

# 헬스케어를 위한 몸동작 인식 시스템의 설계 및 구현

김영운\*, 소인미\*, 강선경\*, 정성태\*

\*원광대학교 컴퓨터공학과

e-mail:kyw1007@wonkwang.ac.kr

## Design and Implementation of Body Action Detection System For Healthcare

Young-Un Kim\*, In-Mi So\*, Sun-Kyung Kang\*, Sung-Tae Jung\*

\*Dept of Computer Engineering, Wonkwang University

요 약

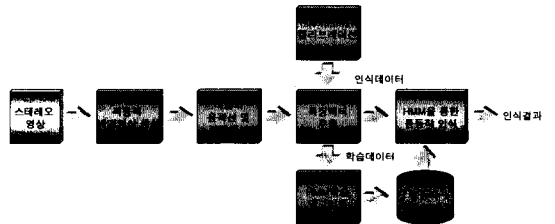
본 논문은 사람의 움직임에 따라 다양한 형태의 몸동작이 발생하고 이러한 몸동작을 이용해 헬스케어 정보로 사용될 수 있도록 몸동작 인식 시스템을 제안하였다. 몸동작 인식 시스템은 입력된 영상으로부터 몸동작의 특징 정보를 추출하여 학습에 필요한 데이터로 사용된다. 학습과정을 거쳐 다양한 몸동작 데이터베이스를 만들고 이를 이용해 몸동작 인식을 수행한다. 본 논문에서 사용된 스테레오 영상 정보는 두 대의 카메라로부터 입력된 영상에서 배경과 객체를 분리하고 분리된 객체에 대해 사각형 영역을 생성한다. 각각의 사각형 정보를 이용해 캘리브레이션 과정을 거쳐 3차원 특징값을 얻을 수 있다. 학습과 인식에 필요한 알고리즘으로 HMM이 사용 되었고 HTK를 이용해 구현 하였다. 실험에 사용된 특징 정보는 몸동작에 따른 사각형의 너비와 높이의 변화량, 사각형 중심점 위치의 변화량 등 2 가지 특징값을 이용해 각각 실험 하였다. 실험 결과 사각형의 너비와 높이의 특징값을 이용하는 것보다 중심점의 3차원 위치 변화량을 이용하는 것이 높은 인식률을 보였다.

### 1. 서론

2000년 고령사회의 진입과 가족 구성원의 보호를 받지 못하는 독거노인의 수가 급증함으로써 개인 중심의 예방 의료가 필요하게 되고 언제 어디서나 서비스를 받을 수 있는 건강관리기술이 필요하게 되었다. 따라서 댁내 거주하는 구성원의 인체 동작 상태를 구분하는 상황인식 기술이 필요하게 된다. 지금까지 연구되고 개발된 상황인식시스템 들을 살펴보면 생체신호 센서를 장착하여 인체동작 상태를 구분 정보로 활용하거나 웨어러블 컴퓨터의 입력 장치를 활용하기도 하고 가속도계를 이용하여 인체동작상태의 상황인식을 하기도 한다[1,2,3,4]. 컴퓨터가 사람의 몸동작이나 손짓을 자동으로 인식이 가능하다면 사람과 컴퓨터간의 편리한 통신 매체 수단이 될 수 있고 좀 더 편리한 인터페이스와 환경을 제공할 수 있을 것이다[5]. 또한 사용자의 생체신호 계측 및 자동진단, 응급경보가 가능한 헬스케어형 의료정보 시스템의 일부분으로 응용되어 사고발생건의 예방 및 발생 직후의 즉각적인 대응을 가능하게 할 수 있

다. 따라서 본 논문에서는 비접촉형 데이터로 영상 데이터를 이용하여 몸동작을 인식함으로써 응급상황을 자동으로 감지하는 시스템을 제안한다.

다음 (그림 1)은 몸동작 인식 시스템의 전체적인 흐름을 나타내고 있다. CCD 카메라를 통해 영상 데이터 획득하여 배경과 객체를 분리하고, 스테레오 카메라를 캘리브레이션하여 각 카메라를 통해 획득된 영상 데이터의 깊이 정보를 포함한 3차원 값을 얻을 수 있다.



(그림 1) 몸동작 인식 시스템

스테레오 카메라 영상에서 각각 배경을 생성한 다음 배경 영상과 입력 영상의 차이를 이용하여 움직이는 객체를 배경영상에서 분리하고 객체의 외곽선을 추출한 뒤 움직이는 객체를 포함하는 사각형 영역을 생성하여 객체를 추적하도록 하였다. 두 카메라

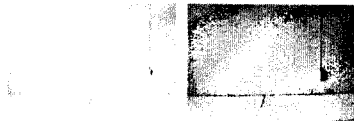
“이 논문은 2006년 정부(교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임” (지방연구중심대학육성사업/헬스케어기술개발사업단)

라의 캘리브레이션 정보를 이용하여 3차원 정보를 추출하였다. 3차원 공간상에서의 사각형 너비와 높이의 변화량과 사각형 중심점 위치의 변화량에 각각에 대하여 동작 인식을 실험을 하였다. 또한 본 논문에서는 학습과 인식 방법으로, 확률에 따른 패턴의 일련을 모델링한 HMM(Hidden Markov Model)을 이용하였고, HTK(Hidden markov Toolkit)를 사용하여 학습과 인식 실험을 하였다.

## 2. 특징정보 추출

### 2.1 배경과 객체의 분리

영상 획득 방법으로 2대의 컬러 CCD 카메라로부터 입력 받아 DVR(Digital Video Recorder) 시스템으로 전송해 채널별 352×240 크기의 영상 정보를 획득한다. 안정적인 배경영상을 생성하기 위하여 먼저 중간값 필터를 이용해 잡음을 제거한 후 컬러영상을 흑백영상으로 변환한다. 전처리된 영상을 N 프레임까지 받아들여 각 픽셀 값들을 누적 한뒤 N으로 나누어 배경영상의 각 픽셀에 대한 평균값을 구한다. (그림 2)은 프레임을 누적해 평균값을 적용한 스테레오 배경 영상을 보이고 있다.



(그림 2) 평균값을 적용한 배경 영상 (좌측/우측)

배경 영상과 몸동작 정보를 분리 하기위해 미리 처리된 배경 영상과 몸동작 영상 정보가 사용된다. 배경 영상의 전처리 방법과 동일하게 처리한 후 흑백으로 변환된 픽셀 값을 임계값과 비교하여 이진 영상을 생성하였다. (그림 3)는 중간값 필터와 흑백 처리가 적용된 영상을 나타내고 있다.



(그림 3) 필터와 흑백 처리된 영상 (좌측/우측)

몸동작 객체를 추출 하기위해 배경영상에서 추출된 각 픽셀 값과 전 처리된 몸동작 영상의 픽셀 값을 이용해 몸동작 영역을 분리해 낸다. 이를 위해 먼저 아래 식과 같이 몸동작 객체 영상을 생성한다.

$$G_{(x,y)} = \begin{cases} 255 & |Background_{(x,y)} - Object_{(x,y)}| > TH \\ 0 & \text{Otherwise} \end{cases}$$

식(1)과 같이 배경영상의 각 픽셀의 평균값에서 몸동작 영상의 각 픽셀 값을 뺀 후 기준 값 보다 크면 흰색(255)으로, 반대로 작거나 같으면 검정색(0)으로

설정한다. 흰색(0)값이 객체이며, 검정색(0)값이 배경이 된다. 그 결과 (그림 4)와 같이 배경 영상과 몸동작 객체가 분리되어 표시된다.



(그림 4) 배경과 객체가 분리된 영상 (좌측/우측)

잡음이 제거된 몸동작 객체 이미지에서 윤곽선(Contour) 정보를 찾는다. 윤곽선이란 움직임 객체의 가장자리 선을 말하는데, 본 논문에서는 객체의 가장자리에 있는 한 픽셀로부터 시작하여 가장자리에 있는 픽셀을 계속 따라가다가 시작점에 도착하면 추적을 멈추는 방법을 사용하였다. (그림 5)는 몸동작 객체의 윤곽선 정보를 추출한 결과 화면이다.



(그림 5) 몸동작 객체의 윤곽선 추출 (좌측/우측)

마지막으로 (그림 6)과 같이 추출된 몸동작 객체의 윤곽선 정보를 이용해 사각형 영역을 추출해 한다. 이 사각형 영역 정보를 이용해 2차원 정보를 획득하고 캘리브레이션 과정을 통해 3차원 정보를 획득해 특징값으로 사용한다.



(그림 6) 윤곽선의 사각형 영역 추출 (좌측/우측)

### 2.2 캘리브레이션

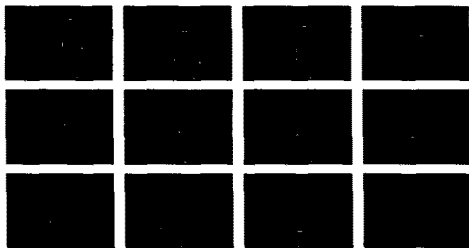
캘리브레이션은 카메라들간의 기하학적인 관계를 활용하여 정확한 3차원 정보를 얻기 위한 방법으로 외부 카메라 모델(extrinsic camera model)과 내부 왜곡 모델(intrinsic distortion model)을 결정하는 과정이다.

일반적으로 카메라 캘리브레이션 과정은 다음의 세 단계에 의해 수행된다. 우선 3차원 실세계 좌표와 그에 대응되는 2차원 영상 좌표의 집합을 획득한다. 다음으로 주어진 좌표 집합에 대해 일련의 방정식을 적용해 각각의 카메라에 대한 카메라 파라미터 값들을 계산한다. 마지막으로 투영 행렬(projection matrix)로부터 epipolar geometry를 구성할 수 있다 [6]. 본 논문에서는 Zhang의 캘리브레이션 알고리즘의 수정 버전에 근거하여 캘리브레이션 과정을 수행하였다[7]. 캘리브레이션 과정은 최종적으로 world coordinate의 기준점으로부터 카메라까지의 회전 및

이동 행렬을 제공한다. 이렇게 얻어진 카메라의 외부 파라미터를 이용하여 실세계에 대한 3차원 좌표를 얻을 수 있다.

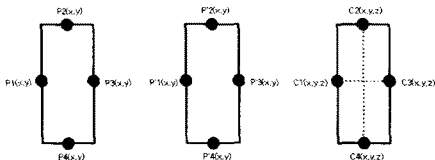
2.3 특징벡터 추출

본 논문에서는 두 종류의 특징값을 추출하여 실험하였다. 첫 번째 특징값은 사각형 영역의 너비와 높이의 크기 변화량에 대한 값이다. 두 번째 특징값은 사각형 영역의 중심점(x, y, z)의 이동 변화량에 대한 값이다. 특징값은 3이상의 변화량을 가질 때에만 특징값으로 사용하였다. 이는 잡음으로 인한 미세한 변화에 대해서는 특징값을 취하지 않도록 하여 오류를 처리하기 위함이다. (그림 7)은 기절동작에 대한 사각형 영역을 표시한 그림이다. 시간의 흐름에 따라 사각형 영역의 너비와 높이가 변화고 중심점이 이동되는 것을 알 수 있다.



(그림 7) 기절동작

(그림 8)은 2차원 좌표점을 이용해 3차원 특징점인 너비와 높이를 계산 하는 방법을 나타내고 있다. 3차원 특징 정보는 좌측 영상의 사각형 선분의 중심점 4개(P1, P2, P3, P4)의 x, y 좌표점과 우측 영상의 사각형 선분의 중심점 4개(P'1, P'2, P'3, P'4)의 x, y 좌표점을 캘리브레이션한 후 4개(C1, C2, C3, C4)의 3차원 좌표점 (x, y, z)값을 획득한다. 이 좌표점을 이용해 사각형 영역의 너비(C1, C3)와 높이(C2, C4)의 변화량을 계산할 수 있다.



(그림 8) 2차원 좌표점을 이용한 3차원 특징값(너비, 높이) 추출

(그림 9)은 2차원 좌표점을 이용해 3차원 특징점인 (x, y, z)값을 계산 하는 방법을 나타내고 있다. 3차원 특징 정보는 좌측 영상의 사각형 중심점 P(x,y)와 우측 영상의 사각형 중심점 P'(x,y)를 캘리브레이션 처리를 통해 C(x, y, z)와 같은 3차원 좌표값을 얻을 수 있다. 이와 같은 3차원 좌표값을 이용해 x축, y축, z축의 객체 이동에 따른 변화량을 계산해

특징값을 획득 하였다.

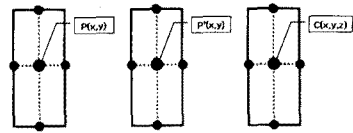
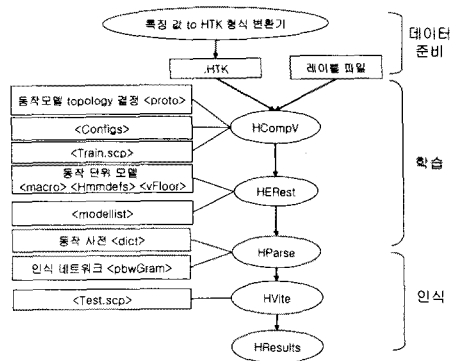


그림 9. 2차원 좌표점을 이용한 3차원 특징값(x, y, z) 추출

2.4 HMM 알고리즘을 이용한 몸동작 학습 및 인식

HMM은 다중 확률구조를 갖는 프로세스를 모델링 하는데 매우 적합한 방법이다. (그림 10)은 본 논문에서 사용된 학습과 인식과정을 잘 나타내고 있다. 먼저 특징값을 추출해 HTK 파일을 생성하고 학습에 필요한 레이블 파일, 스크립트 파일 등을 만든 후 HTK 파일과 레이블 파일 등을 이용해 HCompv, HERest 처리 과정을 거쳐 학습을 수행하고 이 과정의 결과로 몸동작 데이터베이스가 만들어진다.



(그림 10) HTK를 이용한 학습과 인식

학습이 완료되면 인식 테스트를 한다. 테스트를 수행하기 위해 학습된 몸동작 데이터베이스와 동작 사전, 스크립트 파일 등이 필요하며, HParse, HWrite, HResults 처리 과정을 통해 최종 인식결과를 얻을 수 있다.

3. 실험결과

본 논문에서 실험한 실험환경은 일반적인 사무실의 자연 조명에서 영상을 촬영 했으며, DVR장비를 이용해 스테레오 카메라로부터 입력을 받아 실시간으로 영상 데이터를 획득 하였다. 몸동작 인식시스템은 입력 받은 영상에서 특징값을 추출하고 학습과 인식에 필요한 데이터를 생성해 주며, 최종 인식의 결과 값을 확인 할 수 있도록 되어있다. 개발 환경은 Microsoft Visual C++ 6.0 언어를 사용 하였고 Microsoft WindowsXP Platform 환경에서 실험이 수행되었다.

실험에 사용된 몸동작은 총 9개(서 있다가 기절하기, 걷기, 뛰기, 의자에 앉기, 의자에서 일어시기, 허

리 굽혀 줍기, 제자리에서 점프하기, 서 있다가 바닥에 앉기, 앉아 있다가 일어서기)이다. 학습은 9동작 모두 사용 하였으며, 인식실험은 기절동작만 사용하였다. 사람은 7명분이며 각각 9동작씩 총 63개의 동작이 실험에 사용 되었다. 10번 반복 중 첫 번째부터 일곱 번째는 학습 데이터로 사용되었으며, 여덟 번째에서 열 번째는 인식 실험 데이터로 사용되었다. 실험은 특징값 추출방식에 따라 3가지로 구분해 진행하였다. 첫 번째는 객체의 사각형 너비와 높이에 대한 크기 변화량으로 5회 실시 하였으며, 인식 결과는 <표 1>과 같이 평균 90.47%의 인식률을 보였다. 두 번째는 사각형의 중심점 좌표(x, y)의 변화량으로 5회 실시 하였으며, 인식 결과는 <표 2>와 같이 평균 91.42%의 인식률을 보였다. 세 번째는 사각형의 중심점 좌표(x, y, z)의 변화량으로 5회 실시 하였으며, 인식 결과는 <표 3>과 같이 평균 94.28%의 인식률을 보였다. 실험 결과에서 보는바와 같이 몸동작에 따른 움직임 객체의 사각형 중심점 좌표(x, y, z)의 변화량 대한 특징값을 사용하는 것이 더 좋은 인식률을 보였음을 알 수 있었다.

<표 1> width, height 값의 변화량

학 습										테스트	인식률
L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	95.24%	
L3	L4	L5	L6	L7	L8	L9	L10	L1	L2	95.24%	
L5	L6	L7	L8	L9	L10	L1	L2	L3	L4	85.71%	
L7	L8	L9	L10	L1	L2	L3	L4	L5	L6	80.95%	
L9	L10	L1	L2	L3	L4	L5	L6	L7	L8	95.24%	
평균 인식률										90.47%	

<표 2> 중심점(x, y) 이동의 변화량

학 습										테스트	인식률
L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	100%	
L3	L4	L5	L6	L7	L8	L9	L10	L1	L2	80.95%	
L5	L6	L7	L8	L9	L10	L1	L2	L3	L4	85.71%	
L7	L8	L9	L10	L1	L2	L3	L4	L5	L6	100%	
L9	L10	L1	L2	L3	L4	L5	L6	L7	L8	90.48%	
평균 인식률										91.42%	

<표 3> 중심점(x, y, z) 이동의 변화량

학 습										테스트	인식률
L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	100%	
L3	L4	L5	L6	L7	L8	L9	L10	L1	L2	85.71%	
L5	L6	L7	L8	L9	L10	L1	L2	L3	L4	95.24%	
L7	L8	L9	L10	L1	L2	L3	L4	L5	L6	95.24%	
L9	L10	L1	L2	L3	L4	L5	L6	L7	L8	95.24%	
평균 인식률										94.28%	

이런 결과는 캘리브레이션을 통한 3차원 위치 정보를 이용해 x축, y축 특징값 이외 z축의 특징 정보(움직임 객체의 깊이 정보)를 추가해 더 좋은 결과를 얻을 수 있음을 알 수 있다. 또한 움직임 객체의 사각형 너비와 높이를 특징값으로 사용한 것 보다 움직임 객체의 사각형 중심점을 특징값으로 사용하

는 것이 좋은 결과값을 얻을 수 있었다.

#### 4. 결론

본 논문에서는 스테레오 영상으로부터 캘리브레이션 하여 3차원 정보를 추출하여 특징값으로 이용하고 이를 HTK를 통하여 HMM 학습을 하고 인식하는 방법을 제안하였다. 캘리브레이션을 통한 3차원 위치 정보를 이용해 x축, y축 특징값 이외 z축의 특징 정보(움직임 객체의 깊이 정보)를 추가해 동작에 따른 움직임 객체의 사각형 중심점 좌표(x, y, z)의 변화량을 특징값으로 사용함으로써 인식률이 향상됨을 보였다. 이것은 3차원 특징값을 이용하는 것이 견고하고 안정된 인식을 한다는 것을 보이고 있다.

향후 본 시스템은 사고 발생 직후의 즉각적인 대응이 필요한 헬스케어형 의료정보시스템의 일부용으로 응용되어 적용이 가능하리라 본다.

#### 참고문헌

- [1] Baihua Li, Holstein, H., and Qinggang Meng, "Articulated Point Pattern Matching in Optical Motion Capture System", Control, Automation, Robotics and Vision, ICARCV 2002. 7th International Conference, Vol. 1, pp. 298-303, 2002
- [2] Barnes, N.M., Rose, D.A.D., Garner, P.: Lifestyle Monitoring: Technology for Supported Independence. IEE Comp. and Control Eng. J. pp. 169-174, 1998
- [3] Bonner, S.: Assisted Interactive Dwelling House: Edinvar Housing Assoc. Smart Tech. Demonstrator and Evaluation Site In: Improving the Quality of Life for the European Citizen (TIDE), pp. 396 - 400, 1997
- [4] McKenna, S.J., Marquis-Faulkes, F., Gregor, P., Newell, A.F.: Scenario-based Drama as a Tool for Investigating User Requirements with Application to Home Monitoring for Elderly People. In Proc. of HCI, 2003
- [5] 한탁돈, 정철호, 김기섭, 이남규, 이충규, "정보화 사회와 착용형 컴퓨터의 전망", 한국멀티미디어학회지, 제6권, 제1호, pp.88-104, 2002년
- [6] W. Woo, N. Kim and Y. Iwadate, "Stereo Imaging Using a Camera with Stereoscopic Adapter", Proc, of IEEE - Systems, Man, and Cybernetics (SMC) 2000, vol.2, pp.1512-1517. Oct. 2000
- [7] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," Proc. of the Seventh IEEE Int'l Conf., vol. 1, pp.666-673, 1999