

3차원 영상복원 데이터를 이용한 HMM 기반 의도인식 시스템

HMM-based Intent Recognition System using 3D Image Reconstruction Data

고광은·박승민·김준엽·심귀보[†]

Kwang-Enu Ko, Seung-Min Park, Jun-Yeup Kim, and Kwee-Bo Sim[†]

중앙대학교 전자전기공학부

요 약

대뇌 상의 mirror neuron system은 시각 정보에 기반한 모방학습 기능을 담당한다. 관측자의 mirror neuron system 영역을 관찰할 때, 행위자가 수행하는 목적성 행위의 전체가 아닌, 부분적으로 가려지거나 보이지 않는 영역을 포함하는 경우에도 해당 영역의 뉴런이 발화되는 과정을 통해 전체 행동의 의도를 유추할 수 있다. 이러한 모방학습 기능을 3D 비전 기반 지능 시스템에 적용하는 것이 본 논문의 목표이다. 본 연구실에서 선행 연구된 스테레오 카메라를 기반으로 획득된 3차원 영상에 대한 복원을 수행한다. 이 때 3차원 입력영상은 부분적으로 가려진 영역을 포함하는 손동작의 순차적 연속영상이다. 복원 결과를 기반으로 가려진 영역을 내포한 행위에 대하여 LK optical flow, unscented Kalman filter를 이용한 특징검출을 수행하고 의도인식의 수행을 위해, Hidden Markov Model을 활용한다. 순차적 입력데이터에 대한 동적 추론 기능은 가려진 영역을 포함한 손동작 인식 수행에 있어 적합한 특성을 가진다. 본 논문에서 제안하는 의도 인식을 위해 선행 연구에서 복원 영상에서의 객체의 윤곽선 및 특징 검출을 시뮬레이션 하였으며, 검출 특징에 대한 시간적 연속 특징벡터를 생성하여 Hidden Markov Model에 적용함으로써, 의도 패턴에 따른 손동작 분류 시뮬레이션을 수행하였다. 사후 확률 값의 형태로 손 동작 분류 결과를 얻을 수 있었으며, 이를 통한 성능의 우수함을 입증하였다.

키워드 : 의도인식, 손동작 인식, 3차원 영상복원, HMM, Mirror Neuron System

Abstract

The mirror neuron system in the cerebrum, which are handled by visual information-based imitative learning. When we observe the observer's range of mirror neuron system, we can assume intention of performance through progress of neural activation as specific range, in include of partially hidden range. It is goal of our paper that imitative learning is applied to 3D vision-based intelligent system. We have experiment as stereo camera-based restoration about acquired 3D image our previous research Using Optical flow, unscented Kalman filter. At this point, 3D input image is sequential continuous image as including of partially hidden range. We used Hidden Markov Model to perform the intention recognition about performance as result of restoration-based hidden range. The dynamic inference function about sequential input data have compatible properties such as hand gesture recognition include of hidden range. In this paper, for proposed intention recognition, we already had a simulation about object outline and feature extraction in the previous research, we generated temporal continuous feature vector about feature extraction and when we apply to Hidden Markov Model, make a result of simulation about hand gesture classification according to intention pattern. We got the result of hand gesture classification as value of posterior probability, and proved the accuracy outstandingness through the result.

Key Words : Intent recognition, hand gesture recognition, 3D image reconstruction, HMM, mirror neuron system

1. 서 론

Human Computer Interaction (HCI) 연구 분야의 궁극

접수일자: 2011년 11월 19일

심사(수정)일자: 2012년 4월 9일

게재확정일자 : 2012년 4월 9일

† 교신 저자

본 논문은 본 학회 2011년도 추계학술대회에서 선정된 우수논문입니다.

본 논문은 한국연구재단 중견연구지원사업(No. 2011-0029861)에서 지원하여 연구하였습니다. 연구비 지원에 감사드립니다.

적인 목표는 인간과 컴퓨터의 보다 자연스러운 수단을 통한 상호작용에 있다. 이는 주어진 환경에서 사용자의 의도에 대한 인식과정을 위해 다양한 인터페이스 시스템이 개발 연구 수행 중에 있으며, 가장 널리 활용되고 있는 대표적인 사례가 비전 시스템에 대한 연구이다. 최근, 스마트 폰, 태블릿 PC 등과 같은 스마트 인터페이스의 발전은 비전 시스템의 다양한 활용 가능성을 보여주고 있다. 시각적 데이터를 기반으로 처리하는 비전 기반 어플리케이션 등의 형태로, 현재 게임, 소셜 네트워크 등의 엔터테인먼트 관련분야에서 큰 영향을 발휘하고 있다. 또한, 기존의 카메라 모듈에 국한되지 않고 보다 저렴하면서 고성능의 비전 시스템 하드웨어들이 개발되고 있는데, 대표적인 예가 마이크로소프트의 KINECT 시스템이다. KINECT는 보다 저렴한 비용의

로 3차원 스테레오 카메라 못지않은 영상정보획득이 가능하며, 행동인식, 모션 캡처 기능의 구현을 통해 게임, 가상 현실 등 다양한 엔터테인먼트 플랫폼과 연동되어 보다 자연스러운 서비스를 사용자에게 제공할 수 있다 [1].

그 밖에 관심이 집중되고 있는 Brain-Computer Interface (BCI)나 웨어러블 시스템 등의 고가의 정밀센서 시스템이 요구되는 HCI 연구도 있지만, 비전 시스템에 기반한 연구는 영상처리, 기계학습, 지능 시스템 알고리즘의 조합을 통해 보다 자연스럽게 효율적인 방법으로 궁극적인 상호작용의 일부를 구현한다는 장점이 있다. 이와 같은 비전 시스템을 기반한 HCI연구도 감정인식, 행위인식 등 다양한 분류가 있지만, 본 논문에서는 목적을 가지는 임의의 손동작에 대한 인식을 기반으로 의도를 추론하기 위한 핵심 기술들에 대하여 다루고자 한다. 이와 관련하여 비전만을 활용하기보다 다양한 분야의 기술의 집목을 통해 보다 효과적인 의도인지 수행이 가능하도록 하는 연구를 수행하고자 하였다. 이를 위해 초점을 맞춘 연구는 시각적 데이터에 대한 인간의 내재적 반응에 관련된 BCI 분야의 모방학습 기능과 관련된 Mirror Neuron System (MNS) 연구이다.

뇌 과학 분야에서는 대뇌 피질 상 MNS 영역에 대한 다양한 연구가 수행되었다. 이는 BCI 영역에서 뿐만 아니라 심리학, 인지과학 등의 다 학제 간 분야에서 적용되어 왔다. 일반적인 MNS의 정의는 행동유발성(affordance)을 내포하는 행위에 대하여 관측자(observer)와 수행자(executer)를 정의하고 실제 해당 행위를 수행하지 않지만 행위를 관찰하는 과정에서 MNS 영역이 행위를 수행하는 것과 동일한 형태의 발화가 발생하는 것을 의미한다 [2]. 이러한 특성은 행동에 대한 모방학습을 통해 의도를 인지하는 고유의 뇌 기능에 대한 설명을 가능하게 하며, 많은 연구자들이 MNS의 모방학습 기능을 직접적으로 활용하거나, 모델링하는 응용 연구를 수행한 바 있다. 한 예로, mirror neuron에 대한 신경생리학적 선행 연구에서는 원숭이를 대상으로 객체의 가림 여부가 의도인지 과정에 미치는 영향을 분석한 결과, mirror neuron에서 객체 간의 행동을 통해 의도를 인지하는 과정은 입력으로 들어오는 영상정보와 더불어 기존에 인지를 통해 학습한 경험을 토대로 수행됨을 보였다 [2].

이러한 의도 인지 과정을 계산 모델링한 시스템을 구축하기 위해서는 영상입력에 대한 온라인 측정 및 분석과 더불어, 시간에 따라 연속된 상태 별 분류를 수행할 수 있는 방법의 적용이 필요하다. 본 논문에서는 3차원 영상 데이터 기반 재구축 및 움직임 특징 측정을 위하여 선행 연구 [3][11]에서 제안된 방법을 활용하였으며, 시간에 따른 행위 상태 정보를 모델링하기 위한 hidden Markov model (HMM)을 적용한 시스템을 구축하였다. 본 논문의 구성은 다음과 같다. 2장에서는 제안하는 3차원 영상복원과 HMM에 기반한 의도인식 시스템에서 행위 관측 시 활용되는 특징 검출관련 연구에 대한 설명을 하고, 3장에서 의도인식을 위한 손 동작 인식 인식 알고리즘 세부 내용 및 주요 아이디어를 다룬다. 4장에서는 제안 방법에 기반한 의도인식 시뮬레이션을 수행, 그 결과를 분석하고, 마지막 5장에서 최종 결론 및 향후 보완 연구 과제에 대하여 논한다.

2. 행위 관측 특징 검출

2.1 3차원 영상복원 기반 입력영상 전처리

앞서 설명한 바와 같이 MNS은 시각적인 입력에 기반하여 모방학습 기능을 수행한다. 이를 계산 모델링하려면 우선적인 과제가 동적으로 변화하는 행위 입력에 대응하는 특징 검출을 수행해야한다. 본 논문에서 다루고자 하는 행위 관련 내용은 크게 Pose와 Motion으로 구분되어진다. Pose는 정적인 형태의 동작으로 주로 전신 기반의 각 일부의 움직임 변화에 근거하여 추정되며, Motion은 연속된 순차영상으로 표현되는 동적인 형태의 동작으로 안면 근육, 손가락 움직임, 피부 표면의 변화 등과 같이 세밀한 부위 별 측정과 분석까지 가능하다는 장점이 있다 [4]. Pose 관련된 기존의 비전 시스템 연구는 정적으로 주어진 영상에서 표현되는 Pose에 대한 해석을 목표로 하기 때문에 본 논문에서 추구하는 시간에 따라 연속적으로 변화하는 손 동작 변화에 대한 인식 수행에는 부적합하다. 따라서 Motion에 기반한 손 동작 인식의 수행을 위해 그에 해당하는 전처리 및 특징을 검출을 선행할 필요가 있다. 입력영상의 전처리를 위해 본 논문에서 집중하는 바는 실제적인 활용 환경을 고려하였을 때, 3차원 상의 영상입력 데이터와 행위 주체가 일부 가려지는 경우에 대한 복원이다. 본 연구실은 선행된 연구 [3]에서 다룬 Camera Distortion Correction과 Image Level Normalization으로 구성되는 3차원 영상복원 기법을 활용하여 그림 1과 같은 결과를 얻었다.

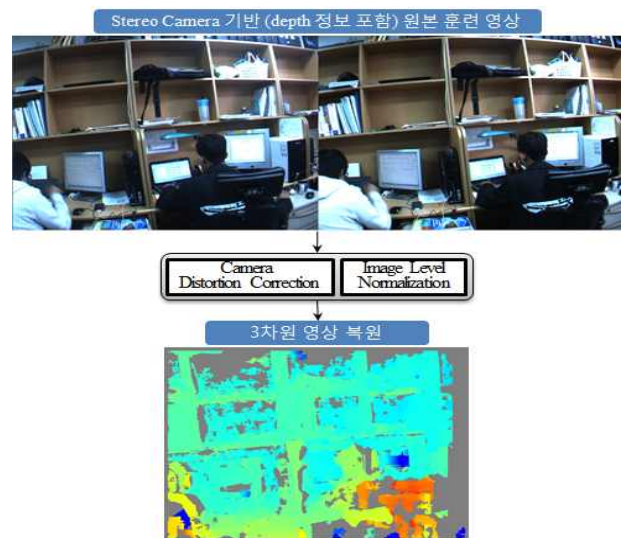


그림 1. 스테레오 카메라 기반 3차원 영상복원

Fig. 1. Stereo Camera-based 3D image reconstruction

2.2 동적영상입력 기반 행위특징 검출

2.2.1 Unscented Kalman Filter 기반 윤곽선 특징 추적

위의 설명된 추적하고자 하는 객체의 움직임 벡터 특징과 더불어 객체의 정확한 위치 추정을 통해 행위 인식 단계에서의 계산량을 줄이고 인식 정확도 향상을 위하여 unscented Kalman Filter (UKF)에 기반한 객체 윤곽선 추적 방법을 활용하였다. 동적으로 변화하여 비선형적 특성을 보이는 순차적 입력 영상에서의 객체 추정을 위해 기존의 여러 연구에서 Kalman Filter를 활용한 바 있다 [5][6]. UKF를 기반으로 윤곽선 추적을 위해 아래의 식 (1)과 (2)를 각각 predictor와 updater로 활용한다.

$$\begin{aligned} \hat{\mathbf{x}}_i^{(t)} &= \mathbf{A}\hat{\mathbf{x}}_i^{(t-1)} + \mathbf{w}_i^{(t-1)} \\ \mathbf{z}_i^{(t-1)} &= \mathbf{P}_i^{(t-1)} + \mathbf{v}_i^{(t-1)} \end{aligned} \quad (1)$$

임의의 시점 t 에서의 프레임 상에서 $\mathbf{x}_i^{(t)}$ 로 표현되는 i 번째 윤곽선 세그먼트의 상태벡터, $\mathbf{z}_i^{(t)}$ 로 표현되는 측정 벡터, 그리고 $\mathbf{w}_i^{(t)}$, $\mathbf{v}_i^{(t)}$ 를 프로세스, 상태 노이즈로 정의한다. 이 때 추정되는 윤곽선 세그먼트의 위치는 $\mathbf{p}_i^{(t)}$ 로 정의한다.

$$\begin{aligned} \mathbf{S}_i^{(t-1)} &= \mathbf{E}_i^{(t-1)} + \mathbf{R} \\ \mathbf{K}_i^{(t-1)} &= \mathbf{A}\mathbf{E}_i^{(t-1)}\mathbf{S}_i^{(t-1)-1} \\ \mathbf{E}_i^{(t)} &= \mathbf{A}\mathbf{E}_i^{(t-1)}\mathbf{A}^T + \mathbf{Q} - \mathbf{A}\mathbf{E}_i^{(t-1)}\mathbf{S}_i^{(t-1)-1}\mathbf{E}_i^{(t-1)}\mathbf{A}^T \\ \hat{\mathbf{x}}_i^{(t)} &= \mathbf{A}\hat{\mathbf{x}}_i^{(t-1)} + \mathbf{K}_i^{(t-1)}(\mathbf{z}_i^{(t)} - \mathbf{A}\hat{\mathbf{x}}_i^{(t-1)}) \end{aligned} \quad (2)$$

$\mathbf{w}_i^{(t)}$ 와 $\mathbf{v}_i^{(t)}$ 에 대한 공분산행렬을 \mathbf{Q} , \mathbf{R} 이라하고 칼만이득 $\mathbf{K}_i^{(t)}$ 와 공분산의 prediction error를 $\mathbf{E}_i^{(t)}$ 로 정의한다.

이렇게 추정된 영상에서의 객체위치는 상태 공간상에서의 탐색 속도와 정확성 향상을 보장한다. UKF에 대하여 보다 상세한 설명은 [6]을 참조한다.

2.2.1 LK-Optical Flow 움직임 벡터

본 논문에서 사용되는 입력영상은 정지영상 기반의 데이터가 아닌 동적 변화하는 순차적 프레임의 연속집합이다. 이러한 입력을 통해 특징 검출을 하기 위해서는 시간에 따라 동적으로 변화하는 특징점 위치 변화에 대하여 파라미터화 된 정보를 생성할 필요가 있다. 본 연구실에서는 선형연구를 통해 동적으로 변화하는 영상의 움직임 벡터 검출에 Lucas-Kanade (LK) optical flow를 활용한 바 있다 [7]. 이를 위해 본 논문에서는 시간적으로 연속되어 다중 프레임으로 구성된 영상에서 추출하고자 하는 특징 점을 UKF를 기반으로 추정된 객체 윤곽선 영역의 픽셀들로 가정하였다. 시점 t 에서의 프레임을 \mathbf{I}_t 라고 한다면, optical flow는 \mathbf{I}_t 와 \mathbf{I}_{t+1} 에 존재하는 특징 점간의 움직임 벡터를 정의한 것으로 볼 수 있다 [8]. 인접 프레임 간의 임의의 픽셀 $\mathbf{p}_i = (x_i, y_i)$ 의 움직임은 \mathbf{p}_i 에 인접한 매우 작은 상수로 근사한다. \mathbf{I}_t 에서의 한 특징점을 $[x_i^t, y_i^t]^T$ 로 가정한다면, 연속프레임일 경우 \mathbf{I}_{t+1} 의 동일 위치 특징점은 $[x_i^t + \delta_x, y_i^t + \delta_y]^T$ 로 표현 가능하며, optical flow는 아래의 식 (3)을 최소화 하는 δ_x, δ_y 에 의한 파라미터로 표현할 수 있다.

$$\epsilon(\delta_x, \delta_y) = \sum_{x=x_i^t-w_x}^{x_i^t+w_x} \sum_{y=y_i^t-w_y}^{y_i^t+w_y} (I_1(x, y) - I_2(x+\delta_x, y+\delta_y)) \quad (3)$$

w_x, w_y 는 적용되는 mask window의 x, y 좌표 값이다.

위 설명을 토대로 본 논문에서 사용하고자 하는 LK optical flow를 식 $\mathbf{A}\mathbf{v} = \mathbf{b}$ 로 표현할 수 있다. 식의 각 parameter는 $\mathbf{A} = [\mathbf{I}_x(p_1), \mathbf{I}_y(p_1); \dots; \mathbf{I}_x(p_n), \mathbf{I}_y(p_n)]^T$, $\mathbf{v} = [V_x, V_y]^T$, $\mathbf{b} = [\mathbf{I}_x(p_1), \mathbf{I}_x(p_2), \dots, \mathbf{I}_x(p_n)]^T$ 를 의미한다. $\mathbf{I}_x(p_i), \mathbf{I}_y(p_i), \mathbf{I}_t(p_i)$ 는 영상프레임 \mathbf{I} 를 x, y, t 에 따라 편미분을 의미한다. 위의 파라미터 중 optical flow \mathbf{v}_i 는 다음의 (4)와 같다.

$$\mathbf{v}_i = \begin{bmatrix} \sum_{i=1}^n \mathbf{I}_x(p_i)\mathbf{I}_x(p_i) & \sum_{i=1}^n \mathbf{I}_x(p_i)\mathbf{I}_y(p_i) \\ \sum_{i=1}^n \mathbf{I}_x(p_i)\mathbf{I}_y(p_i) & \sum_{i=1}^n \mathbf{I}_y(p_i)\mathbf{I}_y(p_i) \end{bmatrix}^{-1} \begin{bmatrix} -\sum_{i=1}^n \mathbf{I}_x(p_i)\mathbf{I}_t(p_i) \\ -\sum_{i=1}^n \mathbf{I}_y(p_i)\mathbf{I}_t(p_i) \end{bmatrix} \quad (4)$$

지금까지 설명한 UKF와 LK optical flow 기법에 기반한 동적영상 입력기반 특징 검출 결과는 그림 2와 같다.



그림 2. 동적영상 입력기반 행위특징 검출

Fig. 2. Dynamic Image Input-based Behavior Feature Extraction

3. 동적영상 입력의 행위특징을 이용한 의도인식 시스템

3.1 목적성 행위 관측 기반 모방학습 모델

동적영상 기반 객체의 행위 특징을 검출하기 위하여 스테레오 카메라를 통해 획득된 영상에 대한 3차원 재구축을 기반으로 하는 전처리 과정과 UKF를 이용한 객체 윤곽선 특징 추정 및 LK optical flow를 기반으로 하는 움직임 벡터 형태의 동적 행위 특징 검출을 수행하였다. 본 논문에서는 이러한 동적 행위 특징을 기반으로 목적성 행위에 대한 의도 추론 및 인지를 목표로 하는 모방학습 기능 계산 모델을 제안한다. 제안하는 계산 모델의 구조는 그림 3과 같다.

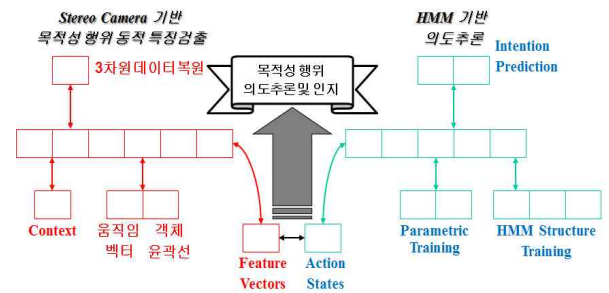


그림 3. 목적성 행위 관측 기반 모방 학습 모델

Fig. 3. Imitative Learning Model based on Object-related Behavior Measurement

3.2 Hidden Markov Model 기반 목적성 손 동작 분류 및 의도추론

앞서 설명한 목적성 행위 관측 기반 모방학습 계산모델에서 의도추론을 위해 본 논문에서는 Hidden Markov Model (HMM)을 활용하고자 한다. HMM의 확률 추정에 따른 상태 분류기능은 실제 환경에서 높은 빈도로 발생 가능한 부분적으로 가려진 경우를 고려하는 목적성 행위 인식에 활용 가능하다 [9]. 또한, 2장의 설명한 기술을 통해 검출된 동적 행위 특징 역시 부분적으로 가려진 경우를 고려하는 과정에서 유용하다. 선행 연구 [3]에서는 이와 관련하여 객체행위 추정에 대한 히스토리 맵을 구축하는 연구를 수행한 바 있다. 본 논문에서는 히스토리에 저장된 동적 영상 특징 데이터를 관측치(observation)로 정의하고, 마르코프 가정을 기반으로 구조, 파라미터 학습을 거쳐 목적성 행위 상태(hidden state) 추정을 위한 HMM을 구축한다.

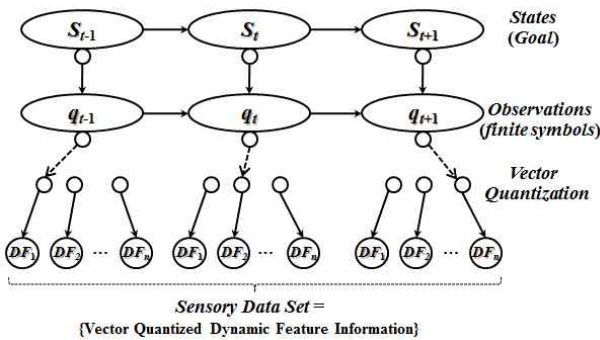


그림 4. 동적영상 입력특징 기반 HMM 구조

Fig. 4. Struction of HMM based on Dynamic Image Input Features

HMM의 입력을 단일 경로 제스처 $g(x, y, z, t)$ 로 정의하고 이를 벡터 양자화 한 값을 q_k 라고 한다. k 는 벡터 양자화 순서를 뜻한다. 분류하고자 하는 손 행위의 해당의도 상태 벡터를 $S = \{s_i\}_{i=1,2,\dots,n}$ 로 정의한다면, HMM의 구조를 그림 4와 같이 구축하였고, HMM을 구성하는 파라미터 $\lambda = (A, B, \pi)$ 는 다음과 같이 정의하였다.

a) A : state transition probability matrix
 $= \{a_{ij} = P(s_j | s_i)\}_{i, j=1,2,\dots,n}$

b) B : observation-state probability matrix
 $= \{b(q_k) = P(q_k | s_j)\}_{j=1,2,\dots,n, k=1,2,\dots,m}$

c) π : initial state, observation probabilities
 $= P(s_i)_{i=1,2,\dots,n}$

n 개의 은닉상태(의도), m 개의 관측치(특징)와 λ 로 모델링된 HMM을 논문에서 수행하고자 하는 목적성 행위 패턴 분류에 적용하려면 Viterbi 알고리즘에 기반한 decoding을 활용한다. 이 때, HMM 기반 행위 상태 확률 추정을 수행하는 식은 다음과 같다.

$$g^* = \arg \max_{all \lambda} P(\lambda | Q), Q = \{q_k\}_{k=1,2,\dots,m} \quad (5)$$

g^* 는 주어진 관측치 집합 Q 에 대하여 학습된 HMM λ 의 확률을 계산하여 최대치를 도출하는 관측치 집합 q_k^* 를 구하

고 이를 해당하는 최적 단일 경로 제스처를 의미한다. 본 논문에서는 그림 4에서의 과정을 통해 동적영상 입력의 목적성 행위 기반 의도 추론 방법에 대하여 제안하였다.

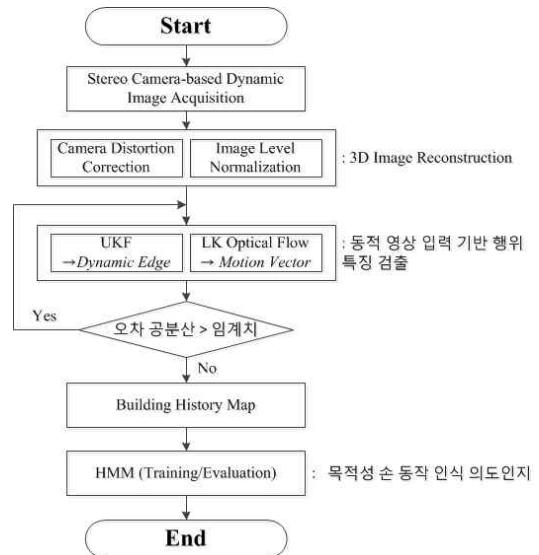


그림 5. 동적영상 입력 기반 의도인식 흐름도

Fig. 5. Flowchart of Dynamic Image Input-based Intent Recognition

4. 시뮬레이션 및 결과분석

본 논문에서 제안하는 3차원 영상복원 데이터를 이용한 목적성 행위 기반 의도인식 시스템의 시뮬레이션을 수행하기 위하여, 2, 3장에서 다룬 전처리, 특징검출, 패턴분류 알고리즘을 각각 구현하였다.

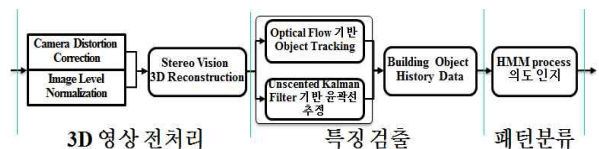


그림 6. 의도인식 시뮬레이션 블록다이어그램

Fig. 6. Block-diagram for Simulation of Intent Recognition

시뮬레이션 시스템의 입력으로 적용되는 3차원 영상데이터는 본 연구실의 보유 장비 중 Point Grey Research Corp.의 BumbleBee 2 스테레오 카메라를 이용하여 수집하였으며, 영상처리 관련 알고리즘의 실질 구현을 위하여 OpenCV 2.3.1 라이브러리를 MSVC 2010 MFC의 Dlg 기반 윈도우 프로그래밍을 수행하였다.

그림 7에 도시된 것처럼 시뮬레이션의 대상이 되는 객체는 4가지 상태의 손동작과 동일한 상태에 부분적으로 가려진 손동작으로 구성된 총 8개의 순차영상 데이터이다. Seq. 1에서 Seq. 4까지는 수행하는 손동작에 대한 부분적 가려짐이 존재하지 않지만, Seq. 5부터 Seq. 8까지는 영상 중앙에 위치한 방해물에 의한 부분적 가려짐을 가정하였다. 시뮬레이션에 쓰일 목적성 손동작에 대한 특징 검출 및 의도 패턴

분류에 대한 학습을 위하여 본 논문에서는 그림 8과 같은 Cambridge Hand Data Set의 클래스 1에서 5까지의 영상을 활용하였다 [10].

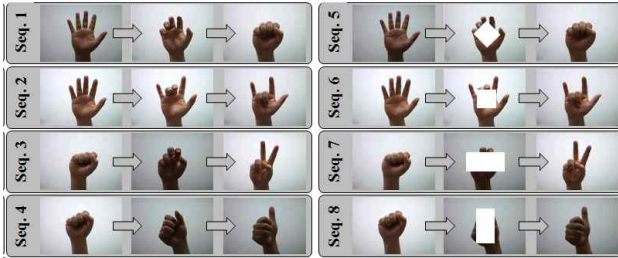


그림 7. 의도인식 시뮬레이션을 위한 손동작 입력 데이터
Fig. 7. Hand Gesture-based Input Data for Intent Recognition Simulation

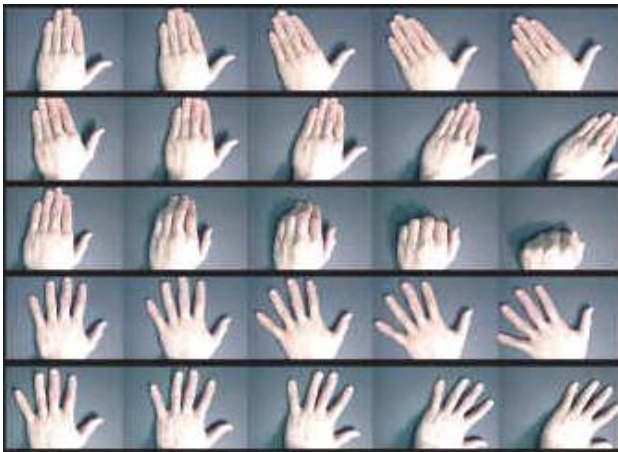


그림 8. 학습에 활용된 손동작 데이터베이스 예시
Fig. 8. Example of Hand Gesture Database for Training

위의 실험 조건과 제안하는 3차원 영상복원 기반 특징 검출 및 HMM를 이용한 손동작 의도 패턴분류를 수행한 결과는 표 1과 같다. 그림 7에 해당하는 각 순차영상 별로 HMM의 결과인 의도패턴 별 우도(likelihood)에 대한 사후 확률 분포를 계산한 결과를 도시하였다. 이 결과는 본 논문에서 제안하는 방법을 이용하여 의도 인지를 수행할 수 있는 HCI 시스템의 개발 가능성을 증명한다.

표 1. 손동작 순차영상 기반 의도패턴 별 사후확률분포
Table 1. Posterior Probability Distribution for Each Hand Gesture Image Sequence

| Image Sequence Index | Probability Distribution (Likelihood) of Intention States (%) | | | |
|----------------------|---|----|----|----|
| | S1 | S2 | S3 | S4 |
| Seq. 1 | 85 | 7 | 2 | 6 |
| Seq. 2 | 9 | 86 | 2 | 3 |
| Seq. 3 | 1 | 8 | 89 | 2 |
| Seq. 4 | 7 | 4 | 6 | 83 |
| Seq. 5 | 72 | 13 | 6 | 9 |
| Seq. 6 | 11 | 76 | 4 | 9 |

| | | | | |
|--------|---|----|----|----|
| Seq. 7 | 6 | 15 | 68 | 11 |
| Seq. 8 | 7 | 5 | 7 | 81 |

5. 결론 및 향후 고찰

본 논문은 비전시스템을 기반으로 사용자의 목적성 행위에 내포된 의도를 인지하는 HCI 시스템의 핵심 기술에 대하여 논하였다. 실제 환경에서 동작인식 등의 비전 시스템 기반 HCI를 구현할 경우, 수집되는 입력 영상에 대하여, 외란이나 방해물로 인한 부분적인 손실이 빈번하게 발생한다. 이를 극복하고자 3단계로 구성되는 목적성 행위인식 기반 의도인식 시스템을 제안하였다. 우선 스테레오 카메라를 통해 수집된 동적 순차영상에 대한 3차원 영상처리 및 복원에 따른 데이터 전처리를 수행함으로써 영상 depth 정보를 토대로 부분적 입력 영상 손실을 줄이고자 하였다. 이렇게 전처리된 영상 데이터를 기반으로 최적의 목적성 행위 특징 데이터를 검출하기 위하여 Unscented Kalman Filter를 통한 동적 객체의 윤곽선을 추적하고 Lucas-Kanade Optical Flow를 검출된 윤곽선 영역에 적용함으로써 계산량의 축소 및 특징 검출 과정에서의 효율성 증대를 도모하였다. 검출된 특징은 Hidden Markov Model로 구성된 패턴분류기로 적용되는데, 기존의 영상처리 기반 HCI에서 주로 활용되는 3차원 영상복원 결과를 얻기 위하여 히스토리 맵을 구축하고 학습과정에서 저장된 동적영상 입력에 대응하는 특징 데이터를 히스토리 맵에 순차적으로 저장하였다. 저장된 특징 데이터를 기반으로 목적성 행위를 입력으로 대응 의도 패턴에 대한 사후 확률 분포를 계산하는 HMM 기반의 패턴분류를 수행함으로써 최종적으로 구하고자 하는 동적 입력 영상에 대한 의도인식을 수행하였다. 제안 방법에 대한 검증을 위하여 4개의 상태로 구성된 손동작 인식 시뮬레이션을 수행하였으며, 제안 방법의 우수성을 도시하였다.

향후에는 목적성 행위에 대한 다양성 확보를 위하여 손 동작 이외의 제스처 인식을 기반으로 하는 연구를 수행할 예정이며, 최근 3차원 영상처리 분야에서 많은 주목을 받고 있는 KINECT 시스템을 기반으로 제안 방법을 구현 후 스테레오 카메라를 기반으로 하는 시스템과의 성능 비교를 통해 고효율 고성능의 시스템 개발을 진행하고자 한다.

참고 문헌

- [1] L. Xia, C.-C. Chen, and J. K. Aggarwal, "Human Detection Using Depth Information by Kinect", *International Workshop on Human Activity Understanding from 3D Data in conjunction with CVPR (HAU3D)*, Colorado Springs, CO, 2011.
- [2] G. Buccino, F. Binkofski, and L. Riggio, "The mirror neuron system and action recognition," *Brain and Language*, vol. 89, no. 2, pp. 370-376, 2004.
- [3] 박준형, 박승민, 심귀보, "3차원 재구성 과 추정된 움티컬 플로우 기반 가려진 객체 움직임 추적방법," *한국지능시스템학회 논문지*, vol. 21, no. 5, pp. 537-542, 2011.

[4] R. Poppe, "Vision-based human motion analysis: An overview," *Computer Vision and Image Understanding*, vol. 108, pp. 4-18, 2007.

[5] Y. Chen, T. Huang, "Parametric contour tracking using unscented Kalman filter," in *Proc. IEEE Int'l Conf on Image Processing 2002*, 2002.

[6] M. Musrshed, M. A. A. Dewan, O. Chae, "Moving object tracking-a parametric edge tracking approach," in *Proc. Int'l Conf on Computer and Information Technology 2009*, 2009.

[7] 고광은, 박승민, 박준형, 심귀보, "ASM-LK Optical Flow 기반 최적 얼굴정서 특징분석 기법", *한국지능시스템학회 논문지*, vol.21, no.4, pp. 512-517, 2011

[8] J. K. Kearney, W.B. Thompson, "Optical flow estimation: an error analysis of gradient-based methods with local optimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 2, 1987.

[9] C. H. Han, C. S. Oh, B. W. Choi, "Recognition of Fighting Motion using a 3D-Chain Code and HMM," *Journal of Institute of Control, Robotics, and Systems*, vol. 16, no. 8, pp. 756-760, 2010.

[10] T-K. Kim, S-F. Wong and R. Cipolla, Tensor Canonical Correlation Analysis for Action Classification, *In Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, MN, 2007.

[11] 고광은, 박준형, 심귀보, "3차원 영상복원 데이터를 이용한 HMM 기반 의도인식 시스템," *한국지능시스템학회 2011년도 추계학술대회*, vol. 21, no. 2, pp. 15-16, 2011.



박승민(Seung-Min Park)

2010년 : 중앙대학교 전자전기공학부 공학사
 2010년 ~ 현재 : 중앙대학교 대학원
 전자전기공학부
 석박사통합과정

관심분야 : Brain-Computer Interface, Intention Recognition
 Soft Computing 등

Phone : 02-820-5319

E-mail : sminpark@cau.ac.kr



김준엽(Jun-Yeup Kim)

2012년 : 중앙대학교 전자전기공학부 공학사
 2012년 ~ 현재 : 중앙대학교 대학원
 전자전기공학부 석사과정

관심분야 : Brain-Computer Interface System, Particle
 Swarm Optimization, Neuro-Robotics 등

Phone : 02-820-5319

E-mail : jy915@cau.ac.kr



심귀보(Kwee-Bo Sim)

1990년 : The University of Tokyo
 전자공학과 공학박사
 1991년 ~ 현재 : 중앙대학교 전자전기공학부
 교수
 2006년 ~ 2007년 : 한국지능시스템학회 회장

관심분야 : 인공지능, 지능로봇, 지능시스템, 다개체시스템,
 학습 및 적응알고리즘, 소프트 컴퓨팅 (신경망,
 퍼지, 진화연산), 인공면역시스템, 침입탐지시스
 템, 진화하드웨어, 인공두뇌, 지능형 홈 및 홈네
 트워킹, 유비쿼터스 컴퓨팅 등

Phone : 02-820-5319

E-mail : kbsim@cau.ac.kr

Homepage URL : <http://alife.cau.ac.kr>

저 자 소 개



고광은(Kwang-Eun Ko)

2007년 : 중앙대학교 전자전기공학부 공학사
 2007년 ~ 현재 : 중앙대학교 대학원
 전자전기공학부
 석박사통합과정

관심분야 : Multi-Agent Robotic Systems (MARS),
 Machine Learning, Context Awareness,
 Emotion Recognition Systems 등

Phone : 02-820-5319

E-mail : kkeun@cau.ac.kr