위치로 사용하여 해당 점을 특징점으로 간주한다.

4. 큐 매칭에 의한 제스처 인식

사람의 제스처는 몸 전체의 움직임에 행위자의 의도가 포함되어 있다. 특히 얼굴과 양손의 이동 경로에 많은 의미를 포함하고 있기 때문에 본 논문에서는 양손의 궤적 정보를 모델 제스처와 비교하여 제스처를 결정하는 방법을 사용한다.

2차원 공간상의 궤적을 매칭하는 방법은 지금까지 다양하게 시도되어 왔다. 특히 필기체 인식과 같은 분야에 많이 이용되고 있지만 제스처 인식의 경우 행위의 시작과 끝을 결정할 수 없기 때문에 효과적으로 적용되기 어렵다. 이는 인식의 수행 시점을 결정하는데 있어서 중요한 역할을 수행하며, "Gesture Spotting" 이라는 주제로 많은 연구가 진행되고 있다[10, 12].

본 논문에서는 복잡한 제스처가 존재하지 않는다는 가정하에 spotting 알고리즘 대신 간단한 큐 비교 알고리즘을 사용한다. 이 방법은 처리 속도에 상당한 이점을 가지며, 구현이 용이하다는 장점을 가지고 있다. 알고리즘의 기본 개념은 다음과 같다. N 개의 모델을 갖는 모델 집합 M 을 가정하자. 각 제스처 모델(G_j)에는 식(8)과 같이 양손의 이동 경로가 k 개의 방향 벡터로 표현되어 차례로 저장되어 있다.

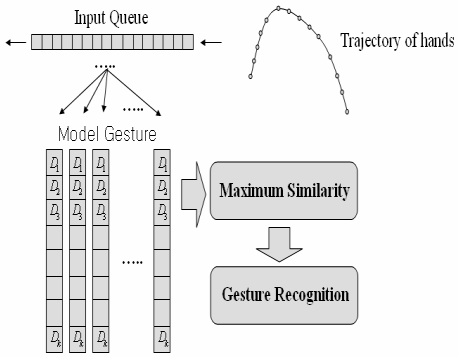

그림 7 모델 제스처와 입력 영상의 큐 매칭 비교

$$Q_{model} = \{N_1, N_2, N_3, ..., N_k\} \quad (8)$$

$$M = \{G_1, G_2, G_3, ..., G_j .., G_N\} \quad (9)$$

$$G_j = \{D_1^j, D_2^j, D_3^j, ..., D_k^j\} \quad (10)$$

$$D_i^j = N_{i+1}^j - N_i^j (N_{i+1}^j, N_i^j \in Q_{model}) \quad (11)$$

한편, 길이가 l 인 입력 큐 I 는 일정 시간 간격으로 입력되는 방향 벡터의 집합이며, 식(14)과 같이 표현된다.

$$Q_{input} = \{N_1, N_2, N_3, ..., N_l\} \quad (12)$$

$$I^j = N_{i+1}^j - N_i^j (N_{i+1}^j, N_i^j \in Q_{input}) \quad (13)$$

$$I = \{I^1, I^2, I^3, ..., I^l\} \quad (14)$$

만약, 입력 큐 내에 행위자의 의도가 담긴 제스처가 포함되어 있다면, 이를 의미 있는 제스처로 인정하고, 모델 큐와의 비교를 위해 식(15)에서와 같이 각각의 제스처 모델 G_j 를 입력 큐 I 와 비교한 뒤, 최대 유사도를 갖는 제스처 G_k ($G_k \in$ M)로 결정한다. 이때 선택된 제스처를 행위자의 제스처로 인식하고 결정하게 된다.

$$R_j = \max_m (\sum_{i=1}^{k} (I^{m+i} \bullet D_i^j)) \quad (15)$$

$$K = \arg\max(R_j) \quad (16)$$

그림 7에 입력된 제스처가 연속적으로 큐에 입력되고 모델 제스처와 비교되면서 제스처가 결정되는 과정이 나와있다.

5. 실험결과 및 문제점

본 논문에서 제안한 방법을 실험하기 위해 Intel Pentium 4 3.0GHz CPU와 512MB RAM을 사용하였으며 스테레오 정보 추출을 하기 위하여 PointGrey 사의 Bumblebee를 이용하였다. 프로그램 환경은 Visual C++ 6.0과 OpenCV 1.0을 사용하였다. 그림 8을 보면 현 시스템에서 각각의 제스처를 취했을 때 인체의 주요 특징점을 추출한 결과를 볼 수 있다.

그림 8 샘플 제스처와 특징점 추출 결과

본 시스템에서는 제스처 인식을 단지 2차원 좌표 값을 사용하여 비교함으로써 문제가 발생했다. 예를 들어 아래에서 수직으로 손을 올리는 제스처와 원 모양을 그리면서 손을 올리는 제스처를 서로 잘 못 인식하는 문제가 자주 발생하였다. 이런 문제는 2차원 좌표 값이 아닌 3차원 좌표 값을 사용하여 제스처 인식에 사용한다면 더욱 효과적으로 제스처를 인식할 수 있을 것이다. 다음으로 악수를 하는 제스처와 잘 가의 제스처의 경우 일부 영역에서 움직임이 존재하지만 궤적이 생기지 않아 제스처를 인식하지 못하는 문제점이 발생하였다. 이러한 문제점은 제스처를 인식하는데 있어 일부 영역에 멈추어 있는 시간정보와 움직임 정보를 이용한다면 제스처를 인식할 수 있을 것이다. 또한 인체의 주요 특징점을 추출하는데 있어 외곽선의 전체 점을 대상으로 검색하는 것보다 Convex Hull 알고리즘을 사용하여 외곽선에서 특징점이 될 수 있는 점들을 먼저 선별한 후 추출된 점에서 특징점을 검색하면 본 논문에서 제시한 방법보다 더 빠르고 정확한 특징점을 얻을 수 있을 것이다.

또한 행위자가 바뀌는 경우 이미 정의된 제스처라 하더라도 행위자에 따라 틀린 궤적 값을 가지는 경우 인식하지 못하는 문제가 발생하였다. 이러한 문제는 먼저 각각의 행위자마다 ID를 부여한 후 해당 행위자의 제스처를 학습하도록 하여 모델 제스처를 구성한다. 그 후 각 행위자의 ID를 판별하여 해당 행위자의 ID로 저장되어 있는 모델 제스처를 검색하여 제스처를 인식한다면 행위자에 따른 제스처 인식의 문제점을 해결할 수 있을 것이다.

6. 결론

본 논문에서는 복수의 인물이 존재하는 경우 각 행위자의 상태를 정의하고 제스처 인식 대상자를 선택하여 제스처를 인식하는 방법에 대하여 기술하였다. 시스템의 효율성과 정확성을 위해 행위자의 영역을 추출하는 부분과 제스처를 인식하는 부분으로 나누었다. 제스처를 인식하는 도중에도 행위자의 주변 상황의 변화를 인식하여 다른 행위자가 진입하여 제스처 인식 대상자를 바꿀 수 있도록 하였다. 또한 제스처를 인식하는데 있어 주요 특징점인 양손의 머리 영역을 Kalman filter를 사용하여 지속적으로 추정하여 제스처를 인

식할 수 있도록 하였다. 또한 제스처 인식을 함에 있어서 간단한 큐 매칭 알고리즘을 사용하여 실시간으로 제스처 인식을 하는데 어려움이 없도록 하였다.

본 논문에서 제안된 방법은 일반적인 서비스 환경에서의 로봇의 일반화된 제스처 인터페이스로 활용될 수 있으며 앞서 언급한 문제점들이 해결된다면 지능형 로봇이 인간과 의사소통을 하는데 있어 자연스러운 커뮤니케이션을 할 수 있을 것이라 생각된다.

Acknowledgement

본 연구는 정통부 및 정보통신연구진흥원의 정보통신선도기반기술개발사업과 한국 과학재단 지정 전남대학교 고품질 전기전자부품 및 시스템 연구센터, 문화관광부 지정 전남대학교 문화콘텐츠 기술연구소의 일부 연구비 지원에 의해 수행되었음

참고문헌

[1] Tao Zhao, Ram Nevatia " Tracking Multiple Humans in Crowded Environment" in Proceedings of CVPR 2004, pp. 1063-6919, 2004.

[2] Bo Wu, Ram Nevatia, “Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors,” in Proceedings of ICCV 2005, vol. 1, pp. 90-97.

[3] Ismail Haritaoglu, David Harwood, Larry S. Davis, “W4: Real-Time Surveillance of People and Their Activities”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No.,8, pp. 809-830, 2000.

[4] Nils T Siebel, Steve Maybank, “Fusion of Multiple Tracking Algorithms for Robust People Tracking”, in Proceeding of ECCV 2002, pp. 373-387, 2002.

[5] Jerome Berclaz Franc, ois Fleuret, Pascal Fua, “Robust People Tracking with Global Trajectory Optimization,” in Proceedings of CVPR 2006, Vol 1, pp. 744-750.

[6] Jerome Berclaz Franc, ois Fleuret, Pascal Fua, “Robust People Tracking with Global Trajectory Optimization,” in Proceedings of CVPR 2006, Vol 1, pp. 744-750.

[7] Hieu T. Nguyen, Qiang Ji, Arnold W.M. Smeulders, “Robust multi-target tracking using spatio-temporal context,” in Proceedings of CVPR 2006, Vol. 1, pp. 578-585.

[8] Junwei Han, George M. Award, Alistair Sutherland, and Hai Wu, "Automatic Skin Segmentation for Gesture Recognition Combining Region and Support Vector Machine Active Learning", in Proceedings of FGR 2006, pp. 237-242.

[9] Hong Li, Michael Greenspan, "Multi-scale Gesture Recognition from Time-Varying Contours", in Proceedings of ICCV 2005, Vol 1, pp. 236-24.

[10] Seong-Whan Lee, "Automatic Gesture Recognition for Intelligent Human-Robot Interaction", in Proceedings of FGR 2006, pp. 645-650.

[11] Nurul Arif Setiawan, Chil-woo Lee, “Gaussian Mixture Model in Improved HLS Color Space for Human Silhouette Extraction”, in Proceedings of ICAT 2006, pp. 732-741, 2006.

[12] Ho-Sub Yoon; Byung-Woo Min; Jung Soh; Young-Iae Bae; Hyun Seung Yang, "Human computer interfacefor gesture-based editing system", Image Analysis and Processing, pp. 969 - 974, 1999.