키넥트 센서 기반 격투액션 게임을 위한 제스처 인식에 관한 연구

김종민 · 김은영*

한국폴리텍대학 제주캠퍼스 융합디자인학과

The Study on Gesture Recognition for Fighting Games based on Kinect Sensor

Jong-Min Kim · Eun-Young Kim*

Dept. of Convergence Design, Jeju Campus, Korea Polytechnics

E-mail : kjm@kopo.ac.kr / key@kopo.ac.kr

요 약

본 논문에서는 키넥트 센서를 이용한 제스처 인식 방법을 개발하고, 이를 이용한 격투액션 제어 인터페이스를 제안한다. 제스처의 패턴 특징을 추출하기 위해서는 단순한 절대 위치 정보를 이용하는 것이 아닌 , 어깨를 중심으로 한 신체 비율 특성을 고려하여 특징을 추출하는 방법을 이용한다. 하지만 동일한 제스처를 수행하더라도 키넥트 센서에 포착되는 각 관절의 위치 좌표값들은 팔의 길이와 방향에 따라 달라질 수 있다는 문제점이 있다. 그래서 논문에서는 제스처를 모델링하고 분석하기 위해 주성분 분석법을 사용하는 방법을 기술한다. 이 방법을 사용함으로서 데이터가 가지는 에러의 영향을 줄일 수 있게 되고, 차원축약의 효과를 얻을 수 있게 된다. 또한 동작 인식 시스템의 동작 제약을 줄이기 위한 방법으로 수정된 매칭 알고리즘을 제안한다.

ABSTRACT

This study developed a gesture recognition method using Kinect sensor and proposed a fighting action control interface. To extract the pattern features of a gesture, it used a method of extracting them in consideration of a body rate based on the shoulders, rather than of absolute positions. Although the same gesture is made, the positional coordinates of each joint caught by Kinect sensor can be different depending on a length and direction of the arm. Therefore, this study applied principal component analysis in order for gesture modeling and analysis. The method helps to reduce the effects of data errors and bring about dimensional contraction effect. In addition, this study proposed a modified matching algorithm to reduce motion restrictions of gesture recognition system.

키워드

제스처인식(Gesture Recognition), 주성분분석(PCA), 키넥트(Kinect)

Ⅰ. 서 론

컴퓨터 기술의 발달과 함께 정보 시스템이 복잡하게 되면서 인간과 정보 시스템 사이에 자연스럽게 정보를 교환할 수 있는 지적 인터페이스에 관한 관심이 날로 커지고 있다. 인간은 일상생활에서 제스처, 표정과 같은 비언어적인 수단을 이용하여 수많은 정보를 전달한다. 따라서 자연스럽고 지적인 인터페이스를 구출하기 위해서는 제스처와 같은 비언어적인 통신 수단에 대한 연구가 매우 중요하다. 이러한 분위기 속에서 보다 인간이 사용하기 쉬운 새로운 사용자 인터페이스 제작 기술로서 제스처 인식기술이 계속적으로 발전하고 있다[1-3]. 컴퓨터를 사용하여 동작을 인식한다는 것은 인체 각 부위가 시간의 흐름에 따라 어떤 모습으로 변화하는 가를 자동으로 분석하고 그 변화를 추상적인 의미로 해석하는 것을 의미한다. 즉 동영상으로

* corresponding author

부터 신체 영역을 추출한 다음 특정 부분을 식별하고 각 부분들이 하나의 의미를 갖기 위해 어떤 변화를 거치는지를 알아내는 것이다. 그러나 인체는 고자유도를 지닌 매우 복잡한 3차원 관절 물체로 2차원의 동영상으로부터 인체부위를 안정적으로 분리해 내고 그 내용을 인식한다는 것은 매우 어려운 일이다[4]. 본 논문에서는 키넥트 센서 환경에서 제스처를 인식하고 이를 격투 액션 게임에 적용하였다. 먼저, 키넥트 센서로부터 획득한 관절 정보를 이용하여 제스처 영역의 움직임특징을 추출하고 인간의 동작을 수학적으로 모델링하고, 이를 주성분 분석법 (Principal Component Analysis)에 의해 고유 공간에 표현한 뒤, 새로운 입력에 대해 모델과 비교(matching)하는 방법으로써, 동작을 효과적으로 모델링하고 분석하는 방법에 대하여 기술한다.

II. 제스처 인식 시스템

본 논문에서 제안하는 제스처 인식 시스템은 키넥트 센서로 부터 미리 정해놓은 동적 제스처(dynamic gesture)중 하나를 취할 때, 키넥트 센서의 깊이 영상(depth image)들로부터 사용자의 관절들의 위치 좌표를 추출한다. 이렇게 추출된 특징점들을 VR Interface로 정의된 움직임에 대해 모델링하고, 주성분 분석법에 의해 고유 공간에 투영시킨다. 고유 공간의 특성상 시각적으로 비슷한 동작의 경우 고유공간에서 비슷한 분포를 갖기 때문에 이를 이용하여 실시간으로 입력되는 입력에 대해 고유공간 내에서 모델 동작과 비교(matching)함으로써 동작을 분석할 수 있게 된다. 이렇게 분석되어진 동작의 정보는 VR Inerface로 변환되어 사용할 수 있다. 그림 1은 제스처 인식 시스템의 전체 구성도를 나타내었다.

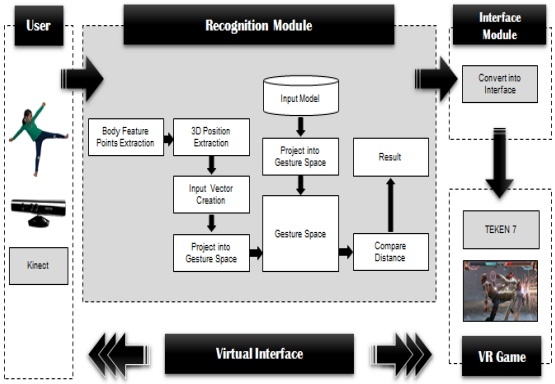

그림 1. 제스처 인식 시스템 구성도

III. 제스처 인식방법

키넥트 센서로부터 획득된 데이터 입력으로부터 제스처를 인식하기 위한 패턴 특징을 추출하기 위해서는, 제스처의 특성을 고려하여 특징을 추출하고 이를 제스처 모델에 적절한 입력으로 표현해주는 것이 필요하다. 본 논문에서는 Kinect SDK의 NUI Skeleton API[5] 이용하여 제스처를 입력하는 사용자에 대한 20개의 관절 정보를 획득한다. 해당 동작에서 그림 2에 표시된 것과 같이 (머리-늑골), (늑골-골반), (오른쪽 어깨-오른쪽 팔꿈치), (오른쪽 팔꿈치-오른쪽 손), (골반-오른쪽 무릎), (오른쪽 무릎-오른쪽 발), (왼쪽 어깨-왼쪽 팔꿈치), (왼쪽 팔꿈치-왼쪽 손), (골반-왼쪽 무릎), (왼쪽 무릎-왼쪽 발)의 10개 관절에 대해 각각 X-Y축, X-Z축의 각도를 계산하여 총 20개의 각도 데이터를 신체 특징점으로 사용한다.

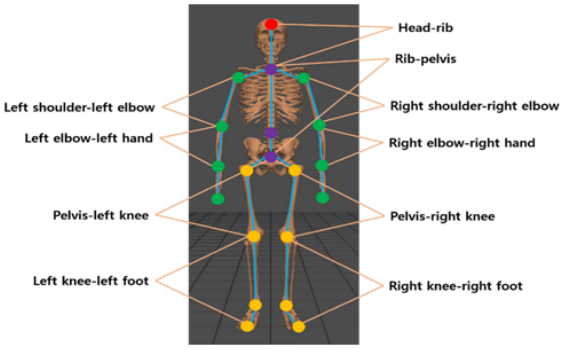

그림 2. 제스처 특징 추출 정보

IV. 제스처의 정의 및 분석

4.1 제스처의 정의

본 논문에서 정의하는 동작은 3차원 액션 게임에서 사용되는 동작들로 표 1과 같이 7개의 동작으로 정의하였다. 정의한 동작들을 구별이 애매하지 않게 수학적으로 모델링 해야하기 때문에 확실한 특징 벡터의 선정이 중요하다. 본 논문에서는 그림 3에서와 같이 머리를 기준으로 하는 양손, 양발의 3차원 상대 위치와 머리와 양발의 이동 속도 벡터를 특징으로 사용한다. 이러한 특징은 식(1)와 같이 총 21차원의 입력 벡터로 생성되며, 이를 분석함으로써 동작을 구별해 낼 수 있다.

$$
\begin{aligned}
M_{raw} &= \{M_{HD}, M_{LH}, M_{RH}, M_{LF}, M_{RF}\} \\
F &= \{F_1, F_2\} \\
&= \{R_{LH}, R_{RH}, R_{LF}, R_{RF}, V_{HD}, V_{LF}, V_{RF}\} \\
F_1 &= \{M_{HD} - M_{LH}, M_{HD} - M_{RH}, M_{HD} - M_{LF}, M_{HD} - M_{RF}\} \\
F_2 &= \{d(M_{HD})/dt, d(M_{LF})/dt, d(M_{RF})/dt\}
\end{aligned}
\qquad (1)
$$

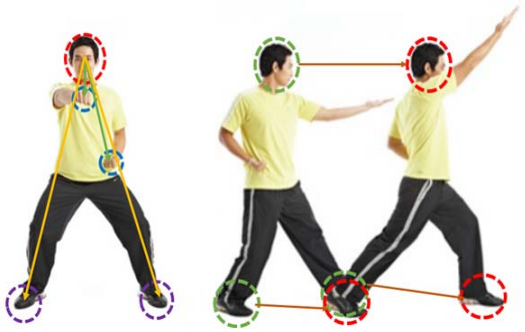

그림 3. 특징 벡터 (a) 상대위치 (b) 속도 벡터

4.2 동작의 모델링

앞 서 정의한 특징벡터들이 매 프레임 마다 얻어지게 되지만, 입력 데이터가 차원이 높고, 일반적인 특징을 갖고 있지 않은 경우도 있기 때문에 이를 효과적으로 모델링 하는 일은 쉽지 않다. 그렇기 때문에 본 논문에서는 주성분 분석법(Principal Component Analysis)을 사용하여 고유공간에서 동작을 모델링 하였다[6].

앞 절에서 구한 특징 벡터를 x_i라 하고, 식(2)와 같이 표현 할 수 있다. 이 벡터의 고유공간을 계산하기 위해서는 먼저 모든 특징 벡터의 평균 벡터를 구하여 각 특징 벡터의 차를 구해야한다. 평균 벡터 c와 새로운 특징 집합 X는 식(3-6)와 같다.

$$x_i = [P_{LH}, P_{RH}, P_{LF}, P_{RF}, V_H, V_{RF}, V_{LF}]^T \quad (2)$$

$$c = (1/N)\sum_{i=1}^{N} x_i \quad (3)$$

$$X \triangleq [x_1 - c, x_2 - c, x_3 - c, \ldots\ldots x_N - c]^T \quad (4)$$

$$Q \triangleq XX^T \quad (5)$$

$$\lambda_i e_i = Q e_i \quad (6)$$

4.3 수정된 매칭 알고리즘

미리 정의된 동작에 대하여 고유공간에서 모델링하고, 새로 입력되는 동작을 고유공간 내에서 가장 가까운 동작으로 결정함으로써 입력 동작을 분석할 수 있다. 하지만 보다 자연스러운 Interface 구현을 위해서는 동작의 수치적인 데이터보다는 동작의 의미에 중점을 두어야 한다. 앞서 정의한 “오른쪽 킥” 동작의 경우를 예로 들면, 양팔의 위치는 이 동작을 구분하는데 필요하지 않다. 다시 말해서 팔의 위치는 동작을 구분하는데 있어서 불필요하기 때문에, 동작 인식에서 제외시켜도 상관없다.

이러한 이유에서 본 논문에서는 필요 없는 특징의 경우 이를 배제하고 동작 인식에 적용함으로써 동작의 제약을 완화하였다. 각 동작과 관계가 있는 특징 값은 표 1과 같다. 표 1의 내용을 기준으로 입력 값에 대하여 각 동작을 가정하고 관계없는 특징 값은 모델의 평균값으로 대체한다. 이렇게 함으로써 불필요한 특징 값을 배제할 수 있고, 동작의 제약을 없앨 수 있다. 그림 4는 수정된 매칭 알고리즘을 나타내었다.

표 1. 각 동작에 관계하는 특징 값

동작	관계하는 특징 벡터
왼쪽 펀치	양손의 상대 위치
오른쪽 펀치	양손의 상대 위치
왼쪽 킥	양발의 상대 위치
오른쪽 킥	양발의 상대 위치
뛰기	머리, 골반, 양발의 위치
앉기	양발의 상대 위치, 머리 위치
날라차기	양발의 위치, 머리 위치

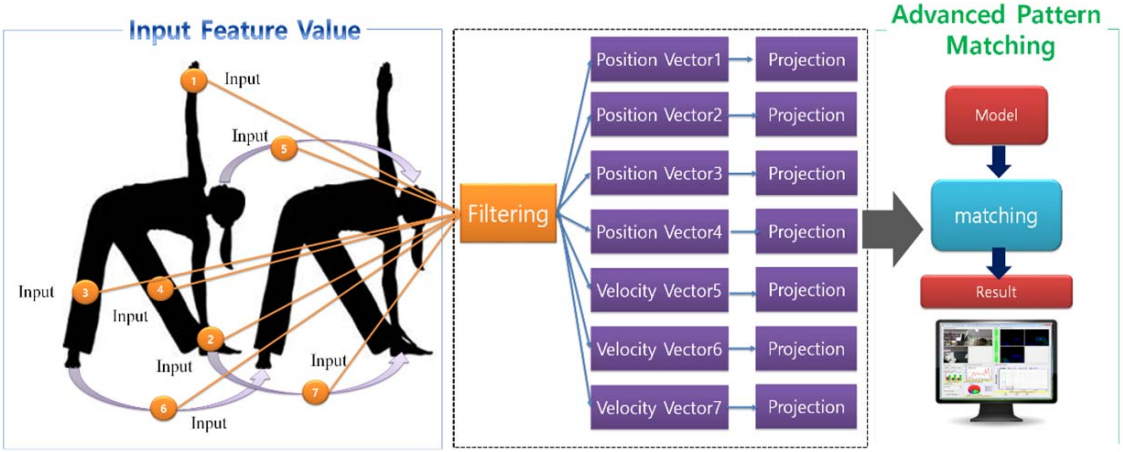

그림 4. 수정된 매칭 알고리즘

Ⅴ. 실험 결과 및 결론

본 시스템은 모델 동작과 입력 동작의 유사도를 측정하기 위하여 각 모델 동작에 대한 고유공간에서의 평균값을 구하고, 이 평균값과의 20차원의 유클리디언 거리를 계산한다. 한편 수정된 매칭 알고리즘에 대한 실험 결과는 그림 5와 같다. 모델 동작과 다르지만 같은 의미를 가지는 7개의 동작 data의 입력 수정전과 입력 수정 후 에 대하여 실험하였고, 시스템은 고유공간 내 에서 모델 동작과의 거리가 가장 가까운 동작 중 임계치(Threshold) 이상의 거리내의 동작을 출력한다.

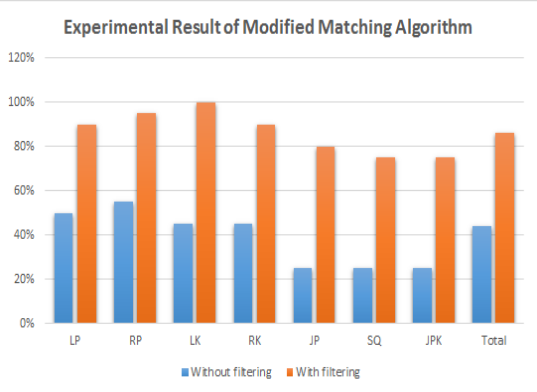

그림 5. 수정된 매칭 알고리즘에 대한 실험결과

본 논문에서는 키넥트 센서로부터 획득한 동작 정보를 이용한 제스처 인식 시스템에 대하여 기술하였다. 인간의 동작을 수치적으로 표현하기 위해 20개의 신체 특징점을 사용했고, 이들 특징점의 정보를 계산하여 동작 정보를 추출하였다. 이러한 데이터를 모델링하기 위해서 주성분 분석법을 사용했으며, 보다 안정적인 인식 결과를 위해 수정된 매칭 알고리즘을 제안하였다. 또한 인식된 결과를 3차원 액션 게임[7]의 인터페이스로 사용함으로써 제스쳐 인식 결과를 사용한 응용프로그램의 예를 제시하였다

References

[1] M. Weiser, "The Computer for the 21st Century", Scientific Anerica, Vol. 265, No. 3, pp. 66-76, 2001.

[2] S. K. Kang, K. Y. Chung, J. H. Lee, "Real-Time Tracking and Recognition Systems for Interactive Telemedicine Health Services", Wireless Personal Communications, Vol. 79, No. 4, pp. 2611-2626, 2014

[3] J. M. Kim, K. Chung, M. A. Kang, "Continuous Gesture Recognition using HLAC and Low-dimensional Space", Wireless Personal Communications, Vol. 86, No. 1, pp. 255-270, 2016.

[4] V. I. Pavlovic, R. Sharma, T. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 677-695, 1997.

[5] Microsoft Kinect SDK,“http://www.microsoft.com /en-us/kinectforwindows”

[6] J. M. Kim, M. A. Kang, "Appearance-based object recognition using higher correlation feature information and PCA", Proc. of the International Conference on Fuzzy Systems and Knowledge Discovery, pp. 1874-1878, 2011.

[7] TEKKEN 7, http://tk7.tekken-net.kr/.