

SVM을 이용한 동적 동작인식: 체감형 동화에 적용

Dynamic Gesture Recognition using SVM and its Application to an Interactive Storybook

이경미

덕성여자대학교 컴퓨터학과

Kyoung-Mi Lee(kmlee@duksung.ac.kr)

요약

본 연구에서는 다차원의 데이터 인식에 유리한 SVM을 이용한 동적 동작인식 알고리즘을 제안한다. 우선, Kinect 비디오 프레임에서 동작의 시작과 끝을 찾아 의미있는 동작 프레임을 분할하고, 프레임 수를 동일하게 정규화시킨다. 정규화된 프레임에서 인체 모델에 기반한 인체 부위의 위치와 부위 사이의 관계를 이용한 동작 특징을 추출하여 동작인식을 수행한다. 동작인식기인 C-SVM는 각 동작에 대해 positive 데이터와 negative 데이터로 구성된 학습 데이터로 학습된다. 최종 동작 선정은 각 C-SVM의 결과값 중 가장 큰 값을 갖는 동작으로 한다. 제안하는 동작인식 알고리즘은 플래시 구연동화에서 더 나아가 유아가 능동적으로 구연동화에 참여할 수 있도록 고안된 체감형 동화 콘텐츠에 동작 인터페이스로 적용되었다.

■ 중심어 : | SVM | 동작 분할 | 동작인식 | 체감형 동화 콘텐츠 | 동작 인터페이스 |

Abstract

This paper proposes a dynamic gesture recognition algorithm using SVM(Support Vector Machine) which is suitable for multi-dimension classification. First of all, the proposed algorithm locates the beginning and end of the gestures on the video frames at the Kinect camera, spots meaningful gesture frames, and normalizes the number of frames. Then, for gesture recognition, the algorithm extracts gesture features using body parts' positions and relations among the parts based on the human model from the normalized frames. C-SVM for each dynamic gesture is trained using training data which consists of positive data and negative data. The final gesture is chosen with the largest value of C-SVM values. The proposed gesture recognition algorithm can be applied to the interactive storybook as gesture interface.

■ keyword : | SVM | Gesture Spotting | Gesture Recognition | Interactive Storybook | Gesture Interface |

1. 서론

현대사회는 정보화 사회라고 불릴 만큼 컴퓨터가 급속도로 발전하고 있다. 특히 공간적, 시간적인 제한이 없는 컴퓨터의 장점은 학습효과의 극대화를 이룰 수 있

어 유아교육 분야에 있어 필수가 되고 있다. 특히, 유아들에게 강력한 전달력과 영향력을 미치는 멀티미디어를 이용한 콘텐츠의 개발 및 활용에 대한 연구가 활발히 이루어지고 있다. 최근에는 동화에 다양하게 적용되고 있는데, 플래시[1]부터 증강현실[2]까지 다양한 방법

* 본 연구는 덕성여자대학교 2012년도 교내연구비 지원에 의해 수행되었습니다.

접수번호 : #130214-007

접수일자 : 2013년 02월 14일

심사완료일 : 2013년 03월 26일

교신저자 : 이경미, e-mail : kmlee@duksung.ac.kr

으로 만들어지고 있다. 그러나 이러한 방법들은 종이책을 단순히 컴퓨터에 옮겨놓은 것으로, 단편적인 체험활동을 중심으로 유아가 수동적 태도를 취하는 단점이 있다[3].

본 논문에서는 언제, 어디서든 읽을 수 있으며, 기존의 종이책보다 유아가 능동적으로 구연동화에 참여할 수 있도록 하는 체감형 콘텐츠를 제안한다. 체감형 동화 콘텐츠에서 동작으로 콘텐츠와 상호작용하는 동작 인터페이스 기술은 Microsoft의 Kinect가 출시되면서 동작인식 기술의 필요성이 대두되고 있다. 카메라의 비디오를 이용한 동작인식은 크게 한 프레임 내에서 이루어진 동작을 인식하는 정적(Static) 동작인식과 여러 프레임에 걸쳐 이루어지는 동적(Dynamic) 동작인식으로 나눌 수 있다. 특히, 동적 동작인식을 위해 신경망[4][5], 결정 트리[6], HMM[7] 등 다양한 인식 알고리즘을 사용할 수 있는데, 시간에 따라 변하는 동적 동작의 특성상 연속된 입력 데이터 인식이 장점인 HMM이 가장 대표적인 방법이다. 그러나, HMM은 다차원의 동작 데이터를 일차원 데이터로 이산화 시켜야 하는 단점이 있다. 본 논문에서는 다차원의 데이터 인식에 유리한 서포트 벡터 머신(Support Vector Machine : SVM)을 이용하여[8] 동적 동작인식을 수행하였다. 제안하는 방법은 연속해서 들어오는 동적 동작데이터에서 인체부위 검출, 동작 분할(Gesture Spotting), 동작 특징 추출, SVM을 이용한 동적 동작인식 등의 기술을 사용하여 보다 쉽게 동작 최적화를 이루고 인식 기능이 뛰어난 방법임을 증명한다.

특히, 본 논문에서는 체감형 동화 시스템에 SVM을 이용한 동작 인터페이스를 구현 및 적용하여 유아 동작의 인식 범위를 넓힘으로써 유아가 동화 속 주인공이 되어 보다 자유롭게 동작할 수 있는 체감형 콘텐츠의 가능성을 보인다. 이는 유아의 교육과 재미 뿐 아니라 운동능력 향상과 집중력을 높일 수 있다.

II. 동작인식을 이용한 체감형 동화 시스템

1. 동작 인터페이스를 활용한 동화 콘텐츠

동작인터페이스를 활용한 ‘체감형 동화’ 시스템은 동

화에 등장하는 기본 동작 요소의 탐색에 기초하여 동작 인식 기술을 활용한 콘텐츠를 개발하여야 한다[3].

본 논문에서는 유아의 발달과 흥미를 고려하기 위해 유아의 일상생활과 친숙한 ‘생일파티’를 주제로 하여 ‘동물’을 소재로 한 동화 시나리오를 작성하였다. 동화의 흐름에 따라 동작을 반영하기 위해 사용자가 주인공 역할을 수행하도록 이야기의 흐름을 구성하였고, 사용자가 수행한 동작에 따라 동화가 진행되도록 하였다. 표 1은 ‘생일잔치 가는 길’이라는 동화 콘텐츠의 동화 내용을 보여준다.

체감형 콘텐츠에 사용되는 동작 인터페이스로서 동작은 동작하기 쉽고, 유일성을 보장하며, 겹침이 없고, 제한된 정확성요구로 정의될 수 있다[9]. 체감형 동화에 사용되는 동작은 동화의 내용을 담아내는 동작으로 가상공간에서 자연스런 상호작용을 할 수 있는 동작으로 선정한다. [표 1]은 ‘생일잔치 가는 길’에서 사용되는 동작을 제시한다.

표 1. 동화 콘텐츠 및 동작 인터페이스[3]

동화 콘텐츠	동작 인터페이스	
(체감형 동화 시작)	두 다리를 벌리고 두 팔을 들기	정적
"색칠하듯이 손을 좌우로 움직여 보세요."	한 팔로 좌우로 움직이기	동적
"머리 위로 손을 올려 사과를 따보세요."	한 손 들기	정적
"장검다리를 건너서 가고 싶은 친구는 그 자리에서 점프해 주세요."	두 발을 들어 점프하기	동적
"수영해서 건너갈 친구는 수영하는 흉내를 내보세요."	두 팔을 앞으로 움직여 수영하기	동적
"고양이 옆에서 큰북을 연주해보세요. 한손으로 쿵,쿵 두드려 볼까요?"	한 팔을 옆으로 움직여 큰북치기	동적
"얼마나 큰 빵을 만들었을까요? 양팔을 이용해 큰 빵을 표현해보세요."	두 팔을 머리 위로 올려 동그라미 그리기	정적
"나비처럼 훨훨 날아볼까요? 양팔을 벌려 훨훨 날아보세요."	두 팔을 양 옆으로 벌려 날개 짓하기	동적
"양손을 머리 위로 올려, 선반위의 모자를 꺼내서 오리에게 주세요."	두 팔 들기	정적

2. 동작 인터페이스를 활용한 체감형 동화 시스템

본 논문에서 제안하는 ‘체감형 동화’ 시스템은 동화에 등장하는 캐릭터가 요구하는 동작을 유아가 수행함으로써 동화를 진행하도록 구성된다. 제안하는 시스템의 구성도는 [그림 1]과 같다. 첫째, 동작인식 부분에서는 카메라에서 들어온 프레임으로부터 인체의 관절값을

받아 동작을 인식하는 부분으로, [표 1]의 동작 인터페이스에 제시된 것처럼 필요한 동작을 정의하고 인식하는 부분이다. 둘째, 체감형 동화 부분은 [표 1]의 동작 콘텐츠에 제시된 것처럼 동화 page들이 제어 신호에 따라 차례대로 보이는 부분이다. 마지막으로 동화 제어부분은 동작인식 부분에서 결정된 동작이 체감형 동화 page에 맞는지 판별하여 체감형 동화의 진행여부를 제어하는 부분이다. [그림 1]과 같은 구성도는 다양한 동화 콘텐츠와 동작 인터페이스로 확장 가능한 시스템 구성도이다.

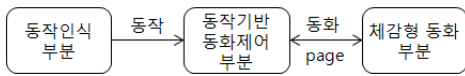


그림 1. 체감형 동화 시스템 구성도

III. 동화 콘텐츠를 위한 동작인식 시스템

‘체감형 동화’ 시스템에서 사용하는 동작인식 부분과 동화제어 부분의 자세한 흐름도는 [그림 2]와 같다.

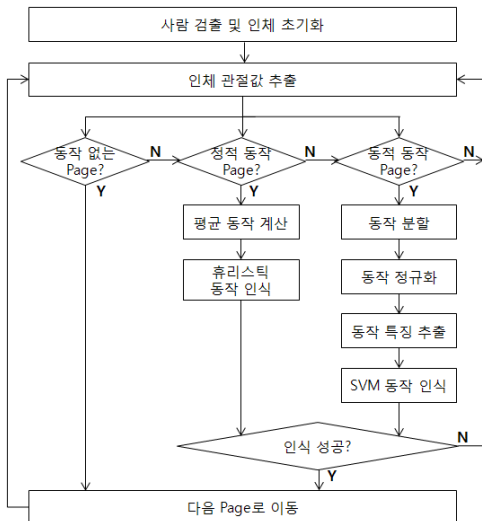


그림 2. 체감형 동화를 위한 동작인식 흐름도

처음에 Kinect 카메라로 사용자를 검출하고 인체모델과 매칭시키는 초기화 작업을 수행한다. 인체모델을 통해 인체의 관절 값을 획득하고, 이를 이용하여 현재

동화에서 요구하는 동작인지 판별하여 동화제어 부분으로 넘긴다. 동화제어 부분은 각 page마다 요구되는 동작이 올바르게 인식되었을 때만 다음 page로 진행한다. 잘못된 동작 등으로 오인식이 일어나면 시스템은 다시 동작을 수행할 것을 요구한다.

1. 사람 검출 및 인체 모델 초기화

들어오는 비디오 입력 영상에서 사람을 검출하기 위해선 우선 배경 영상에서 사람을 분리하여야 한다. 본 논문에서는 Kinect의 적외선 거리 센서를 사용하는데, 초당 30프레임의 데이터가 입력되는 센서를 통해 세그멘테이션 맵을 구성하고 이 맵을 통해 사용자와 배경을 분리해서 사용자를 인식한다.

다음으로 검출된 사람을 미리 정의한 인체모델의 각 관절과 일치시키는 초기화 작업을 진행시킨다. 초기화 작업은 사람의 관절을 추적하여 동작을 인식하는데 중요한 과정이다. 초기화를 쉽게 하기 위해서 사용자가 다리를 어깨 넓이만큼 벌리고 두 손을 드는 마치 역기를 드는 듯한 자세인 PSI 자세를 취하면, 캘리브레이션 (Calibration)이 수행되고 각 관절의 값을 찾게 된다.

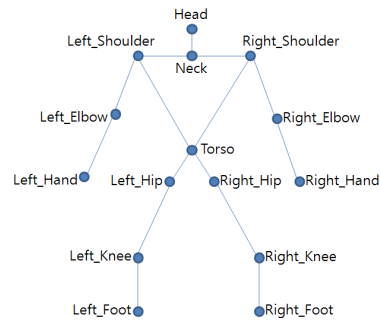


그림 3. Kinect 인체 관절 모델

본 논문에서는 Kinect에서 사용된 관절 모델을 사용하여 [그림 3]처럼 15개의 관절로 이루어진 인체 모델 h_i ($i = 1 \dots 15$)을 사용했다. 이 인체 모델은 Kinect에서 사용자가 분리되면 먼저 입체감을 분석하여 골격의 형태를 포착하는 것으로 한 번에 전신 골격을 포착하여 구성하게 된다. 이렇게 구성된 인체 관절 모델을 이용

해 연속된 프레임에서 찾은 각 관절의 x, y, z 값을 이용하여 관절의 위치 값을 추출한다.

2. 정적 동작인식

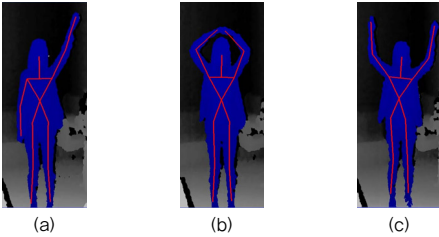


그림 4. 정적 동작에 대한 평균 동작 예 : (a) 한 손 들기, (b) 두 팔을 머리 위로 올려 동그라미 그리기, (c) 두 팔 들기

정적 동작은 한 프레임에서 이루어지는 동작을 말하지만, 한 프레임의 정보만을 이용하면 관절의 순간적인 떨림이나 부분 자세로서의 정적 동작을 포함한 경우, 잘못 동작을 취한 경우 및 잡음에 민감할 수 있다. 따라서 본 논문에서는 이와 같은 인터페이스로서의 민감성을 조절하기 위해서 사용자에게 일정 시간 동안 정적인 동작을 취하도록 요구하고, 그 시간 동안 취한 동작들에 대한 평균 동작을 계산한다. 본 논문에서는 3초간 정적인 자세를 유지하여야 유효한 동작으로 인정하고, 이때 들어 온 3초간의 인체 모델로 평균 동작 모델 μ_i 을 계산한다.

$$\mu_i = \frac{\sum_{t=0}^{T-1} h_i^t}{T}, \quad i = 1 \dots 15$$

여기서 h_i^t 는 t 번째 프레임에서 인체모델의 i 번째 관절값이고, T 는 일정 시간동안, 즉 3초 동안 들어온 프레임의 수이다. [그림 4]는 본 논문에서 사용하는 3개의 정적 동작에 대한 평균 동작 예를 보여주고 있다.

입력 프레임에서 평균 동작을 계산한 후, 평균 동작 모델의 관절들 사이의 상대적 위치관계를 반영하는 평가함수를 이용하여 동작을 인식한다. 예를 들어, [표 1]

에서 제시하는 동작 중 ‘한 손 들기’의 경우 한 손(hand)의 높이가 머리(head)의 높이보다 위에 있고 다른 손의 높이가 어깨(shoulder) 아래에 있는 상태이면 ‘한 손 들기’ 동작으로 인식한다. 이 때, 왼손과 오른손을 구별하지 않는다. 즉, (평균 동작의 왼손의 y 좌표 > 머리의 y 좌표) 또는 (평균 동작의 오른손의 y 좌표 > 머리의 y 좌표)이다. ‘두 팔 들기’는 양 손의 높이가 머리의 높이보다 위에 있다고 정의한다. 즉, (평균 동작의 왼손의 y 좌표 > 머리의 y 좌표) 그리고 (평균 동작의 오른손의 y 좌표 > 머리의 y 좌표)이다. 또한, ‘두 팔을 머리 위로 올려 동그라미 그리기’는 양 손의 높이가 머리의 높이보다 위에 있으며 팔꿈치(elbow) 사이의 x 축 간격이 두 손의 간격보다 큰 경우를 말한다. 즉, 두 손이 머리 위에 있으면서, (양 팔꿈치의 x 좌표의 차이) > (양 손의 x 좌표의 차이+ α)이고, 본 논문에서 α 는 60이다.

3. 동적 동작 분할(Dynamic gesture spotting)

정적 동작과 달리 동적 동작은 여러 프레임에 걸쳐 이루어진 동작을 의미한다. 따라서 동적 동작을 인식하기 위해서는 연속적인 프레임에서 어느 프레임이 동적 동작의 시작이고, 어느 프레임이 동작의 끝인지를 파악하여 그렇지 않은 프레임을 제거하여 의미있는 동작 프레임만을 가려내는 작업이 필요하다.

일반적으로 동작 분할은 동작을 수행하는 도중 발생할 수 있는 가장 큰 특징을 이용하여 분할된다. 이 방법에는 첫째, 속도를 기준으로 하여 동작 시작 전과 후의 일정 시간 동안 멈추는 동작의 특징을 이용하는 방법과 둘째 동작의 시작과 끝에 많이 나타나는 동작의 특징을 이용하는 방법, 마지막으로 심한 굴곡 포인트를 가진 프레임을 기준으로 하는 방법이 있다[10]. 우선 본 논문에서는 동작 분할을 용이하게 하기 위해서, 하나의 동적 동작이 끝나는 순간 사용자는 그 모습 그대로 잠시 멈추도록 하여 동적 동작의 끝에 멈춤을 추가하였다. 여기에 동작의 크기, 속도, 굴곡이 크게 변한 프레임을 찾아 동작 분할의 기준으로 삼았다. 특히, 굴곡률은 양 어깨, 팔꿈치, 엉덩이(hip), 무릎(knee)에 대해 30도 이상의 변화가 일어난 프레임을 이용했다. 본 논문에서는 단일 동작만을 대상으로 하므로, 동작의 중복은 허용하

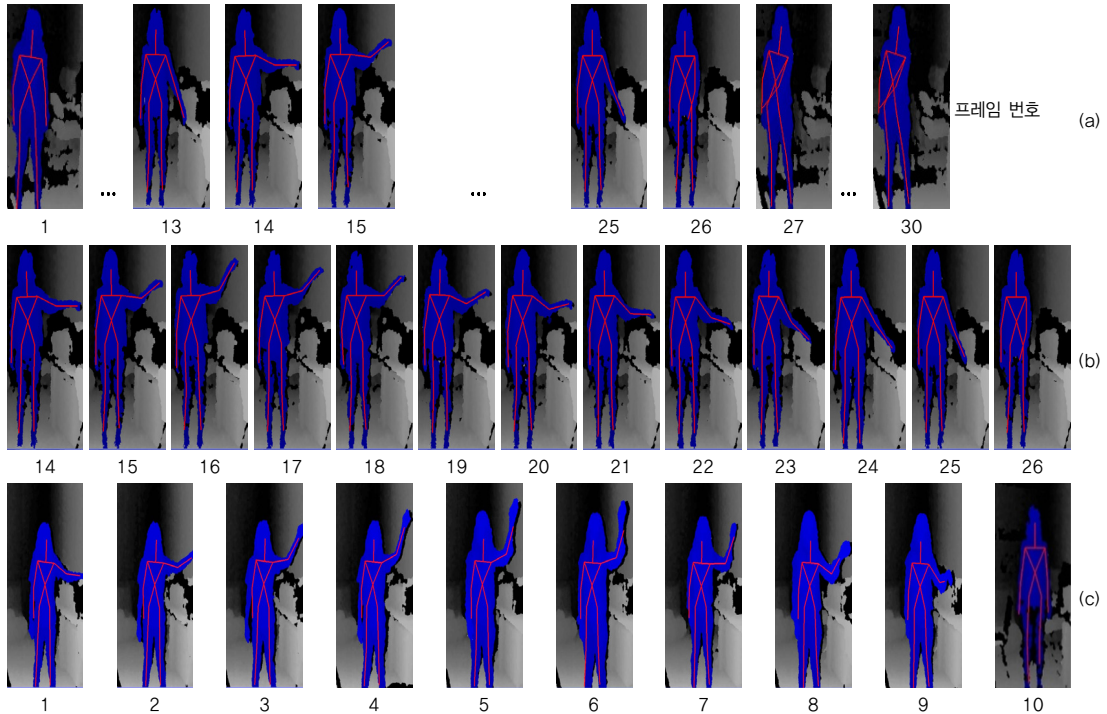


그림 5. ‘한 팔을 옆으로 움직여 큰북치기’ 동작 : (a) 모든 프레임, (b) 동작 분할, (c) 동작 정규화

지 않는다. 즉, 한 동작의 끝 프레임을 찾은 후에만 다음 동작의 시작 프레임을 찾을 수 있도록 한다.

[그림 5]는 ‘한 팔을 옆으로 움직여 큰북치기’ 동작에 대한 동작 분할 과정을 보여주고 있다. [그림 5(a)]는 연속된 동작 프레임의 일부를 보여주고 있다. [그림 5(b)]는 [그림 5(a)]의 동작 프레임 중 동작 분할 조건을 만족하는 프레임들만 보여주고 있다. 즉, 프레임 13과 프레임 14 사이와 프레임 26과 프레임 27 사이에서 동작 분할 기준을 만족하였다. 따라서 본 논문에서는 프레임 14에서 프레임 26까지의 프레임들을 ‘한 팔을 옆으로 움직여 큰 북치기’ 동작의 프레임으로 분할할 수 있었다.

4. 동작 정규화 및 특징 추출

동적 동작에 대한 분할이 이루어지면, 동작을 구성하는 의미 있는 프레임들의 집합을 얻게 된다[그림 5(b)]. 그러나 같은 동작이라도 동작을 할 때마다 분할된 프레임의 수 T 는 다를 수 있다. 이는 동작인식 알고리즘이

동일한 크기의 입력 정보를 요구하는 경우에는 적절하지 않다. 따라서 본 논문에서는 N 개의 동일한 프레임 개수를 가지게끔 프레임 수 정규화를 수행한다.

알고리즘 I은 동적 동작의 프레임 수 정규화에 대한 알고리즘을 보여주고 있는데, 입력된 동작 프레임에서의 인체 모델 h_i 는 정규화를 통해 정규화된 인체모델 λ_i 를 생성하게 된다. 정규화된 인체모델의 첫 프레임 λ_i^0 는 동작의 첫 프레임을 포함하며, 이후의 정규화된 동작 모델의 j 번째 동작 λ_i^j 은 인접한 두 프레임의 동작을 보간하여 구할 수 있다.

동적 동작의 프레임의 수를 동일하게 정규화시킨 후엔, 동작의 특징을 추출하여야 한다. 동작 특징 추출은 동작을 표현하는 공통적인 성질을 추출하는 것으로 복잡하고 많은 양의 자료 대신 두드러지는 특징만을 통해 동작을 인식하는 것을 말한다. 가장 일반적인 동작 특징추출 방법은 인체의 특정 부분을 찾아 인식하는 것으

로 인체의 각 부위의 위치와 각 부위간의 위치 관계를 이용할 수 있다. 본 논문에서는 인체의 각 부위들의 3차원 위치와 각 부위들 간의 관계를 이용하였다. 부위들 간의 관계는 [그림 3]과 같은 부모-자식의 관계를 갖는 트리형태로 구성되는 계층적 인체 모델을 이용한다. [표 1]의 동작들은 양 팔과 다리의 위치와 관계가 중요한 반면, 머리의 위치나 관계는 동작의 특징을 구별하는데 큰 역할을 못한다. 따라서 본 논문에서는 [표 1]의 15개 인체 부위 중 머리와 목(neck)을 제외한 13개 부위의 3차원 위치값과 양 팔꿈치 각도와 양 무릎 각도를 이용한다.

알고리즘 1. 동적 동작 프레임 수 정규화

```

for( j = 0 ; j < N ; j++ )
    (1) j 가 0이면,  $\lambda_i^0 = h_i^0$ 
    (2) j 가 0이 아니면,
        i) float  $x_1 = \frac{j-1}{N-1} * T(\text{분할된 프레임 수})$ 
        ii) int  $x_2 = \text{floor}(x_1)$ 
        iii) int  $x_3 = x_2 - 1$ 
        iv)  $\lambda_i^j = (x_2 - x_1) * h_i^{x_3} + (x_1 - x_3) * h_i^{x_2}$ 
    
```

5. SVM을 이용한 동적 동작인식

정규화된 프레임으로부터 특징값을 추출한 후, 동작 인식을 수행한다. 본 논문에서는 SVM 학습알고리즘을 통해 동적 동작을 인식하였다. SVM은 1995년 V. N. Vapnik에 의해 개발된 통계적 학습이론으로[8], 현재 알려져 있는 많은 인식 알고리즘 중 가장 인식 성능이 뛰어난 학습 모델 중 하나이다. SVM은 데이터들을 주어진 공간보다 높은 차원의 공간으로 변환하고 나서, 각 클래스들을 구분하는 최적 분리 경계면을 구하기 위해 분리 경계면과 분리 경계면에 가장 인접한 점인 support vector와의 거리를 최대화한다. SVM은 비선형의 복잡한 결정경계를 가지는데 이 때문에 학습시간이 오래 걸리지만, 그만큼 매우 정확하다. 또 신경망을 이용할 경우 입력층/은닉층/출력층의 개수와 학습률을 정해야하는데 과정이 번거롭고 최적화하기 힘든 반면

SVM은 비중을 크게 줄 목적분류와 커널함수, 매개함수만 정하면 되기 때문에 최적화가 상대적으로 쉽고 분류 문제에 있어서 일반화 기능이 뛰어나 많은 분야에서 응용되고 있다.

본 논문에서는 C-SVM을 이용하여 5개의 동적 동작 각각에 대한 이진 분류를 수행한다[11]. 각 동작에 대한 C-SVM을 학습시키기 위한 학습데이터는 positive 데이터와 negative 데이터로 구성된다. positive 데이터는 올바르게 취한 동작에서 추출할 수 있으므로, 동작 분할과 정규화 과정을 통해 자동으로 특징 데이터를 구할 수 있다. 그러나 대부분의 negative 데이터는 별도로 만들어야 하는데, positive 데이터와 너무 다르다면 예측 성능이 떨어질 수 있으므로 positive 데이터와 유사하면서 negative 데이터로 사용할 수 있어야 한다[12]. 본 논문에서는 positive 데이터의 일부만 포함시키거나 동작 분할 전후의 동작을 포함시켜 negative 데이터를 만들었다. 이렇게 만들어진 positive와 negative 동작 데이터에 대해, positive/negative 비율을 1/1로 학습시켰다. 마지막으로 5개의 C-SVM에서 각 동작에 대한 예측 결과가 나오면, 가장 큰 값을 가진 동작을 최종 동작으로 선정한다[그림 6].

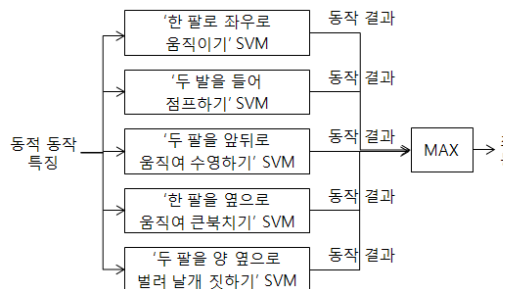


그림 6. SVM을 이용한 동적 동작인식

IV. 실험 결과

1. 데이터

Kinect 카메라로부터 들어오는 비디오 프레임에서 동작인식을 수행하기 위해서 정적 동작과 동적 동작으로 나누어 테스트하였다. 동작 데이터를 위해 10명의

사용자로부터 정적 동작은 동작별로 각 2개씩 총 20개를, 동적 동작은 동작별로 각 5개씩 총 50개를 획득하였다. 동적 동작은 학습이 필요하므로, 30개는 학습 데이터로, 나머지 20개는 테스트 데이터로 사용하였다. 각 30개의 학습 데이터는 positive 데이터로 사용하였고, III-5장에서 언급한 것처럼 positive 데이터를 이용하여 30개의 negative 데이터를 만들었다. [그림 7]은 실험에 사용된 5개 동적 동작의 positive 데이터 예로, 동작 비디오 프레임 중 일부 프레임만을 보여주고 있다.

2. 인식률

본 논문에서 사용하는 3개의 정적 동작에 대한 인식 결과는 96.67%로, [표 2]와 같다. 대부분의 정적 동작에 대해서 매우 잘 인식하였으나, ‘두 팔을 머리 위로 올려 동그라미 그리기’ 동작에 대해서 머리(head) 위에 양 손(hand)이 위치하지만, 양 손의 간격과 양 팔꿈치(elbow)의 간격이 비슷하게 되면, ‘두 팔을 들기’로 오인식되는 경우가 있었다. [표 1]에 소개된 5개의 동적 동작에 대한 인식 결과는 91%로 [표 3]과 같다. ‘두 발을 들어 점프하기’는 유일한 다리 동작으로 100%의 인식률을 보였다. ‘한 팔을 옆으로 움직여 큰북치기’는 한 손(hand)이 가로로 한 번 북을 친 후에 멈추지 않고 되돌리는 연속 동작이 포함되는 경우엔 ‘한 팔로 좌우로 움직이기’로 오인식 되었다. ‘두 팔을 앞뒤로 움직여 수영하기’는 두 손을 x축을 중심으로 회전하여 움직일 때 양 손의 x축 간격이 넓게 벌려 수영하는 동작을 취하는 경우가 종종 발생해서 ‘두 팔을 양 옆으로 벌려 날개 짓하기’와 혼동을 일으켰다.

비교 실험을 위해 같은 데이터를 HTK(HMM Tool Kit)에 기반한 HMM에 적용하여[13] 78%의 인식률을 얻었다. 일반적으로 HMM은 적은 수의 학습 데이터에서는 SVM보다 우수한 인식 결과를 보이지만, 학습 데이터의 수가 증가할수록 SVM의 인식률이 좋아진다고 알려져 있다[14]. 이는 본 논문에서 SVM이 동작 별로 30개의 positive 데이터를 사용하는 것이 각 동작이 가진 다양한 변이를 충분히 반영하면서 다른 동작과 구별할 수 있는 유사성을 가지고 있음을 의미하여, SVM이 가진 우수성을 보여주고 있다.

표 2. 정적 동작인식 결과

입력 동작	결과 동작	한 손 들기	두 팔을 머리 위로 올려 동그라미 그리기	두 팔 들기
	한 손 들기	100%	-	-
두 팔을 머리 위로 올려 동그라미 그리기	-	90%	10%	
두 팔 들기	-	-	100%	

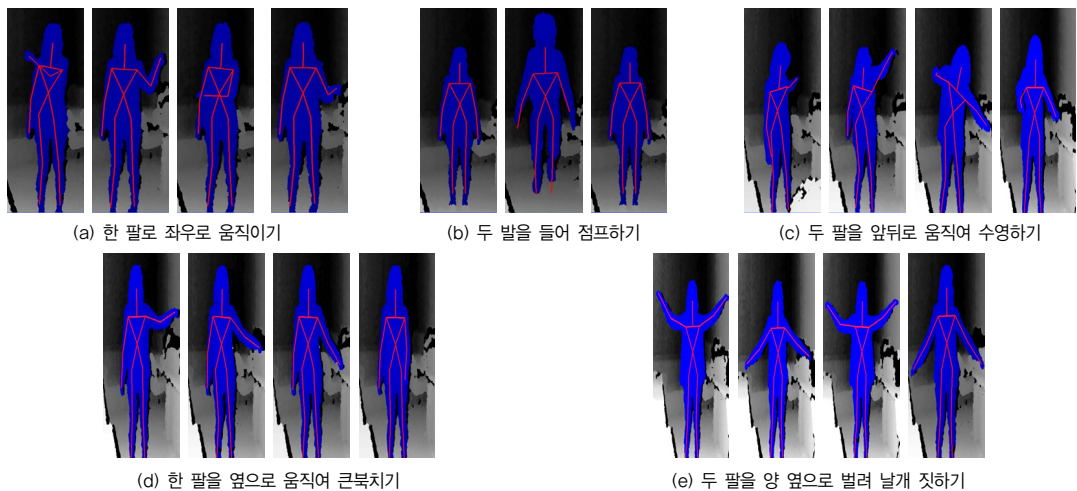


그림 7. 동적 동작의 positive 데이터 예

표 3. 동적 동작인식 결과

입력 동작	결과 동작	한 팔로 좌우로 움직이기	두 발을 들어 점프하기	두 팔을 앞뒤로 움직여 수영하기	한 팔을 옆으로 움직여 큰북치기	두 팔을 양 옆으로 벌려 날개 짓하기
한 팔로 좌우로 움직이기		95%	-	-	5%	-
두 발을 들어 점프하기		-	100%	-	-	-
두 팔을 앞뒤로 움직여 수영하기		-	-	75%	-	25%
한 팔을 옆으로 움직여 큰북치기		10%	-	-	90%	-
두 팔을 양 옆으로 벌려 날개 짓하기		-	-	5%	-	95%

V. 결론

본 논문에서는 휴리스틱 평가함수를 이용하여 정적 동작을 인식하고 SVM을 이용하여 동적 동작을 인식하는 방법을 제안하였다. 이 동작인식 시스템은 체감형 동화에 동작 인터페이스로 적용되어 사용자의 몰입도와 흥미를 높일 수 있었다. 제안하는 시스템의 장점은 다음과 같다.

본 논문에서는 SVM을 이용하여 시간변이 데이터에 대한 인식 방법을 제안하였다. 시간변이 데이터에 유리한 HMM을 사용하는 대신, 다차원 데이터 분류에 유리한 SVM을 적용하였다.

제안하는 방법은 동적 동작을 인식하기 위해 동작 분할(Gesture Spotting) 및 정규화를 통한 동작 특징 추출 및 충분한 수의 학습데이터를 이용하였다.

또한, 일반적으로 인체 부위를 이용하여 동작 특징을 추출하는 것은 정확한 동작인식에 도움이 되지만, 정확하게 부위를 찾아야 한다는 단점이 있다. 이와 같은 인체 부위 독립적인 방법의 문제점을 극복하기 위해서 본 논문에서는 인체 부위로 구성된 인체 모델을 이용하였다.

제안하는 연구는 SVM을 이용한 동적 동작인식을 통해서 유아 동작 인식의 범위를 넓힘으로써, 유아가 동화 속 주인공이 되어 보다 자유롭게 동작할 수 있는 체감형 콘텐츠를 가능케 하는 장점이 있다. 이는 유아의 교육과 재미 뿐 아니라 운동능력 향상과 집중력을 높일

수 있습니다.

제안하는 동작인식 시스템은 향후 다양한 동작을 추가하여 다양한 동작 인터페이스를 사용한 체감형 동화 개발에 도움을 줄 수 있을 것이다. 또한, 동적 동작의 인식률을 향상시키기 위해 동작들을 분명하게 구분해 줄 수 있는 특징을 추출하는 방법에 대한 연구가 필요하다.

참고 문헌

- [1] P. Boland, D. Kavanagh, and J. Squires, "iMARK Interactive Media Annotation Resource Kit," 2010, File Available at: <https://dspace.ndlr.ie/handle/10633/5508>.
- [2] T. G. Kirner, F. M. V. Reis, and C. Kirner, "Development of an interactive book with augmented reality for teaching and learning geometric shapes," in proc. of Iberian conf. on Information Systems and Technologies, pp.1-6, 2012.
- [3] 이경옥, 이경미, 이상희, "유아 동작에 기초한 interactive 동화콘텐츠 개발 연구", 어린이미디어 연구논문지, 제11권, 제2호, pp.133-154, 2012.
- [4] N. A. Ibraheem and R. Z. Khan, "Vision based gesture recognition using neural networks approaches: a review," Int'l J. of Human Computer Interaction, Vol.3, No.1, pp.1-14, 2012.
- [5] 이동규, 이기정, 황보택근, 임혁규, "신경망을 이용한 동작분석과 원격 응급상황 검출 시스템", 한국콘텐츠학회논문지, 제6권, 제9호, pp.50-59, 2006.
- [6] M. J. Jeon, S. W. Lee, and Z. Bien, "Hand gesture recognition using multivariate fuzzy decision tree and user adaptation," Int'l J. of Fuzzy System Applications, Vol.1, Issue 3, pp.15-31, 2011.
- [7] W. Xiaojuan, "Dynamic gesture track recogni

tion based on HMM," in proc. of IEEE conf. on VSLI Design and Video Technology, pp.169-174, 2005.

- [8] V. N. Vapnik, The nature of statistical learning theory, Springer-Verlag, 1995.
- [9] 이경미, "가상공간에서 동작을 이용한 체감형 게임 인터페이스 설계", 한국컴퓨터게임학회논문지, 제24권, 제1호, pp.133-144, 2011.
- [10] H. Kang, C. Lee, and K. Jung, "Recognition-based gesture spotting in video games," Pattern Recognition Letters, Vol.25, No.15, pp.1701-1714, 2004.
- [11] C. Chang and C. Lin, LIBSVM: a library for support vector machines, Software Available at: <http://www.csie.ntu.edu.tw/~wcjlin/libsvm>.
- [12] 이제근, 김수진, 장병탁, "학습을 위한 네거티브 데이터가 존재하지 않는 경우의 microRNA 타겟 예측 방법", 한국컴퓨터종합학술대회 논문집, 제35권, 제1(C)호, pp.212-216, 2008.
- [13] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, V. Valtchev, and P. Woodland, "The HTK Book," Microsoft Corporation and Cambridge University Engineering Department, 2009.
- [14] E. J. R. Justino, F. Bortolozzi, and R. Sabourin, "A comparison of SVM and HMM classifiers in the offline signature verification," Pattern Recognition Letters, Vol.26, No.9, pp.1377-1385, 2005.

저 자 소 개

이 경 미(Kyoung-Mi Lee)

정회원



- 1993년 2월 : 덕성여자대학교 전산학과(이학사)
 - 1996년 2월 : 연세대학교 전산학과(이학석사)
 - 2001년 2월 : 아이오와 주립대학교 전산학과(이학박사)
 - 2003년 3월 ~ 현재 : 덕성여자대학교 컴퓨터학과 교수
- <관심분야> : 영상처리, 패턴인식, 멀티미디어, HCI