

A Bio-Inspired Modeling of Visual Information Processing for Action Recognition

JinOk Kim[†]

ABSTRACT

Various literatures related computing of information processing have been recently shown the researches inspired from the remarkably excellent human capabilities which recognize and categorize very complex visual patterns such as body motions and facial expressions. Applied from human's outstanding ability of perception, the classification function of visual sequences without context information is specially crucial task for computer vision to understand both the coding and the retrieval of spatio-temporal patterns. This paper presents a biological process based action recognition model of computer vision, which is inspired from visual information processing of human brain for action recognition of visual sequences. Proposed model employs the structure of neural fields of bio-inspired visual perception on detecting motion sequences and discriminating visual patterns in human brain. Experimental results show that proposed recognition model takes not only into account several biological properties of visual information processing, but also is tolerant of time-warping. Furthermore, the model allows robust temporal evolution of classification compared to researches of action recognition. Presented model contributes to implement bio-inspired visual processing system such as intelligent robot agent, etc.

Keywords : Bio-Inspired Visual Information Processing, Action Recognition, Spatio-Temporal Correlation, Recognition of Visual Sequence

생체 기반 시각정보처리 동작인식 모델링

김진옥[†]

요약

신체 동작, 얼굴 표정과 같이 아주 복잡한 생체 패턴을 인식하고 분류하는 인간의 능력을 모방한 정보처리 컴퓨팅 관련 연구가 최근 다수 등장하고 있다. 특히 컴퓨터비전 분야에서는 인간의 뛰어난 인지 능력 중 상황정보 없이 시각시퀀스에서 동작을 분류하는 기능을 통해 시공간적 패턴 코딩과 빠른 인식 방법을 이해하고자 한다. 본 연구는 비디오 시퀀스상의 동작인식에 생물학적 시각인지과정의 영향을 받은 생체 기반 컴퓨터비전 모델을 제시하였다. 제안 모델은 이미지 시퀀스에서 동작을 검출하고 시각 패턴을 판별하는 데 생체 시각처리과정의 신경망 구조 단계를 반영하였다. 실험을 통해 생체 기반 동작인식 모델이 인간 시각인지 처리의 여러 가지 속성을 고려했을 뿐 아니라 기존 동작인식시스템에 비해 시간 정확성이 뛰어나며 시간 변화에 강건한 분류 능력을 보임을 알 수 있다. 제안 모델은 지능형 로봇 에이전트와 같은 생체 기반 시각정보처리 시스템 구축에 기여할 수 있다.

키워드 : 생체 기반 시각정보처리, 동작인식, 시공간적 상관관계, 시각적 순서 인식

1. 서론

뇌의 정보 인지능력에 대한 동작 원리를 연구하고 있는 생물심리학, 신경생리학과 더불어 인지과학과 인공지능 분야에서는 인간의 뇌를 거대한 정보처리 시스템으로 보고 인간의 뇌 기능을 모방하여 정보처리를 하는 지능형 컴퓨터를

만드는 노력을 계속하고 있다[1][2]. 특히 컴퓨터비전과 로봇틱스 연구에서 인지과학을 적극적으로 컴퓨터 시스템에 접목하기 위해 인간의 인지 기능과 정보 처리과정을 컴퓨터에 응용하는 방안을 모색하고 로봇에게 인간의 감정 표현 방법과 유사한 기능을 제공하기 위해 감정 매카니즘을 담당하는 뇌 영역과 신경망 간의 시간동기화에 대한 연구를 다수 제안하고 있다[3]. 또한 컴퓨터비전 연구자들은 뇌에서 어떻게 외부 시각 자극 데이터를 입력받는지, 어떤 데이터를 어떤 식으로 선택하고 뇌 내부에서 어떤 과정으로 시각정보화를 하는지에 대해 신경생리학을 통해 도출된 뇌 정보처리 기능

[†] 종신회원 : 대구한의대학교 모바일콘텐츠학부 교수
Manuscript Received : March 3, 2014
First Revision : May 26, 2014; Second Revision : June 17, 2014
Accepted : June 20, 2014
* Corresponding Author : JinOk KIM(bit@dhu.ac.kr)

을 컴퓨터 지능에 반영하고자 한다.

생체 기반의 시각정보처리 방식을 이용한 본 연구는 동작 인식에 뇌의 생체 시각처리 시스템 기능을 모방한 매카니즘을 적용하여 동작을 효과적으로 인식하는 방법을 제안하고자 한다. 뇌가 동작을 인식하는 과정을 분석한 다음 이를 컴퓨터시스템의 동작인식에 적용함으로써 컴퓨터비전의 주요 연구주제인 동작인식 연구에 뇌의 시각 인지 능력을 모방한 기능을 접목한 것이다.

연구에서 제시하는 동작인식 모델은 인간 두뇌에서 구동하는 시각인식 기능의 시간감지 속성을 반영한 시공간적 시각 시퀀스의 코딩과 판별을 목적으로 하고 있으며 비디오 시퀀스에서 추출한 몇 개의 동작 특징점으로도 동작 분류를 처리하고 3D 형태의 동작을 비디오 시퀀스의 2D 코딩을 통해 인식하여 생체 기반의 영리한 동작인식 방법을 제안함으로써 컴퓨터비전 기술 발전에 기여한다.

또한 동작인식 시 복잡한 신체 모델을 선구축하여 동작을 분류할 때 전체적으로 제시되는 비디오 시퀀스 전체를 모두 이용하는 기존 동작인식 방법과 달리 명확한 신체 모델 없이 이미지 시퀀스를 시간 변화에 따라 분류하고 부분적 동작 특징을 조합하여 전체 동작으로 인식하는 과정에서 시퀀스 입력 초기에 이미 동작인식 결과를 빠르게 제시하여 인간의 시각인지 능력과 유사한, 빠른 동작인식 결과를 보인다는 점에서 기존 연구와 차별성이 있다.

본 연구는 2장에서 시각인지 특성과 컴퓨터비전에 대한 개요를 기술하고 기존 연구 관련 내용을 기술하였다. 3장에서는 제안 동작인식 모델을 설명하였으며 4장에서 실험 결과를 제시한 후 5장에서 결론을 맺었다.

2. 관련 연구

생물학과 신경생리학에서는 뇌의 운동 시각 인식과정을 밝히기 위해 뇌의 대뇌피질 영역 중 시각영역인 V1(primary visual area)과 V5영역인 중간 관자엽 MT(middle temporal area)에 분산 분포하는 뉴런이 어떻게 자동조직화를 통해 수렴하면서 협력 및 경쟁적 상호작용을 하는지에 대한 뉴런 기능 기반의 인지능력 연구[4][5]와 동작인식과 움직임의 구성 요소 검출을 담당하는 내측상관자엽(MST: medial superior temporal area)영역의 기능에 대한 연구를 진행하여 꾸준히 동작인식 처리과정의 매카니즘을 규명하고 있다[6].

이러한 생물학 및 신경생리학 연구에서 영향을 받아 1990년 브룩스(Brook)가 센서모터 경험을 환경에 따라 능동적으로 구축하고 활용하는 로봇 에이전트에 대한 연구를 시도한 이래로 다양한 신경생리학적 운동을 컴퓨팅 모델로 설명하는 연구가 제시되고 있다[7]. 요한슨(Johansson)은 사람의 동작이 여러 가지 동시에 이뤄지는 상태에서 사람의 형태와 윤곽로부터 생물학적 동작 정보를 분리하기 위한 점광 디스플레이를 고안하여 200msec의 아주 짧은 시간 동안 12개의 점광만으로 움직이는 인간의 형태를 충분히 인식할 수 있음을 밝혔다[8]. 또한 점광 디스플레이를 이용하여 각종 인간

의 동작 및 걸음걸이 패턴을 관측한 다음 해당 관측 데이터를 통해 뇌의 인식 패턴을 분석하여 생물학적 동작인식 매카니즘을 밝힌 많은 인지과학 및 패턴 인식 연구 사례가 있다[9][10].

신경생리학적 동작인식과정을 컴퓨터비전 기술과 융합한 사례로는 해리스(Harris)의 특징점 연산자 연구[11]가 있으며 이를 바탕으로 랩테브(Laptev)가 시공간 관심점(STIP: Spatio-temporal interest points)을 이용해 비디오에서 인간의 동작을 구조적으로 검출하는 연구[12]를 제시하여 요한슨(Johansson)의 생체 동작에 대한 점광 패턴 연구와 컴퓨터비전의 동작인식 접근 방식이 통합될 수 있음을 보여 주었다.

한편, 기즈와 포기오(Giese & Poggio)[13]는 인간을 비롯한 생물의 운동 인식 성능이 어떻게 발달되었는지에 대한 내용을 바탕으로 인간은 전체 신체, 입, 손의 운동, 얼굴표정을 선택적으로 담당하는 뉴런이 있어 동작과 움직임을 관찰할 때 뇌의 뉴런이 같이 응답하는 구조임을 밝히고 이를 컴퓨터모델로 설명하였다. 이와 같이 생물학과 신경생리학 연구를 통해 인간이 시각정보를 인지하고 처리하는 과정에 대한 다양한 결과들이 제시되고 있으며 인공지능, 컴퓨터비전 분야 역시 인간을 닮은 컴퓨터를 만들기 위해 지속적으로 생체 기반 정보처리 연구를 제안하고 있다.

컴퓨터비전에서는 다양한 알고리즘들을 바탕으로 비디오에서 동작 특징을 인식하고 패턴을 분류한다[14][15][16]. 특징 추출의 다양한 기술 때문에 일반화하기는 어렵지만 동작 특징으로는 신체 형태가 가장 많이 이용되며 픽셀의 강도, 경계선, 외형선, 색상과 이 모든 특징을 조합한 결과 역시 특징으로 이용한다[17]. 추출된 특징은 발생빈도를 확률로 계산하는 PCA, ICA, SOM, VQ 등의 처리 알고리즘을 통해 가장 연관성 높은 공간을 정의하고 이를 바탕으로 HMM, SVM 등의 특징분류 패턴기를 통해 동작을 분류하는 방법이 일반적이다.

하지만 시쿨러(Sekuler)[18]의 연구와 같이 동작 검출과 분석을 신경작용의 연관관계로 보고 시각장의 한정된 영역 내에서 부분 동작 신호를 검출한 다음, 대상 동작의 방향과 속도를 전역적으로 통합하여 대상의 동작을 판별하는 생물학적 동작인식 방법을 컴퓨터비전의 동작인식 과정에 적용하면 기존 컴퓨터비전 방식의 동작인식에서 간과하고 있는 동작과 시간 정합성 간의 상관관계, 이미지시퀀스가 모두 제시된 후에 동작인식 결과를 확인하는 대신 빨리 동작인식 결과를 확인할 수 있는 등 기존 동작인식의 문제점을 지능형으로 개선할 수 있다.

본 연구는 동작 특징 추출과 동작 패턴 인식에 대해 인간의 시각정보처리 기능에 기반한 새로운 인식 방법을 제시하기 위해 기존 컴퓨터비전 동작인식 기능에 생체 시각인지 기능 특성을 반영하고 인간의 인지기능 한계를 개선한 시각인지 기능을 추가한 동작인식을 시도하였다. 동작인식 대상이 담긴 이미지시퀀스가 모두 제시되기 전에 미리 인식 분류 결과를 도출하며 입력 시퀀스의 시간변화와 속도변화에 대응해서 우수한 동작인식결과를 보임을 확인하도록 실험에

서 테스트 비디오를 보고 동작을 판별할 때 비디오 시퀀스를 세분화한 후, 이를 시각적 자극으로 전달하고 처리하는 방법을 설정하여 테스트하였다. 사용자 행동을 이용한 감정과 의도 인식을 수행한 사전연구를 토대[19][20][21]로 하였으며, 동작인식 과정을 수행하기 위해서는 동작인식 대상의 초기화와 추적과정이 필요하지만 대부분 인식 대상이 비디오 시퀀스에서 이미 설정되어 있고 시퀀스 장면 내에 피해야 할 방해요소나 해석해야 할 컨텍스트가 없으므로 연구 범위에서 초기화와 추적 과정은 제외하였다.

인간을 닮은 컴퓨터 에이전트를 구현하고자 인간의 시각 인지 과정과 이의 동작원리를 컴퓨터비전, 인공지능, 동작인식기술의 컴퓨팅 모델로 구조화하여 응용하려는 아이디어의 확장은 필연적이며, 본 연구 역시 그러한 연구과정의 일환으로 생체 기반의 컴퓨터 에이전트 연구범위 확대에 기여한다.

3. 동작인식 구조

제안 방법은 사람이 생체 동작을 인식하는 과정을 기반으로 동작인식을 처리하는 것으로 생체 시각정보처리 구조를 컴퓨터비전의 동작 패턴 분류에 응용하여 신경장(neural field) 기반의 동작인식 과정을 설정하고 컴퓨터비전에서 적용해 온 사전 동작 모델 없이 동작을 인식하는 방법으로 구성하였다.

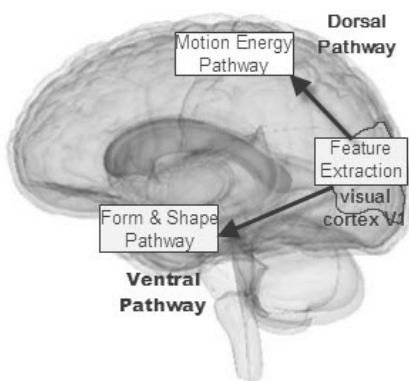
3.1 생체 시각정보처리 과정

생체 시각정보처리 과정은 Fig. 1(a)와 같이 뇌 회색피질의 각 영역이 해당되는 정보처리를 담당한다. 망막에 맺힌 이미지의 부분 동작은 전자 신호로 바뀐 후 배외측 슬상핵(LGN: lateral geniculate nucleus)에 전달되고 전체 이미지보다 작게 설정된 원도형태의 감수영역이 있는 시각정보처리의 중심영역인 V1에서 처리된다. V1에서 추출된 뉴런을 통해 시각자극이 특정 방향을 향해 움직이는지를 확인하고, 측면 상호작용을 통해 첫 번째로 동작 검출과 예측을 시도

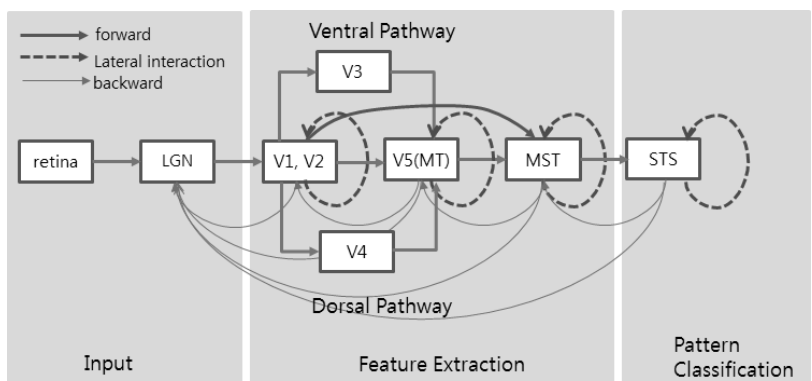
한다. 5차 시각영역인 V5(MT)는 부분 운동을 감지하고 광류를 담당하는 MST 영역에서 동작 특징을 추출한다. 특징이 추출되면 운동에 반응하는 상측 두구(STS: superior temporal sulcus)/방추모양 얼굴 영역(FFA: fusiform face area)에서 운동 패턴을 분류한다.

하지만 동작의 시각적 인지가 온전히 신경계 감수영역의 부분 응답으로만 결정되는 것은 아니고 Fig. 1(b)와 같이 V1에서 수집한 정보 외 MT 영역에서 수집한 내용을 이용하여 속도 정보를 확인하고 개별적 뉴런 응답의 불확실성을 해결하기 위해 신호를 조합한다. 그래서 시각처리를 담당하는 각 영역에 순방향으로 정보가 전달되고 해당 영역에서 측면 상호작용을 통해 정보를 모은 다음 역방향으로 시각 인지 처리 결과를 확인한다. 인식 결과 뇌의 시각처리과정에서 도출된 생체 인식 특성은 다음과 같다.

- 신호 강건성 : 시간적으로 신호의 양이 현저히 줄어도 점광(Point-Light)과 같은 단순한 자극만으로 시각적 패턴 분류가 잘 이루어지므로 점광과 같은 몇 개의 관심점 표시로 여러 동작을 구분할 수 있다[13][14].
- 시점 종속성 : 시각적 패턴 인식은 시점 각도에 달려 있다. 제시된 패턴이 회전변환하여 패턴 방향이 바뀔 경우 인식 성능은 낮아지지만 경험을 통해 학습하면 회전변환 때문에 저하된 인식성능은 개선할 수 있다.
- 시간 민감성 : 시퀀스를 분류하는 뇌의 시각적 시퀀스 패턴 매칭은 시간의 상관관계에 아주 민감해서 시간정합성을 고려해야 한다[6].
- 처리 경로 : 시각 정보가 전달될 때 LGN을 통해 V1, V2를 거치는 신호는 두 경로로 나뉘어 처리한다. 운동과 깊이 자극은 배측 경로(dorsal pathway)를 통해 V5(MT)로 전달하는 반면, 자세, 얼굴 표정, 손 자세와 같은 형태와 색깔자극은 정적 특징으로 복측 경로(ventral pathway)를 통해 V4로 전달한다.



(a) Visual processing areas in brain[22]



(b) Biological vision processing structure

Fig. 1. Action recognition structure in brain cortex

Fig. 1의 (b)는 컴퓨터비전의 특징추출과 패턴분류를 생체 시각자극 처리 과정에 대응하여 설명한 것으로 컴퓨터비전의 동작 검출과 특징 추출 단계를 뇌에서는 V1 시각피질영역과 V5(MT)에서 담당하여 운동의 방향과 움직임 특징을 추출한다. 패턴 분류는 생물학적 운동 패턴을 분류하는 STS와 FFA 영역에서 수행한다. 뇌의 배측 경로는 운동자극과 깊이 자극만을 담당하므로 자극은 광류신호로 전달된다. 이러한 생체 시각정보 인식 과정은 입력 자극의 개별 특징을 식별한 다음 전체 특징을 조합하여 대상을 인지하고 장면에서 대상의 컨텍스트를 확인하는 것으로 정리할 수 있다.

제안 동작인식 구조는 생체 시각인식과정에서 영향을 받은 다음 세 가지 아이디어를 적용하였다. 먼저 컴퓨터비전 동작인식 과정에 동작인식을 담당하는 주요 시각 피질 계층의 기능적 속성을 모방하였다. 두 번째, 패턴 인식 단계에서 피질계층 활성화를 모방하여 모델링하였다. 세 번째, 제안 동작인식 모델의 평가를 생물학적 동작 실험데이터로 대상으로 수행하였다. 이를 위해 사람의 걷기 동작을 단일 궤적으로 인식하는 실험을 수행하고 복잡한 시공간적 시퀀스로 다양하게 구성하여 여러 상황의 동작으로 통합 인식하였다. 이는 사람이 생체 동작을 인식할 때 부분적으로 동작으로 인식한 후 부분을 통합하여 전체 동작으로 이해하고 인식하는 과정과 동일하게 하였다. 인식된 결과는 복잡한 팔 동작 시퀀스에 응용하여 인식 효과를 검증하는 내용으로 구성하였다.

3.2 모델 구축 관련 전제사항

컴퓨터비전 패턴 인식 기술에 생체 시각정보처리를 담당하는 뉴런 분포 개념을 반영하려면 시퀀스 분류 모델 구축에 몇 가지 특성을 전제해야 한다. 먼저 인식 영역의 일부분에 특이성을 보이는 감수영역(receptive field) 단위로 코딩하며 코딩은 2D로 이루어진다. 비디오 시퀀스는 시간흐름을 따르는 것이 아니라 특정 시점에 종속적으로 처리해야 한다. 또한 복잡한 시퀀스를 정확하게 모델링하는 문제는 인식과정에서 발화되는 샘플링 원도 역할의 감수영역의 크기가 아니라 시스템이 동작을 분류하는 가중치 커널함수 ' $w(\cdot)$ '를 어떻게 정의하고 선택하느냐에 달려 있다는 점을 고려해야 한다.

본 연구에 적용한 동작 분류 코딩 방법은 분산 매카니즘으로써 복잡한 시퀀스를 짧은 여러 개의 시공간적 궤적 집합으로 나누어 진행하는 방식이다. 즉, 개별 특징을 식별한 다음 식별 특징을 조합하여 전체적으로 대상을 인지하는 방식이다. 이 방식으로 모델링한 결과는 4장에서 짧게 나눈 몇 개의 궤적을 대상으로 테스트하여 전체 동작을 판별하여 인식성능을 나타낸다.

생체 시각정보처리에서는 비디오 시퀀스를 스냅샷으로 코딩하는지 암묵적 템플릿으로 코딩하는지 불분명하다. 또한 움직임 처리 특징에 따라 배측과 복측으로 처리경로가 분리

되어 있다. 하지만 컴퓨터비전에서는 명확한 코딩과정이 제시되어야 하고 동작과 깊이 처리를 하는 배측 경로 처리와 정적 특징을 처리하는 복측 경로의 동작자세 뉴런상태를 같이 고려해야 한다. 만약 모든 자세를 표현하는 템플릿이 있으면 각 동작의 스냅샷 당 개별 감수영역을 할당하고 새로운 인식수행 과정에서는 동작 자세별로 인식을 처리해 생체 시각정보를 처리하는 뇌 영역의 각 과정을 쉽게 설명할 수 있으나 실제로 모든 동작 자세를 개별 스냅샷으로 설정하면 이에 따른 비용소요가 커진다. 반대로 동일한 신경 맵으로 여러 가지 동작 패턴을 설명하는 방법을 통해 작은 뉴런 공간으로 동작 전체를 설명할 수 있으나 이 방법으로는 동작 식별이 아주 어려워진다[23]. 이런 점을 고려하여 본 연구에서는 동작을 결정하는 신체 부위별로 부분 패턴을 검출하고 이를 해당 감수영역에 표시하는 방법으로 감수영역 수를 줄인다.

3.3 동작 특징 추출

신경생물학 구조상 뇌에서는 비디오 시퀀스 인식 단계에서 부분 동작 정보만으로 패턴 분류를 충분히 수행해내며 코딩 또한 2D 형태로 처리한다. 또한 동작 패턴을 판별하는 인간의 능력은 다양한 속도변화에도 잘 대처하며 모든 패턴이 제시되기 전에 이미 패턴을 파악할 수 있다. 다만 부분적이라도 대상이 회전 변환 상태라면 분류를 원활하게 처리하지 못하고 시간변화에 아주 민감하다. 이와 같은 생체 패턴 인식 특성을 고려한 결과, 패턴 분류과정에서 제시시간에 고려해야 할 특징들은 몇 개의 연관 궤적에만 한정되어 있기 때문에 임계시간 내에 비디오 시퀀스에서 특징 점의 위치를 포착할 수 있는지 감안한 후 동작 궤적을 달리하여 분석할 필요가 있다.

뇌에서 처리하는 패턴 분류과정을 반영하여 제안 모델에서는 비디오 시퀀스가 입력되면 먼저 전체 동작 영역을 검출한 다음 부분 동작패턴을 검출한다. 검출한 패턴은 패턴 분류기가 분류하도록 해당 분류 층으로 전달한다. 제안 모델구조에서는 부분 패턴 검출이 가능함을 가정하고 패턴 분류를 수행했다. 한 동작에서 동작특징들의 서로 다른 이동 궤적을 확인하기 위해서는 Continuum Neural Field Theory (CNFT)[24]를 이용한 계산 모델을 적용하여 시각 신호 입력 값이 2D 뉴런 형태의 해당 감수영역에 매핑하여 분포되도록 한다. 이를 위해, 이미지시퀀스에서 보행자의 동작 특징과 주요 형태를 추출하기 위해 이미지에 점광 자극형태의 관심점 특징을 표시하고 감수영역 크기를 설정한 다음 이미지 시퀀스의 전체 프레임에 대해 감수영역을 겹쳐서 추출한 광류의 방향을 휘도 분포로 설정하고 여기에 PCA를 적용하여 다차원 데이터공간에서 해당 특징을 추출한다. 추출한 형태 특징은 휘도 값에서 얻는 공분산행렬의 가장 큰 기저 값에 해당하는 기저벡터로 정의한다. 주요 기저벡터는 다차원 특징 공간에 휘도 값으로 주어진 최대변화량 특징으로서,

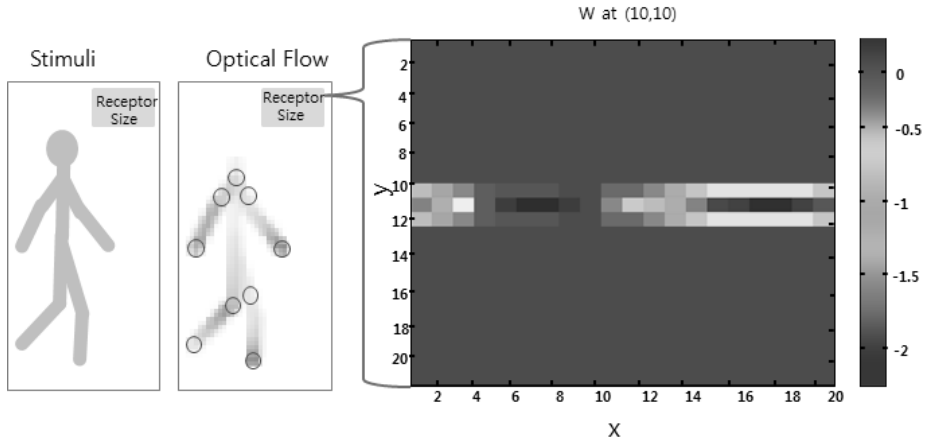


Fig. 2. Kernel function w at $p=(10,10)$ for possible locations χ

감수영역 역할을 하는 샘플링 윈도우에 휘도성 분포를 색상 코드로 나타내어 시각화한다.

Fig. 2는 이와 같은 동작 특징 추출과정을 나타낸 것으로, 단일 프레임의 신체 동작 특징을 결정하는 뉴런 활성화 정도를 식 (2)의 $w(\chi, p)$ 커널 함수로 표시한다. 이때, 주요 동작 특징 계산에 사용할 감수영역의 크기를 정하고 입력된 동작의 광류는 비디오의 하위 시퀀스 프레임을 통해 계산한다. 생물학적 처리와 마찬가지로 중간단계에서 동작을 결정하는 신체 특징 부위를 설정하고 휘도 분포에 PCA를 적용하여 추출한 동작 특징을 감수영역 형태의 샘플링 윈도우에 나타낸다.

3.4 동작 패턴 분류

시각 신호의 패턴을 분류하기 위해 2D 뉴런 맵으로 구성된 분류시스템을 구축하여 각 맵을 망막위상 형태로 조직화하고 뉴런의 활성화는 부분 입력과 다른 뉴런에서 받은 자극과 억제에 따라 이루어지도록 한다. 이런 상호작용 상태는 식 (1)을 통해 감수영역별로 동작의 부분 패턴이 나타나는지 활성화 여부를 모델링하여 나타낸다. 인식하려는 각 동작 패턴에는 한 개씩의 뉴런 맵 m 을 할당한다[21].

$$\frac{\partial m(\chi, t)}{\partial t} + \tau m(\chi, t) = \left[\int_0^{x_f} w(\chi, \chi) m(\dot{\chi}, t) d\dot{\chi} + l(\chi, t) \right]^+ \quad (1)$$

식 (1)에서 $m(\chi, t)$ 는 시간 t 에 맵 중심위치 χ 에서 뉴런의 활성화를 나타내는 스칼라 양이다. $w()$ 는 시스템이 입력 자극을 어떤 동작으로 선택분류 할지를 결정하고 $[]^+$ 는 0을 최대치로 나타낸다. \int 함수는 입력과 개별 뉴런 전체의 활성화 비율 간의 관계를 비선형적으로 나타낸다. τ 는 역학의 시간 상수이다. $w(\chi)$ 커널함수는 감수영역 측면 연결의 억제와 활성화 정도를 나타낸다.

커널함수를 사용하면 차원이 서로 다른 경우에도 2차원

계산만으로 매핑함수 값을 구할 수 있기 때문에 본 연구에서는 가우시안 커널함수를 이용한다. 동작 패턴에서 단순 궤적을 포착하기 위해 궤적 축을 따라 가우시안 커널함수와 주기함수로 나타낸다.

$$w(\chi, p) = \alpha \exp\left(-\frac{(y-p_y)^2}{2\sigma}\right) (J_0 + J_1 \cos(2\pi(p_x - x)/l - \beta)) \quad (2)$$

식 (2)에서 β 는 비대칭성, σ 는 커널의 공간 크기, l 은 궤적 경로의 전체 길이를 나타내는 주요 파라미터들이다. J_0, J_1 은 단위의 y 축 값이다. 현재 단위위치 p 에서 커널함수 $w()$ 는 χ 위치에 대한 연결강도를 결정한다. Fig. 3을 보면 실제 동작 패턴 궤적에서 멀리 있는 커널함수 $w()$ 값은 연결강도 가중치가 없어 거의 0에 가깝다는 것을 알 수 있다. 또한 커널의 위치 p 가 결정되기 전에 활성화되어야 할 뉴런에만 양의 값을 할당하여 제안 동작인식 모델이 입력 순서를 확인하도록 한다. 각 패턴의 최종 동작을 결정하기 위해 식 (3)을 통해 일정 시간 동안의 평균값을 계산한다[4]. 뉴런의 상태 방정식 결과를 현재 뉴런 맵 m 보다 나중에 나타내기 위해 활성화 감쇄는 τ 보다 더 작게 $\tau/2$ 으로 하였다.

$$\frac{\partial s(t)}{\partial t} + \tau s(t)/2 = \int m(\dot{X}, t) d\dot{X} \quad (3)$$

4. 실험 및 결과

제안 연구가 생체 시각정보처리 기능을 모방하여 동작 패턴을 검출하고 분류하였기 때문에 신경과학에서 수행하는 실험내용을 확장하여 적용하였다. 실험에서는 걷기 동작을 단일 직선궤적을 통해 모델 테스트를 하고 모델링 한 내용은 팔 동작으로 시뮬레이션하였다. 단일 궤적은 복잡한 시퀀스를 해체하는 데 유용하기도 하지만 가장 간단한 패턴이므로 동작의 직선 궤적 시뮬레이션의 목적인 여러 변수를

제어하면서 서로 다른 방향의 궤적을 판별하기 위함이다. 팔의 구부러짐을 펴는 동작을 대상으로 동작인식을 테스트한 것은 모델이 다른 속도로 여러 개의 동작 궤적을 동시에 처리할 수 있음을 설명하고 실제 이미지에 가까운 궤적을 시나리오 상에서 판별할 수 있음을 테스트하기 위해서이다.

4.1 입력 데이터

동작 패턴을 분류하기 위해 왼쪽에서 오른쪽으로 걷는 비디오 시퀀스와 오른쪽에서 왼쪽으로 걷는 비디오 시퀀스를 입력 데이터로 이용한다. 이와 같은 단순한 동작은 신경생리학과 실험심리학[12]에서 적용하는 실험 패러다임으로, 비디오 시퀀스에 관심점을 특징으로 표시하여 왼쪽과 오른쪽을 향해 걷는 동작을 구분한다. 실험에 적용한 입력 값은 식 (4)와 같이 정의한다.

$$I(\chi, t) = \exp\left(-\frac{(x-ut)^2 + (y-y_0)^2}{2\sigma^2}\right) + N(0, \Sigma) \quad (4)$$

$I(\chi, t)$ 에서 χ 는 (x, y) 좌표 위치, v 는 입력속도, y_0 는 y 축에서의 위치, σ 는 입력 크기를 의미한다. 이 이미지 시퀀스 입력에 Σ (잡음 수준)와 y_0 (y 축 위치), v (입력 속도)를 달리 할 때의 변화정도를 50회의 실험을 통해 분석하고 평균 0, 분산 Σ 의 가우시안 잡음을 추가하여 잡음상태에서의 동작인식 강건성을 확인한다. 식 (2)의 커널함수 $w()$ 의 매개 변수 $v=5, \Sigma=0.005, y_0=l/2, \tau=0.15, J_0=-9.8, J_1=-13.5, \beta=2, \sigma=0.001, l=20$ 로 설정한다. 식 (1)의 적분 시뮬레이션은 두 프레임 간을 최소한 10회 반복하여 $dt=0.1$ 인 4차 룽게-쿠타(Runge-Kutta)방법에 따라 수행하였고 사인 곡선 파라미터의 확장분석은 [24]를 따랐다. Fig. 3은 입력 속도별 시퀀스의 커널 예이다.

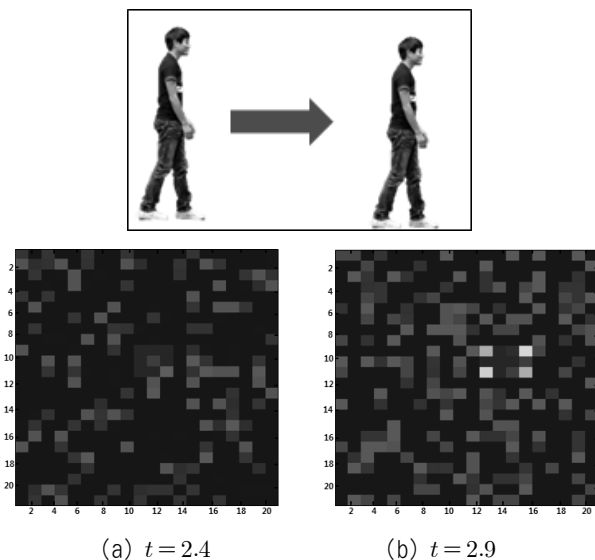


Fig. 3. Input Sequence Example (noise $\Sigma=0.005$)

4.2 동작의 단일 직선궤적 대상 실험

1) 시간정합성(Time Warping) 테스트

생물학적 시각 자극처리에서는 시간길이가 다른 즉 움직이는 속도가 다른 연속 동작도 정확하게 분류한다. 따라서 속도가 다른 두 시퀀스 사이의 유사성을 측정한 실험을 통해 제안 모델 역시 우수한 시간정합기능을 제공하는지 확인할 필요가 있다. 시간정합은 패턴인식에서 DTW(Dynamic Time Warping) 알고리즘을 통해 시간 길이가 다른 시퀀스를 효과적으로 인식하는 방법이다.

제안 모델의 시간 정합성을 확인하기 위해 입력 시퀀스의 시간 길이를 달리 하여 원래 입력 시퀀스의 특정 부분을 느리게 입력하여도 모델이 동일하게 동작 패턴을 잘 찾을 수 있는지를 테스트하였다. 이 테스트에서는 식 (4)의 입력 속도 v 를 다양화해 입력시퀀스의 속도 v 를 $[0,100]$ 범위 내에서 $-v$ 로 달리 설정하여 실험하였다. 모델이 입력 시간이 다른 입력 시퀀스 패턴을 동일하게 잘 판별하면 시간 정합성이 우수하다고 할 수 있다. 잡음의 영향을 피하기 위해 50회를 테스트해서 결과 값을 평균화했다.

Fig. 4의 결과를 보면 제안 패턴분류 방식이 한 개의 궤적을 대상으로, 동일한 시퀀스의 다양한 입력 속도 변화에도 잘 인식하였다. 모델의 속도는 커널함수에서 $v=5$ 로 설정하였다. 입력 속도가 증가하면 v 와 $-v$ 간의 절대 값 차이도 역시 커져 모델이 쉽게 구별할 수 있기 때문에 제안 모델이 v 와 $-v$ 두 가지 속도를 구별하여 속도의 변화량에 상관없이 분류를 잘 해냄을 설명할 수 있다.

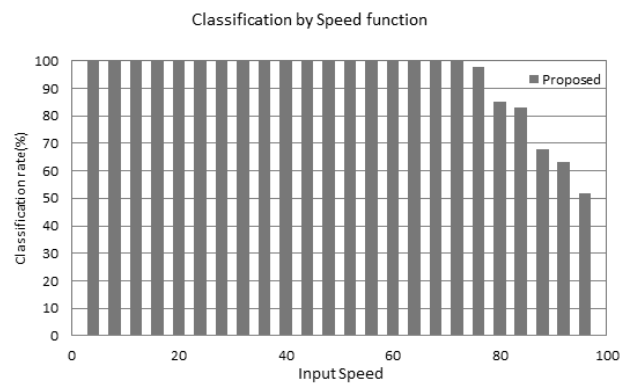


Fig. 4. Classification performance by speed (50 trials, noise $\Sigma=0.005$)

2) 시간 변화에 따른 분류 반응 테스트

이 실험을 통해서도 속도가 변화하는 상태에서 동작을 잘 분류하는지 확인한다. 생물학적 시각 자극처리에서는 연속적인 단계별 패턴 분류 뿐만아니라 순간적인 패턴 분류 역시 정확하게 처리하기 때문에 제안 모델 또한 시퀀스에서 순간적으로 패턴을 분류하는지를 설명할 필요가 있다. 시퀀

스는 정해진 속도 v 에 인식하도록 설정하고 v 또는 $-v$ 속도로 정확하게 시퀀스를 입력하여 모델이 특정 속도에 입력 시퀀스를 잘 판별하면 잘 분류한다고 할 수 있다. 잡음 허용성 실험을 통해 잡음이 증가하면 성능이 떨어짐을 알기 때문에 잡음 값을 $\Sigma \in \{0, 0.05, 0.1\}$ 로 달리하여 잡음 효과를 평균화시켰다.

실험 결과 Fig. 6과 같이 초기에는 제안 방식의 순간적 동작 분류 성능이 낮았지만 점차 안정적인 결과를 보였다. Fig. 3의 입력시퀀스 경계 문제를 피하면서 시퀀스 공간의 끝인 $x=20$ 에서 시작점인 $x=1$ 까지 주기적으로 입력 값을 움직였다. $t=40$ 과 $t=80$ 에 입력 값을 움직이기 전에 이미 분류 성능은 최고에 달했으며 잡음이 나타나는 변곡점 ($t=40, 80$)에서만 일시적으로 성능이 저하되었다.

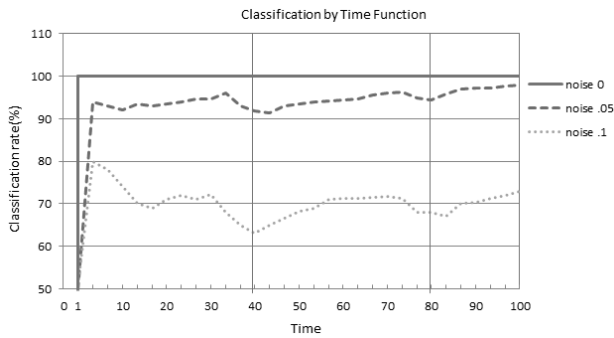


Fig. 5. Classification performance by time (50 trials, noise $\Sigma = 0, 0.05, 1$)

4.3 시공간적 상관관계(STC: Spatiotemporal correlation)와 비교 실험

제안 모델의 성능을 설명하기 위해 신경생리학의 동작 예측에서 적용하는 가장 일반적인 모델인 시공간 상관관계 매카니즘(STC) [25]와 비교하였다.

매시간 t 별로 완전한 입력 템플릿을 준비하고 여기에 식 (3)을 통해 시간 시퀀스 관계성을 부여하여 입력 템플릿 간에 시공간적 상관관계를 설정한 STC를 구축한다. STC는 동작 분류를 수행하는 가장 간단한 방법으로 공간적으로 잘 배치된 입력과 정확한 시간 순서(운동 방향)가 제시되면 높은 동작분류 값을 결과치로 나타내는 장점이 있다. STC는 식 (5)(6)으로 정리할 수 있다.

$$C(t) = \sum I(x,t)I(x,t) \tag{5}$$

$$\frac{\partial S(t)}{\partial t} + \tau S(t)/2 = C(t) \tag{6}$$

식 (5)에서 T 는 템플릿을 의미한다. STC는 제안 모델과 마찬가지로 식 (6)과 같이 시간이 갈수록 공간적 상관관계는 줄어들면서 순차성을 확인하는 매카니즘이다.

시공간 상관관계 방식은 매시간 t 별로 모든 시퀀스 스냅샷을 기록하여 이 스냅샷들의 시공간관계를 상관적으로 분석하여 패턴을 분류한다. 제안 모델처럼 부분 특징만을 이용해 패턴 분류를 시도하는 방식과 모든 시퀀스 스냅샷을 매순간 기록하는 방식을 비교하여 제안 방식의 효과를 설명함으로써 제안 구조의 단순성과 정확성을 확인할 수 있다.

1) 잡음 허용성 비교

식 (4)의 잡음 정도 Σ 를 다양화시켜 잡음 허용 정도를 테스트했다. 제안 모델과 STC 분류시스템이 속도 v 로 주어진 입력을 인식하도록 설정하고 v 또는 $-v$ 의 속도로 입력이 이루어지도록 했다. 모델이 두 입력 값을 판별하면 정확하게 분류한다고 할 수 있다. 50회의 시도를 통해 잡음효과를 평균화시켰다. 잡음이 낮으면 제안 모델과 STC가 동일한 분류 성능을 보였지만 잡음이 증가하면서 Fig. 6과 같이 $\Sigma=0.15$ 에서는 제안 모델의 성능이 50%에 도달할 정도로 낮아졌다. 잡음이 없을 때는 제안 모델의 인식률이 높으나 이미지 시퀀스의 잡음 수준이 높아지면 제안 모델의 성능이 나빠지는 이유는 제안 모델이 커널함수의 크기와 위치에 따라 동작 특징의 범위를 설정하는 데 잡음 때문에 동작 특징 위치를 제대로 포착하지 못해서이다. 인식률은 커널함수 $w()$ 값에 의해 결정되므로 커널 크기 σ 를 증가시키거나 커널함수 값이 동작 궤적 밖의 0 값이 아니라면 이런 종류의 잡음에 따른 인식률은 개선할 수 있다.

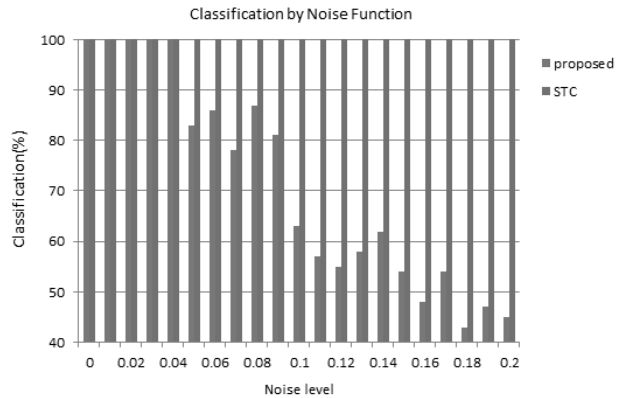


Fig. 6. Classification performance by noise levels (input speed $v = 5$)

2) 위치변화 강건성 비교

신경생물학적 시각처리 방법에서는 인식대상의 위치가 회전변환 될 경우 인식성능이 아주 나빠진다. 반면 컴퓨터비전의 동작 패턴인식 방법은 대상의 회전변화에 아주 강건하다. 제안 모델 역시 동작하는 대상의 위치가 달라져도 잘 인식하는지를 확인하기 위해 식 (4)에서 y_0 궤적에 수직인 축의 위치를 다양화시켜 다른 동작위치에서도 동일한 분류 성능을 보이는지 실험하였다. 이를 위해 제안 모델과 STC

분류 시스템이 이미지 시퀀스 y 축 위치인 y_0 변수 값을 변화시켜 서로 다른 위치에서 동작 궤적을 측정하였다. 이때 시퀀스 입력은 v 또는 $-v$ 의 속도로 이루어지도록 하였다. 모델이 위치를 달리한 다른 두 입력 시퀀스를 잘 판별하면 정확하게 분류한다고 할 수 있다. 50회의 시도를 통해 잡음 효과는 평균화시켰다. 그 결과 STC 분류 시스템은 공간 변화에 아주 민감하여 위치가 바뀌면 상관관계도 달라져 인식 성능이 급속히 떨어졌다. STC 분류 시스템은 신경생물학적 시각처리방식을 이용하였기 때문에 사람은 인식대상의 동작이 회전변환하거나 동작 부위 위치가 변경되면 동작을 제대로 인식하지 못함을 반영한 것이다. 제안 방식 역시 이와 유사한 속성을 보이며 Fig. 7과 같이 분류 성능이 떨어졌다. 하지만 제안 모델의 위치변화 강건성 문제는 커널함수 $w()$ 의 위치좌표를 재설정하여 해결할 수 있다. 커널의 크기 σ 값으로 공간 위치변화를 허용하기 때문이다. 다만 σ 를 변형하면 잡음에 대한 허용치도 같이 낮아진다.

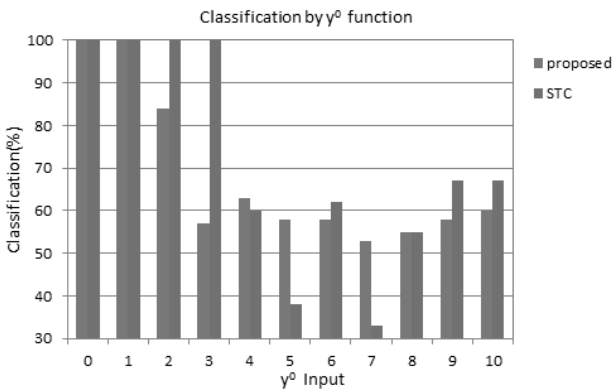


Fig. 7. Classification performance by y_0 trajectory (noise $\Sigma=0.005$, input speed $v=5$)

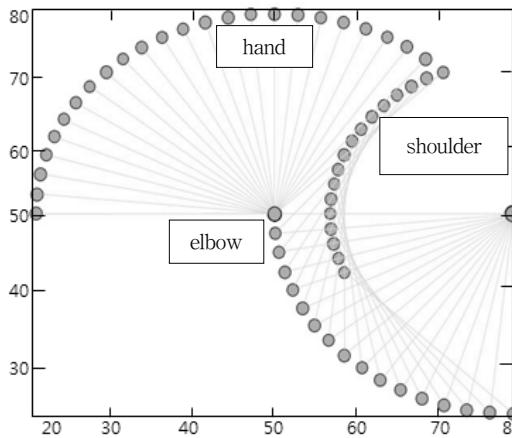
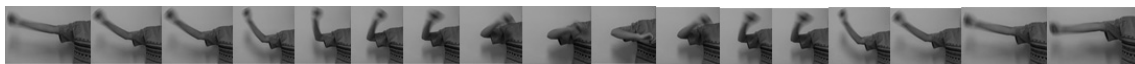
4.4 팔 동작 대상 실험

실제 동작을 대상으로 제안모델이 동작 패턴을 잘 분류하는지 확인하기 위해 Fig. 8과 같이 팔을 구부렸다 수평방향으로 펴는 동작 경로를 입력 값으로 한두 개의 m 뉴런 맵을 구축했다. 입력 시퀀스를 단순화하기 위해 4.1에 의거, 팔꿈치와 손의 궤적만을 이용했다. 같은 시간 동안 팔과 손 두 부위가 동작을 취하도록 했고, 식 (7)과 같이 w 의 정의를 변형하여 궤적을 다양하게 포착하도록 했다.

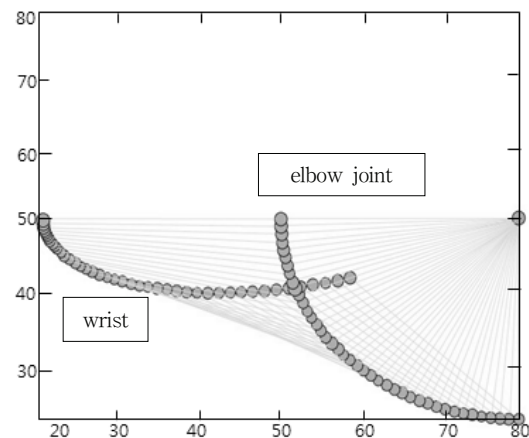
$$w(x, p, v) = \exp\left(\frac{y_1^2 + x_1^2}{2\sigma^2}\right)(J_0 + J_1 \cos(2\pi x_1/4\sigma - \beta)) \quad (7)$$

x_1 은 $\tan^{-1}(v_x/v_y)$ 에 의한 x 의 회전변환 결과 값이다. 이미지 입력 시퀀스는 4.1의 가우시안 부가 잡음을 적용한 시퀀스와 동일하다.

Fig. 8은 구부렸다 펴는 팔 동작 경로로 그림에서 원의 위치는 손, 어깨, 팔꿈치를 의미한다. 같은 위치에서 시작해서 동작을 마치는 팔의 관절동작을 구별하기 위해 팔꿈치와 손, 어깨 관절 동작을 테스트한 결과, 오른쪽 동작 즉, 수평방향으로 펴는 동작의 뉴런 분포 평균 s 가 더 많이 활성화된 것을 확인할 수 있다. 같은 팔 동작에서 왼쪽 그림의 '경로 A'는 팔꿈치 동작 경로이고 오른쪽 '경로 B'는 손목과 어깨 관절 경로이다. A와 B 경로를 입력 값으로 하여 시스템이 여러 잡음 정도에 따라 정확하게 팔꿈치 동작과 손목, 어깨 동작을 구분하는지 확인한 결과, 입력 값으로 A 경로를 이용하고 $\Sigma=0$ 일 때 정확하게 100%의 팔 동작인식 결과를 보였다. $\Sigma=0.005$ 로 잡음의 정도를 증가시키면 인식률은 90%로 낮아졌다. 분류 결과는 10회 시도하여 계산한 평균값이다. 이 결과로 제안 모델이 겹침과 곡선 궤적으로 유사하면서도 복잡한 서로 다른 시공간 패턴을 표현하고 분류할 수 있음을 알 수 있다.



(a) A path moves first elbow (path A)



(b) A path extended both joints (path B)

Fig. 8. Arm action from flexion into the horizontal extension position

5. 결 론

본 연구에서는 생체 시각정보처리 매카니즘을 바탕으로 비디오 시퀀스 내 동작을 효과적으로 인식하는 생체 기반 동작인식 모델을 제안하였다. 생체 시각자극 처리 매카니즘의 속성을 반영한 실험 결과, 제안 모델의 패턴 분류 방식이 일반적인 생체 시각처리 기능보다 시공간적 시퀀스를 잘 분류함을 보여주었다. 특히, 시간길이가 다른 연속 동작도 정확하게 인식하는 시간 정합성 및 연속적인 패턴 인식에서 우수한 결과를 보였다. 또한 순간적인 패턴을 정확하게 인식하여 분류하는 기능 외 잡음 허용 정도와 동작의 위치가 회전변환하여 달라져도 잘 인식하는지에 대한 위치변화 강건성 실험을 수행하여 제안 모델이 생물학적 시각자극 처리 방식의 장점은 물론 위치변화에도 강건하게 동작 패턴을 인식함을 확인하였다. 더불어 제안 모델은 점광 자극과 같은 관심점 표시만으로 매순간의 명시적 스냅샷 없이 동작을 인식하였다.

제안 모델을 가장 기본적인 시공간 상관관계 구조인 STC와 비교했을 때 잡음이 증가하면 분류 성능과 위치변화 강건성이 저하되는 등 시공간 상관관계와 유사한 공간적 속성을 나타냈지만 커널 크기를 조절하여 이 문제를 해결할 수 있다.

결과적으로 제안 방식은 시간 정합성, 온라인 분류와 같이 인간의 뇌가 시각적 패턴을 분류 처리하는 방법과 유사한 기능과 속성을 보였으며 잡음의 유무 또는 궤적 간의 부분적 유사성에 상관없이 다수의 궤적을 처리하여 잘 분류함을 알 수 있다.

연구 내용을 동작 이미지시퀀스를 인식하는 최근 컴퓨터 비전 기술과 비교했을 때 컴퓨터비전에서 수행하는 방식인 대상의 특징 추출과 분류를 위해 신체 동작 모델을 선 구축하여 적용하는 것과 본 연구 내용처럼 단순하게 부분적이고 정적인 동작을 이용하여 특징을 추출하고 분류하는 것과는 아직 많은 차이가 있어서 향후 해결해야 할 것들이 많고 기술적 격차도 크다. 하지만 생체 기반 시각처리 동작인식 기술은 기존의 기계적인 패턴 인식 방식에 인간의 생체 시각정보처리 속성을 도입함으로써 컴퓨터비전을 이용한 동작인식기능에서 생체 기능과 유사한 상호 작용성을 도모할 수 있으며 지능적인 동작인식 기술 발전에 기여할 것으로 기대한다.

동작인식에 대한 본 연구는 지능형 로봇, HCI 등 생체 기반 시각처리 시스템 구현에 적용하여 효용 가치를 높일 수 있다. 제안 모델을 실제 다양한 동작 비디오 시퀀스를 대상으로 테스트하고 인식성능을 높이는 작업이 향후 계속 해야 할 과제이며 신체 전체가 표현하는 동작 신호에서 어떤 특징을 선택하고 이를 어떻게 효과적으로 코딩할 것인가는 계속 연구해야 할 내용이다.

Reference

- [1] Falk Fleischer, Antonino Casile, Martin A. Giese, "Bio-inspired approach for the recognition of goal-directed hand actions," International Conference on Computer Analysis of Images and Patterns, pp.714-722, 2009.
- [2] Byoung-Jun PARK, Kyu-Sang SHIN, Myung-Ae CHUNG, Sang-Hyeob KIM, "Bio-mimetic Light Sensor Based Cognitive System Technology", Journal of the Korean Physical Society, Vol.19, No.9. pp.2-8, 2010.
- [3] M. Lagarde, P. Andry, P. Gaussier, "Distributed real time neural networks in interactive complex systems. "International conference on Soft Computing as transdisciplinary science and technology, pp.95-101, 2008.
- [4] W. A. Fellez, J. G. Talyor, "Establishing retinotopy by lateral-inhibition type homogeneous neural fields," NeuroComputing, Vol.48, pp.313-322, 2002.
- [5] E. Mingolla, "Neural models of motion integration and segmentation," Neural Networks, Vol.16, pp.939-945, 2003.
- [6] R. Blake, M. Shiffrar, "Perception of human motion," Annu. Rev. Psychol., Vol.58, pp.47-73, 2007.
- [7] R. A. Brooks, "A Robot that walks: emergent behaviors from a carefully evolved network," Neural Computation, Vol.1, No.2, pp.253-262, 1989.
- [8] G. Johansson, "Visual perception of biological motion and a model for its analysis," Perception and Psychophysics, Vol.14, No.2, pp.201-211, 1973.
- [9] C. Curio, M. Giese, "Combing view-based and model-based tracking of articulated human movements," IEEE Workshop on Motion and Video Computing, Vol.2, 2005, pp.261-268.
- [10] K. Schindler, L. van Gool, "Action snippets: How many frames does human action recognition require?," Computer Vision and Pattern Recognition, pp.1-8, 2008.
- [11] C. Harris, M. Stephens, "A combined corner and edge detector," Albey vision conference, 1988, pp.147-152.
- [12] I. Laptev, "On space-time interest points", International Journal of Computer Vision, Vol.64, No.2, pp.107-123, 2005.
- [13] M. A. Giese, T. Poggio, "Neural mechanisms for the recognition of biological movements," Nat. Rev. Neurosci., Vol.4, pp.179-192, 2003.
- [14] A. Casile, M. A. Giese, "Critical features for the recognition of biological motion," Journal of Vision, Vol.5, No.4, pp. 348-360, 2005.
- [15] A. Puce, D. Perrett, "Electrophysiology and brain imaging of biological motion," Philosophical Transactions of the Royal Society of London, Series B. Vol.358, pp.435-445, 2003.
- [16] K. Fugushima, S. Miyake, T. Ito, "Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition," IEEE Trans. on System. Man and Cybernetics, Vol.13, No.5, pp.826-834, 1993.

- [17] T. B. Moeslund, E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, Vol.81, No.3, pp.231-268, 2001.
- [18] R. Sekuler, S. N. J. Watamaniuk, R. Blake, "Motion Perception," *Steven's Handbook of Experimental Psychology*, Vol.1, pp.121-176, 1998.
- [19] JinOk Kim, "Effective Pose-based Approach with Pose Estimation for Emotional Action Recognition," *The KIPS Transactions: Part B*, Vol.2, No.3, pp.1-10, 2013.
- [20] JinOk Kim, "Agent's Activities based Intention Recognition Computing," *The KSII Transactions*, Vol.13, No.2, pp.87-98, 2012.
- [21] JinOk Kim, "A Study on Visual Perception based Emotion Recognition using Body-Activity Posture," *The KIPS Transactions, Part B*, Vol.18, No.5, pp.305-314, 2011.
- [22] http://www.helpautismnow.com/Understanding_Children_With_Autism_PP.pdf
- [23] N. Rougier, J. Vitay, "Emergence of Attention within a Neural Population," *Neural Networks*, Vol.19, pp.573-581, 2006.
- [24] X. Xie, M. A. Giese, "Nonlinear dynamics of direction-selective recurrent neural media," *Physical Review E-statistical, nonlinear and soft matter physics*, Vol.65, No.1, pp.1-11, 2002.
- [25] I. Ahmad, W Zheng, J. Luo, M. Liou, "A fast adaptive motion estimation algorithm," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol.16, No.3, pp.420-438, 2006.



김진옥

e-mail : bit@dhu.ac.kr

1989년 성균관대학교(학사)

1998년 성균관대학교 정보통신공학과(석사)

2002년 성균관대학교 전기전자 및 컴퓨터 공학과(박사)

2004년~현재 대구한의대학교 모바일콘텐츠학부 교수

관심분야 : 멀티미디어공학, 패턴인식, 영상처리, HCI