

# 퍼지 집합과 신경망을 이용한 동작 인식

## (Motion Recognition using Fuzzy Sets and Neural-nets)

남 성 현 \* 김 광 용 \*\* 김 현 속 \*\*\* 황 중 선 \*\*\*\* 양 영 규 \*\*\*\*\*  
(Sung Hyun Nam)(Kwang Yong Kim)(Hyun Sook Kim)(Chong Sun Hwang)(Young Kyu Yang)

**요약** 비디오 영상에서 자연스럽게 움직이는 객체에 대한 인식방법을 제안한다. 이 객체는 자유자재로 움직이며, 움직임의 방향을 예측하기 어렵고 빠르게 움직인다. 객체의 동작은 그 객체가 고유하게 가지고 있는 움직임의 특성과 객체의 주변상황에 밀접한 관계가 있다. 본 논문은 자연스럽게 동적인 상황에서 객체의 동작을 분석하고 인식하는 방법을 제안한다. 객체의 동작을 인식하기 위해 동작을 구분하고 분석하여 동작의 대표특징을 추출한다. 그리고, 하나의 객체 주위에 다가오는 다른 객체와의 상관관계를 표현한다. 또한, 동작인식을 위한 알고리즘을 제안한다. 즉, 구간별 포아송(Poisson) 난-제로(Non-zero) 확률 분포를 통한 방법과 퍼지와 신경망의 결합에 의한 인식방법을 제안한다. 그리고 제안된 방법들로 구현된 인식결과를 비교 분석하고 평가한다. 우리의 알고리즘은 축구영역에서 시험된다. 우리는 동작인식의 성공뿐 아니라 실패의 경우도 설명하며 미래에 연구할 과제를 제안한다.

**Abstract** We propose approaches to moving objects in a natural and dynamic video scene. The objects are non-rigid and difficult to presuppose of moving direction. Their motion is erratic and they change shape rapidly between frames.

In this paper, We extract representative features for recognition of moving object and describe correlation between objects. And we propose approaches to recognition, i. e. using Poisson non-zero probability distribution in a section and fuzzy sets with neural-nets.

Also we test, analyze, and evaluate proposed approaches. Our algorithm is tested on soccer games. We include successfully recognition results as well as a few example, of the algorithm fails. Finally, we suggest some future extensions.

### 1. 서 론

데이터 저장장치, 이미지처리, 통신과 데이터 압축 기술이 빠르게 발전함에 따라 비디오는 정보시스템의 새로운 세대에서 중요한 요소가 되었다. 비디오 데이터는 시간적, 공간적 정보를 내포하고 있으며, 이들은 텍스트나 그래픽 또는 이미지가 주는 정보보다 많은 정보를

제공한다. 비디오 데이터는 위치, 거리등 시간 공간적 관계가 암시적으로 포함되어있다. 오락, 비주얼(Visual) 통신, 멀티미디어, 교육, 의료, 동작예측, 과학적 조사, 그리고 스포츠 같은 영역에서 비디오 시퀀스(Sequence)의 사용은 텔레비전과 VTR의 사용이 점점 그 폭을 넓혀감에 따라 지속적으로 증가한다. 축구는 지구상에서 가장 널리 알려진 스포츠 중의 하나이다. 특히, 한국에서는 2002년 월드컵 축구를 개최하기로 되어있다. 또한, 축구의 비디오 분석은 비디오 분석 분야에서 많은 관심 거리를 제공한다.

축구게임 분석은 크게 3가지로 분류할 수 있다. 첫째, 비디오 장면 변화 탐색과 인덱싱(indexing), 둘째, 움직이는 다중 객체 추출 및 표현, 셋째, 영상의 파노라마 합성이다. 표 1에서 보는 바와 같이 움직이는 객체의 추출 및 표현 문제는 축구게임에서 움직이는 객체 인식과 축구게임 포메이션(Formation) 인식을 포함한다. 더욱

\* 비 회 원 : 두원공과대학 소프트웨어개발과 교수  
shnam@doowon.ac.kr

\*\* 비 회 원 : 한국전자통신연구원 영상처리연구부 연구원  
kykim@computer.etri.re.kr

\*\*\* 정 회 원 : 신성대학 전산정보처리과 교수  
khyuns@avhana.shinsung.ac.kr

\*\*\*\* 중 심 회 원 : 고려대학교 컴퓨터학과 교수  
hwang@disys.korea.ac.kr

\*\*\*\*\* 정 회 원 : 한국전자통신연구원 영상처리연구부 부장  
ykyang@etri.re.kr

논문접수 : 1999년 9월 29일  
심사완료 : 2000년 9월 6일

이 동작 인식과 포메이션 인식은 축구게임 분석에서 기본 요소이다.

본 논문은 자연스럽게 동적인 환경에서 객체의 움직임 예측이 어려우며, 때때로 움직임이 매우 빠른 객체의 동작을 분석하고, 동작의 인식을 설명하며 구현한다. 객체의 동작을 인식하기 위해 동작을 구분하고 분석하여 동작의 대표특징을 추출한다. 그리고, 객체 주위에 다가오는 다른 객체와의 상관관계를 표현한다. 또한, 동작인식을 위한 알고리즘을 제안한다. 즉, 구간별 포아송-네제로 확률 분포를 통한 방법과 퍼지 집합과 신경망의 결합에 의한 인식방법을 제안한다. 그리고 제안된 방법들로 구현된 인식결과를 분석하고 평가한다. 우리의 알고리즘은 축구영역에서 시험되며 동작인식의 성공뿐 아니라 실패의 경우도 원인을 규명하고 미래에 연구할 과제를 제안한다.

논문의 구성은 1장에서 서론, 2장은 관련 연구 및 연구 방법 제안, 3장은 동작영역과 대표특징, 4장은 객체의 동작인식을 설명하고, 5장 실험결과 및 평가에 이어 6장은 결론과 미래의 연구과제를 제안한다.

표 1 축구 게임의 연구 분야

연구분야	아이템
A. 장면변화 탐색 및 인덱싱	1.자르기 2.상호작용 인덱싱
B. 움직이는 객체인식	1.분할과 추적 2.동작과 포메이션 인식
C. 파노라믹 뷰 (Panoramic View)합성	1.파노라믹 뷰 합성 2.모자이크(Mosaicking)

## 2. 관련 연구와 연구방법 제안

### 2.1 관련 연구

움직이는 객체에 대한 연구가 스포츠 영역에서 점점 많은 관심을 보이며 그 연구의 폭이 넓어지고있다. M.I.T.의 Stephen S. Intille[1]은 복잡하고 동적인 환경에서 객체의 추적 문제를 설명했다. 이들의 기법을 클로즈드 월드(closed world) 추적이라 한다. 이들은 클로즈드 월드 분석을 어떻게 하고 콘텍스트(context) 특징 추적이 미식축구 선수의 추적에 어떻게 사용될 수 있는지 설명하고 구현하였다. 이 연구는 자연스럽게 움직임이 빠르며, 움직임의 방향예측이 어려운 미식축구에서 객체 움직임의 분석에 관한 연구 수행이지만 동작인식이 아닌 추적에 관한 연구이다. Elisabeth Andre[2]는 축구 분석 시스템인 SOCCER 시스템을 연구하였다.

이들은 점진적 이벤트 인식 그리고 언어 생성의 적절한 결합에 대한 동시적 음성 자동생성을 논의했다. 이 논문은 축구의 이벤트에 따른 언어의 자동생성에 관한 연구이다. Chueh-Wei chang과 Suh-Yin Lee[3]는 스포츠 동작 분석에 대한 비디오 정보 시스템에서 몇 가지 중요한 방법을 설명했다. 첫째로 비디오 시퀀스에서 움직이는 객체 위치의 효율적 계산 방법을 설명하고, 객체 추적과 특징 추출에 활용했다. 또한 두 비디오 시퀀스에서 객체 움직임의 차이를 비교하기 위한 매칭 매커니즘을 제안했다. 이 논문은 이동 물체의 추적에 대한 설명과 추적된 객체의 비교 논문으로 동작분석의 성격이 강하지만 인간의 동작을 인식하는 논문은 아니다. Dennis Yow[4]는 비디오 이미지 내용을 분석해 축구의 하이라이트를 자동으로 탐색하고 추출하는 기법을 설명했다. 그리고 선택된 이벤트를 파노라마로 재구성하여 슛 동작을 표현했다. 이 분석은 축구의 특징을 인식하고, 볼을 추적하며, 효율적 인식을 위한 카메라 보정과 파노라믹 뷰의 구성을 포함한다. 이 논문은 객체의 추적과 복원에 관한 논문이다.

축구게임에서 그룹 행위에 대한 정량적 정성적 표현에 대한 연구도 몇 편 발견할 수 있다. T.Taki[5][6]는 고정된 여러대의 카메라에서 촬영한 축구경기 영상에 대해 집단행동 중에 각 선수들이 만드는 일종의 세력범위(motion trajectories; 공 패스를 유지하기 위한 동작, 또는 팀워크)를 특징량으로 하여 집단행동을 분석하는 연구이다. 두 논문은 축구영역에서 선수 개개의 행위분석이라기 보다는 그룹행위의 분석이다.

그리고, 사람의 동작인식에 관한 몇 가지 연구를 살펴보자. Koh Kakusho[7]는 실 이미지 시퀀스와 음악의 소리신호로부터 사교댄스의 종류를 인식하는 방법에 대해 논했다. 이들은 댄싱 할 때 음악의 강약과 소리를 추적하였으며, 댄싱 시 몸의 회전과 두 댄서의 위치 그리고 방향을 추적함으로써 댄싱의 종류를 인식했다. 그러나 이 논문은 사람 동작 자체의 인식보다는 댄서의 스텝과 위치 그리고 음악의 종류에 대한 인식을 다루고 있다. Lee Campbell과 Aaron Bobick[8]는 X,Y,Z 축의 추적 데이터로부터 전통 발레의 동작을 인식하는 시스템을 개발했다. 페이즈(phase) 공간의 하부 공간에서 공간 곡선에 기초한 동작 표현에 대한 기법이다. 이 페이즈 공간은 몸의 굴절되는 부분의 각과 몸통의 위치 그리고 자세의 축을 갖는다. 또한 하부 공간의 축은 페이즈 공간의 축에 대한 부분집합이다. 이들은 두 명의 댄서로부터 수행되는 9가지의 전통적 발레 동작을 인식하였다. 이 논문은 우리가 추구하는 객체의 동작 인식과

유사한 면이 있다. 하지만 이 논문은 고정된 객체 위치와 카메라 그리고 정형화된 동작을 인식한다.

Jia-Ching Cheng[9]은 비디오 영상에서 사람의 동작을 인식했다. 이들은 동적인 환경에서 모델기반 인식 방법을 제안하였다. 이들은 사람을 탐색하고 위치를 표시했으며 사람의 몸체와 걷기를 모델링 요소로 표현하고 모델링 요소에 동작요소를 비교하여 인식하였다. 하지만 이 논문은 똑바로 걷는 30 프레임의 비디오 영상에서 사람의 걷기를 인식하는데 그쳤다. Yaser Yacoob[10]은 많은 수의 일시적 파라미터가 이미지 시퀀스로부터 추출될 때 행동 모델링과 인식에 대한 매개변수적 모델을 제안하고 테스트했다. 이들은 이미지 시퀀스의 집합을 PCA(Principal Component Analysis) 선형 변화기법을 이용하여 걷기, 행진, 라인 걷기, 걸으며 차기 및 입술의 변화에 따른 언어를 인식하였다. 이 논문에서 사용된 데이터의 배경이 자연스럽고, 걷기 등의 인식범위는 우리의 인식범위와 상당히 유사한 면을 가지고 있으나, 이 논문은 실내에서 정형화된 움직임을 인식하며 객체의 움직임이 예측가능하고 움직임의 속도 또한 동작에 따라 커다란 차이를 보이지 않는다.

국내에서의 비디오 영상 분석에 대한 연구는 초보적인 수준이다. 포항공대[11][12]는 축구 게임 분석에 대한 작업을 수행했다. 이들은 자동 축구 분석을 위해 선수와 공의 추적 방법을 개발했다. 이들은 필드 모델 상에 선수와 공의 궤적을 구축했다. 또한 그들은 축구게임의 흑백 이미지 시퀀스로부터 축구공의 3차원 위치를 찾는 방법을 제안하였다. 공의 높이 전환은 그라운드상의 볼의 주어진 시작 위치와 끝 위치에 대한 간단한 삼각형의 기하학적 관계를 사용하여 쉽게 계산하였다. 김현숙[13]은 축구경기에서 선수들과 공의 위치정보를 추출하고, 그룹 포메이션 정보를 기초로 BP 뉴럴넷 알고리즘을 이용하여 축구경기 하이라이트 장면의 자동추출을 위한 공격패턴에 대한 자동 분류기법을 제안했다. 이 논문은 축구의 좌/우/중앙과 코너킥 등의 공격 패턴을 자동으로 분류한 논문으로 그룹행위 분석이다.

## 2.2 제안된 연구방법

움직이는 객체의 인식은 동작인식에 관심이 있는 수많은 연구자들로부터 크게 관심을 끌고 있다. 그러나 한정된 배경 하에 정형화된 움직임을 인식하는 것이 지금까지 대부분의 움직임은 객체의 인식 연구였다. 자연스러운 실세계의 환경에서 넓은 범위를 빠르게 움직이며, 객체의 움직임은 방향 예측이 어려운 형태의 객체 동작 인식은 그 연구사례가 많지 않다. 더구나, 자유자재로 움직이는 스포츠 영역에서의 동작인식은 그 연구 사례

가 극히 드물지만, 점점 많은 관심을 갖는 분야로 발전하고 있다.

동작인식은 어떤면에서는 동작의 분류를 위한 내재되어 있는 정량화된 통계적인 바탕이 존재한다. 통계적 동작인식은 통계적인 가정에 바탕을 두고 분류 업무를 수행한다. 사람의 동작은 동작 자체에 모호한 요소가 항상 잠재해 있다. 예를 들어, 일반적으로, 뛰기는 걷기보다 다리각이 크게 벌어지지만 어느 만큼 다리각이 크면 뛰는 것이고, 어느 만큼 다리각이 작으면 걷는 것인지 명확히 정의하기란 불가능하다. 따라서 우리는 이를 퍼지론으로 처리하여 3가지 범주로 나눈다. 또한 퍼지 처리는 수작업으로 수행한 내재된 오류를 보완한다.

HMM(Hidden Markov Model) 역시 통계적 방법론의 한 방법이다. HMM은 최근 제스처의 시공간적 변화를 모델링하기 위한 유용한 틀로 많은 연구자들에게 관심을 끌어들였다. HMM은 시간 공간적 정보를 자연스럽게 모델화할 수 있는 학습과 인식에 대한 훌륭한 알고리즘을 제공한다. 하지만, HMM은 넌-제스처(non-gesture) 패턴을 표현하는데 취약함을 보이고 있다.[14]

통계적 동작인식과 신경망을 이용한 동작인식을 명확히 구분하는 것은 어렵다. 특정한 인식 문제가 주어졌을 때 학습 능력의 적합성뿐 아니라 문제에 내재되어 있는 통계적 요소나 구조의 분석을 바탕으로 통계적 인식과 신경망을 이용한 패턴인식 등 여러 접근 방법중의 하나를 선택할 수 있다.[15] 신경망은 구별 가능한 특징을 분류하는 문제에 있어서 우수한 성능을 보이며, 패턴연상의 응용에 적합하고, 정보의 부분적 손실이나 왜곡, 잡음 등에 강한 장점을 보인다. 신경망은 정적 패턴을 인식하는데 효과적인 점을 보여주고 있지만 동적 패턴에서는 적합하지 않다. 이러한 신경망의 문제는 동적 패턴을 어떻게 정적 패턴 모델로 변환하는가가 관심사이다.[14]

한편, 사람의 움직임은 동적인 환경에서 한 동작의 구분이 정해진 시간에 발생되는 것이 아니며, 한 동작의 시작과 끝이 일정한 시간 간격으로 나누어지지 않고 매 동작에 따라 가변적이다. 동적인 동작 영역이란 이와 같이 한 동작의 시작과 끝의 가변적인 영역으로 정의한다. 그리고 정의된 동적 영역에서 동작을 구분하기 위해 동작의 대표적인 특징을 정의할 수 있다.

우리의 동작인식은 동적 특성에서 정적 특성을 추출하여 통계적 방법과 신경망을 사용해 인식하고 각 방법을 비교 설명한다. 본 논문은 동적인 동작 영역에서 대표특징을 기반으로 그 영역에서의 움직임이 무엇인가를 인식한다. 우리는 동적인 동작 영역의 정의와 동적 영역

에서의 대표특징을 제안하고 인식을 위한 방법을 제안한다. 동작영역에서의 대표특징을 통해 프레임간 객체 상태의 분석과 프레임간 연관 관계의 계산 량을 줄여 연산속도를 개선하고 잡음에 강한 인식을 도모한다. 인식을 위한 방법으로는 구간별 포아송 난-제로 확률 분포를 통한 방법과 퍼지 집합과 신경망을 혼합한 방법을 제안한다. 또한, 제안된 알고리즘으로의 구현 및 테스트와 평가를 통해 제안된 방법의 타당성을 검증한다.

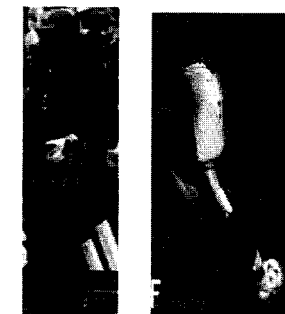
우리의 실험 환경은 축구와 미식축구의 차이는 있지만 Stephen S. Intille[1]가 수행한 실제 경기 상황에서의 실험과 유사한 면이 있으며, 동작 인식의 측면에서 보면 Lee Campbell과 Aaron Bobick[8]의 발레 동작 인식과, Yaser Yacoub[10]의 움직이는 객체 동작인식과 한 축에 있는 연구라 할 수 있다.

### 3. 동작영역과 대표특징

축구에서의 동작들은 그림1에서 보는바와 같이 대표적으로 드리블, 뛰기, 걷기, 서있기, 킁, 드리블 후 킁 등으로 구별할 수 있으며, 이들 동작들은 허리 이하 다리각의 변화를 가지고 구분할 수 있다. 그림 2는 축구선수의 다리각이다.



④ 드리블      ⑤ 뛰기      ⑥ 걷기



⑦ 서있기      ⑧ 킁

그림 1 축구선수의 동작

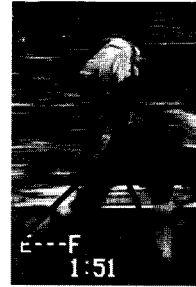


그림 2 축구선수의 다리각

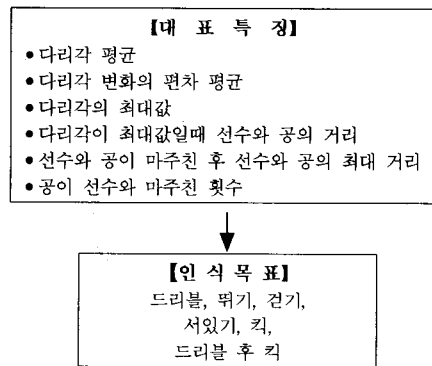


그림 3 대표특징과 인식목표

축구동작의 인식을 위해 그림3에서 보는바와 같이 다리각의 평균, 프레임간 다리각의 편차 평균, 다리각의 최대값, 다리각이 최대값일 때 선수와 공의 거리, 선수와 공이 마지막으로 마주친 후 선수와 공의 최대 거리, 선수와 공이 마주친 횟수를 한 동작의 비디오 장면에서 대표특징으로 정의한다.

위에 열거한 각각의 특징을 살펴보면 다음과 같다.

- 1) 다리각의 평균: 다리각의 변화는 선수의 스피드를 알 수 있는 척도이다. 다리의 각이 넓게 벌어진다는 것은 선수의 스피드가 빠르다는 것을 암시한다.
- 2) 다리각 변화의 편차 평균: 선수가 가만히 서있을 경우에는 다리의 움직임이 거의 없기 때문에 다리각의 변화가 매우 적다. 선수가 다리를 벌리고 서있건 그렇지 않건 간에 두 다리각의 변화 편차가 적은 것은 선수가 거의 움직이지 않는다는 단서가 된다.
- 3) 다리각의 최대값: 두 다리각의 최대값이 크면 빨리 뛸 확률에 대한 힌트가 되며, 공을 강하게 킁할 때에도 두 다리각은 일시적으로 심하게 벌어진다.
- 4) 다리각이 최대값일 때 선수와 공의 거리: 선수가 공을 킁하기 위해서는 일시적으로 한쪽다리를 뒤로 높

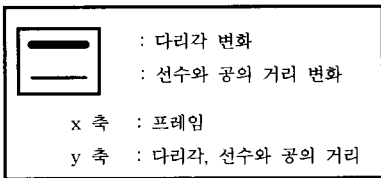
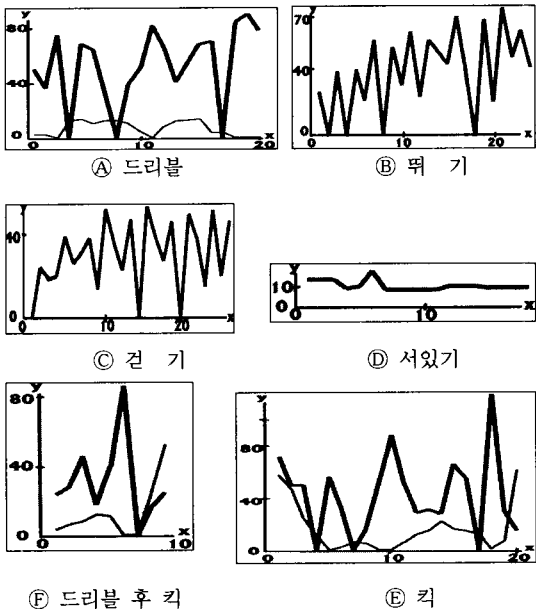
이 들어야 하며, 이때 보통 한 켄에서 다리각이 최대가 된다. 또한, 공을 킂하기 위해서는 공이 선수와 붙어 있어야하며, 이것은 공을 킂할 때의 단서가 된다.

5) 선수와 공이 마주친 후 선수와 공의 최대거리: 선수가 공을 치는 행위는 선수와 공이 일시적으로 만난 후 공이 멀리 날아가 버린다. 선수와 공이 마주친 후 선수와 공의 거리는 선수가 공을 킂했는지의 여부를 판가름하는 중요한 힌트가 된다.

6) 선수와 공이 만난 횟수: 드리블은, 선수가 움직이며 공을 소유하고 있기 때문에 여러번 선수와 공의 거리가 0 이 된다. 또는 공과 선수의 거리가 단 한번만 0이 됐다면 패스나 킂의 힌트가 되며, 공이 여러 번 선수와 만난 후 멀리 날아가 버리면 드리블 후 킂의 힌트가 된다.

그림 4는 다리각의 변화와 공과 선수의 거리변화를 표현한 그래프 예이다.

어떤 주어진 시간(또는 공간) 동안 발생하는 사건의 수를 다룬다. 다음 두 조건을 만족하는 확률변수는 포아송 분포를 따른다.



㉔ 범례

그림 4 다리각과 선수와 공의 거리 변화 그래프, 예

송 분포를 따른다. 첫째, 같은 주어진 시간(또는 공간)에 사건 출현 확률은 같다. 둘째, 어떤 주어진 시간의 사건 출현 또는 미 출현은 다른 주어진 시간의 사건 또는 미 출현과 독립이다. 이러한 포아송 분포는 다음 식을 갖는다.

4. 객체의 동작 인식

우리는 인식을 위한 방법으로 제안된 구간별 포아송 분포로 확률 분포를 통한 방법과 퍼지 집합과 신경망을 이용한 방법을 제안하고, 실험을 통하여 두 방법을 비교 분석한다. 그림 5는 동작인식 전체의 흐름도이다. 본 논문에서는 객체의 탐색과 추적 과정은 논외로 한다.

4.1 확률분포

포아송 분포

$$f(x) = (\mu^x e^{-\mu}) / x!$$

$$(x = 0, 1, 2, 3, \dots)$$

여기서 x는 확률변수,  $\mu$ 는 주어진 시간의 평균 발생 수이고,  $e = 2.71828$  이다.

가우스 분포

가우스분포 확률밀도함수는 다음과 같다.

$$f(x) = (1/\sigma \sqrt{2\pi}) \times e^{-a}$$

$$(a = (x-\mu)^2/2\sigma^2, -\infty < x < \infty)$$

확률변수 X가 평균  $\mu$ , 표준편차  $\sigma$ 인 가우스분포를 따를 때  $X \sim N(\mu, \sigma^2)$ 으로 표시한다. 이러한 확률변수 X에 대하여 다음과 같이 놓으면, Z는 평균이 0이고 표준편차가 1인 가우스분포를 따른다. 다시 말해서  $Z \sim N(0, 1)$ 이다.

$$Z = \frac{X - \mu}{\sigma}$$

이러한 확률변수 Z를 표준정규확률변수라 하고, 이 분포가 표준가우스분포이다.

신경망

신경망은 기존의 컴퓨터와는 달리 연속적인 아날로그 신호를 처리한다. 기존의 컴퓨터가 ON/OFF 형태의 신호를 취급하여 의사를 결정하고, 프로그램에 의해서 정해진 순서대로 처리하는 반면, 신경망은 의사 결정을 융통성 있게 표시할 수 있고 절차적인 프로그램 없이 학습에 의해 스스로 처리될 수 있다. 정보 검색에 있어서도 기존의 컴퓨터는 정확한 정보를 다루기 때문에 정보의 부분적 손실이나 잡음 등에 취약한 반면, 신경망은 가장 적절한 정보를 다루기 때문에 정보의 부분적 손실이나 왜곡, 잡음 등에 강한 장점을 보인다. 또한 신경망은 뉴우런(Neuron)과 뉴우런의 연결 강도를 조절하여 정보를 저장하는 분산 저장의 형태로 이루어진다.

### 4.2 동작인식

#### 통계적 가중치 인식

통계적 가중치 인식방법으로 구간별 년-제로 포아송 확률분포를 통한 인식 방법을 제안한다.

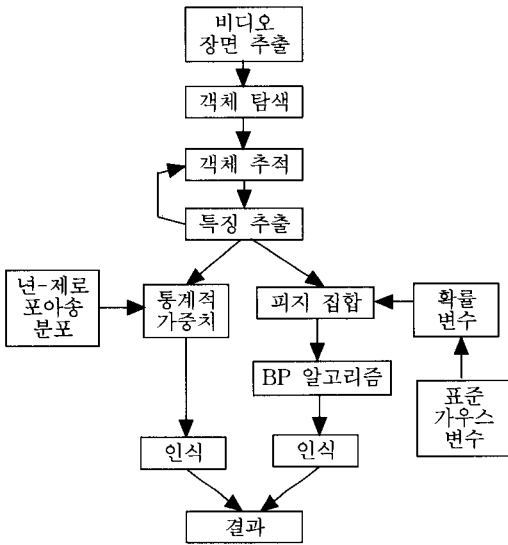


그림 5 동작인식의 전체 흐름도

가중치를 이용한 방법이란 추출한 각각의 특징 값에 따라 가중치를 부여하고, 각각의 인식동작에 가중치를 연산하여 최종 결과 값에 따라 동작을 구별하는 방법이다.

그림 6은 가중치에 의한 인식을 보인 것으로 대표특징 6가지를 의미한다.

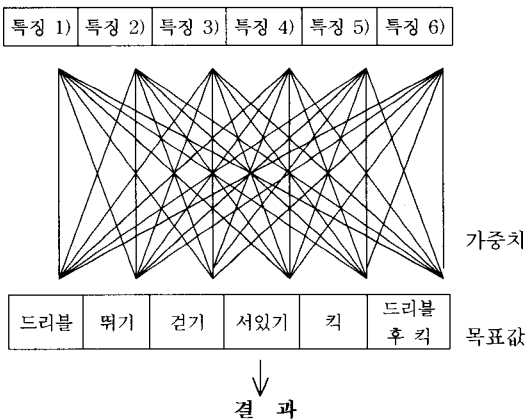


그림 6 가중치에 의한 인식

포아송분포의 사건발생 공간을 10단위의 구간으로 분류를 한다(단, 선수와 공의 마주친 횟수는 1단위의 구간으로 분류). 예를 들면, 뛰기에서 다리각의 평균이 46.7이었다면, 뛰기의 5번째에 분포하게 되는 것이다. 이러한 분포에 따라 포아송 확률분포 함수를 이용하여 각 분포에 따라 주어진 구간에서 사건발생이 0이 될 확률을 구한다. 0이 될 확률이  $a$  라 하면 0보다 클 확률은  $(1.0 - a)$ 이다. 이 값이 가중치이다. 예를 들면, 뛰기의 다리각 평균이 40.0과 49.9 사이의 분포가 0일 포아송 확률 분포 값이 0.2라면 1 이상일 확률은 0.8이며 가중치가 0.8이 된다.

$\mu$ 를  $y$ 구간(예, 뛰기의 다리각 평균을 45.7이라 하면  $40 \leq y < 50$ , 즉, 구간은 뛰기의 5번째가 된다)의 평균 값이라 하면, 포아송 분포에서 0일 확률은 다음과 같이 함수로 표현한다.

$$f(x, y, \mu) = (\mu^x e^{-\mu}) / x!$$

$$(x = 0), (\mu \text{는 } y \text{ 구간에서의 평균값}),$$

그러므로, 위 함수에 따른  $y$  구간에서의 가중치함수는 다음과 같이 표현된다. 이것이 구간별 년-제로 포아송 확률분포이다.

$$W(y) = 1.0 - f(x, y, \mu)$$

$\text{Min}(f(x, y, \mu))$ 을  $f(x, y, \mu), (y > 1)$ 의 최소값으로 정의하면, 선수와 공이 마지막으로 만난 후 공과 선수의 최대거리, 그리고 선수와 공이 만난 횟수의 가중치 함수는 다음과 같이 표현한다.

$$W(y) = 1.0 - \text{Min}(f(x, y, \mu)), (y > 1)$$

$W(y)$ 가 구해지면 각각의 입력값에 따른 가중치를 목표값에 곱하여 가장 큰 수가 나오는 것이 결과값이다. 그림 7은 년-제로 확률분포를 통한 인식을 그림으로 나타내었다.

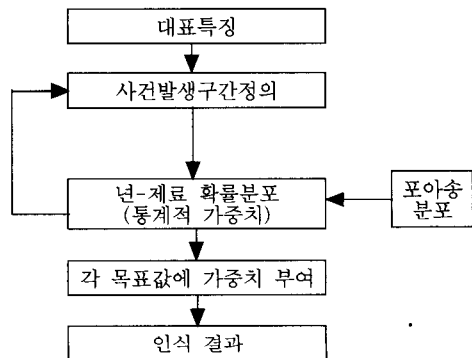


그림 7 년-제로 확률분포를 통한 인식

**퍼지집합과 신경망을 이용한 인식**

표준 가우스 확률변수를 이용한 퍼지 집합을 구성하고 설계된 퍼지 집합과 신경망을 이용한 인식 방법을 제안한다. 신경망은 역전파(BP:backpropagation) 알고리즘을 이용한다. 퍼지 집합을 통해 추론된 퍼지 소속 정도값을 신경망의 입력으로 사용하여 동작인식을 위한 학습을 수행한다. 퍼지 집합을 만들기 위한 소속 함수로는 가우스 분포 함수를 사용한다.

인식 특징값을 작다(S), 중간(M), 크다(L)로 정의한다. 또한 작다와 중간사이, 중간과 크다 사이를 각각 (SM)만큼 작고 (MS)만큼 중간값을 갖는다고 표현하며, (ML)만큼 중간이고 (LM)만큼 크다고 표현한다. 즉, 평균을 중심으로 평균값 가까이 있지 않으며 평균보다 작은 것은 작다로 정의하고, 평균값 주변의 값은 중간으로 정의하며, 평균 가까이 있지 않으며 큰 것은 크다고 정의한다. 그러나, 최대각일 때 선수와 공의 거리는 작다(S)와 크다(L)로 정의한다. 퍼지 집합이 가우스 분포를 따를 때 평균은 표준확률분포의 0 이 된다. 그러므로 X를 확률변수, Z를 표준정규확률 변수라 하고, S, SM, MS, M, ML, LM, L을 위한 기준 표준 정규 확률 변수를 z', z''이 라하며, 확률변수 X를 구하는 함수를 f(Z)라 하면 확률변수 X를 구하는 확률변수 함수는 다음과 같이 표현된다.

$$f(Z) = \mu + Z * \sigma,$$

$$(-3.0 < Z < 3.0, -3.0 < z' \pm z'' < 3.0, z' \geq z'')$$

만일 (Z <= -(z'+z''))이면 S,  
 만일 (-(z'+z'') < Z <= -(z'-z''))이면 SM,

MS,

만일 (-z'' < Z <= z'')이면 M,  
 만일 (z'-z'' < Z <= z'+z'')이면 ML, LM,  
 만일 (Z > z'')이면 L 이다. 그림8 ㉠.

SM, MS의 정도를 표현하면 다음과 같다.

$$SM = \frac{(\mu - z' * \sigma) - f(Z)}{(\mu - (z'-z'') * \sigma) - (\mu - (z'+z'') * \sigma)},$$

$$MS = 1.0 - SM.$$

또한, ML, LM의 정도 표현은 다음과 같다.

$$ML = \frac{(\mu + (z'+z'') * \sigma) - f(Z)}{(\mu + (z'+z'') * \sigma) - (\mu + (z'-z'') * \sigma)},$$

$$LM = 1.0 - ML.$$

그리고, 최대각일 때 공과 선수의 거리에 따른 퍼지 집합을 살펴보면 다음과 같이 조건이 다르게 나타난다. 그림8 ㉡.

만일 (Z <= 0)이면 S,  
 만일 (0 < Z <= z')이면 SL, LS,

만일 (z' < Z)이면 L이다.

SL, LS의 정도 표현은 다음과 같다.

$$SL = \frac{\mu - f(Z)}{z' * \sigma},$$

$$LS = 1.0 - SL.$$

그리고, 공과 선수가 만난 횟수의 조건은 다음과 같다. 만난 횟수를 X라 하면,

만일 (X <= 0)이면 S,  
 만일 (0 < X < 2)이면 M,  
 만일 (X >= 2)이면 L이다. 그림8 ㉢.

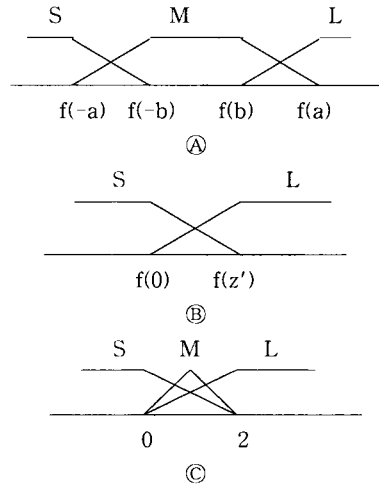


그림 8 표준 가우스 확률변수를 이용한 퍼지 집합(단, a=z'+z'', b=z'-z'')

**5. 실험 결과**

**5.1 자료의 샘플링 및 제약조건**

본 논문의 자료는 1998년 프랑스 월드컵의 축구경기의 결승전으로 프랑스와 브라질 전, 3-4위 전으로 네델란드와 크로아티아 전, 그리고 1999년 3월 한국과 브라질의 서울 친선 경기에서 비디오로 녹화된 테이프에서 6프레임에 하나 꼴로 전체 5,650프레임을 샘플링하였다. 이는 인식하고자하는 6개 동작 각각에 대하여 60개씩 360개의 장면이다.

본 논문의 특징추출은 다리의 각, 공과 선수의 거리로 추출이 된다. 본 논문의 동작 영역은 동작이나 비디오 장면의 변화에 따라 자동으로 인텍싱 되어야 하지만 그의 구체적인 방법과 구현은 논외로 하고, 본 논문에서는 이를 수 작업으로 수행하였다. 선수와 공의 탐색, 다리 각과 거리의 계산 또한 자동으로 추출되어야 하지만,

그 작업이 방대한 시간을 소요하는 작업이기 때문에 우선 포토샵으로 자동화의 가능성만을 검토한 후 선수의 다리각과 선수와 공의 거리는 수동으로 추출하였다. 자동으로의 추출은 반드시 수행해야 하는 차후 과제이다. 수작업으로 하다보니 작은 오차들이 잠재해 있다. 이의 극복을 위해 다리각의 경우 다리 뒤꿈치에 맞추었으며 공과의 거리는 발의 앞부분과 공의 최단거리부분을 기준으로 정밀을 기했으며, 퍼지를 통해 극복했다. 또한 약간의 오차는 인식에 별로 영향을 끼치지 않았다.

정의된 동작인식의 범위에서(즉, 드리블, 뛰기, 걷기, 서있기, 킥, 드리블 후 킥), 인식을 위한 방법으로 구간별 포아송 난-제로 확률 분포를 통한 방법, 퍼지 집합과 신경망을 결합한 방법등 2 가지를 제안하고 실험한다. 신경망을 위한 학습데이터와 통계적 가중치에 의한 통계 데이터는 동작 인식 특징 각 20개씩 120개의 장면을 사용하며, 실험 데이터는 인식 특징 각 40개씩 240개의 장면을 사용하였다.

**5.2 실험결과 및 고찰**

**실험 결과**

인식결과를 살펴보자. 우선, 통계적 가중치를 이용한 방법의 경우, 표 2에서 보는 바와 같이 뛰기의 6개가 걷기로 오 인식되었고, 1개의 미 인식이 있었다. 또한 킥에서는 1개가 드리블 후 킥으로 오 인식되었으며, 11개가 미 인식되었다. 그리고, 드리블 후 킥은 10개의 미 인식이 있었다. 둘째, 퍼지 집합과 신경망의 결합을 이용한 경우, 표 3에서 보는 바와 같이 드리블에서는 1개가 드리블 후 킥으로 인식하였다. 뛰기는 4개가 걷기로 인식하였고, 2개는 서기로 인식하였다. 걷기는 4개가 서있기로 인식하였으며, 킥은 1개가 드리블 후 킥으로 인식하였다. 그리고 드리블 후 킥은 3개가 드리블로 인식되었다.

**평가 및 고찰**

구간별 포아송 난-제로 확률 분포를 통한 통계적 가중치 방법과 퍼지와 신경망을 결합한 방법으로 실험하

표 2 통계적 가중치 인식결과

	오인식	미인식	인식률
드리블	0	2	95%
뛰기	6 -> 걷기	1	82.5%
걷기	0	0	100%
서기	0	0	100%
킥	1 > 드리블 후 킥	11	70%
드리블 후 킥	0	10	75%
종합	7	24	87.1%

표 3 퍼지와 신경망

	오인식	미인식	인식율
드리블	1 -> 드리블 후 킥	0	97.5%
뛰기	4 -> 걷기 2 -> 서있기	0	85%
걷기	4 -> 서있기	0	90%
서있기	0	0	100%
킥	1 -> 드리블 후 킥	0	97.5%
드리블 후 킥	3 -> 드리블	0	92.5%
종합	15	0	93.8%

표 4 인식을 종합표

	인식율	
	통계적 가중치	퍼지,신경망
드리블	95%	97.5%
뛰 기	82.5%	85%
걷 기	100%	90%
서있기	100%	100%
킥	70%	97.5%
드리블 후 킥	75%	92.5%
계	87.1%	93.8%

였다. 제안한 방법으로 테스트한 결과 표 4에서 보는바와 같이 각각 87.1%, 93.8%의 인식결과를 얻었다. 비록 퍼지와 신경망을 결합한 방법이 걷기에서 통계적 가중치보다 많은 오류가 발생했지만, 킥과 드리블 후 킥의 인식에서 통계적 가중치의 오류를 현저히 감소시켰으며, 오류가 전체적으로 특정 동작에 치우치지 않고 고르게 나타났다. 위와 같은 실험결과로 종합해 볼 때 퍼지와 신경망의 결합이 통계적 가중치를 이용한 방법보다 우수하다고 말할 수 있다.

뛰기를 걷기로 오인식 한 것이 많았다. 뛰기가 걷기로 오인식 되는 것은 매우 천천히 뛰었기 때문에 뛰기의 특징이 명확히 나타나지 않아서이다. 이때는 다리각이 많이 벌어지지 않아 다리각이 조금 빨리 걷는 경우와 유사하게 나타나기 때문이다. 그리고, 걷기가 서있기로 오인식 된 것은 대부분 천천히 걸을 때이다. 서있기는 전부 인식하였다. 이는 서있기의 중요한 특징인 다리각의 편차 평균이 뚜렷이 구별되는 특징을 보이기 때문이다. 또한 드리블, 킥, 드리블 후 킥은 전체적으로 서로 서로를 오 인식하여 인식하지 못하는 결과를 보였다. 이러한 이유는 드리블, 킥, 드리블 후 킥의 구별 요소가 최대각일 때 공과 선수의 거리, 선수와 공이 마지막으로 만난 후 공과 선수의 거리, 공과 선수가 만난 횟수 등으



로 구별하는데, 드리블, 킥, 드리블 후 킥이 모두 공과의 상관관계가 있기 때문에 이들의 상관관계가 뚜렷이 구별되지 못했기 때문이다.

## 6. 결론

비디오로 녹화된 실세계의 자연스러운 환경에서 빨리 움직이거나 움직임을 예측하기 어려운 동작에 대한 접근법을 설명하였다. 동작의 범위는 축구 선수의 드리블, 뛰기, 걷기, 서있기, 킥, 드리블 후 킥 등이다. 동작분석을 위해 한 동작의 동적인 동작영역을 정의하고, 동적인 동작영역에서 대표특징을 추출하였다. 동적인 동작영역에서의 대표특징을 통한 인식은 앞 뒤 프레임간 객체상태의 분석이 간단하고 프레임간 연관관계의 계산이 필요없기 때문에 계산량을 획기적으로 줄여 연산시간을 개선하였다. 대표특징으로는 두 다리 사이각의 평균, 두 다리각 변화의 편차 평균, 두 다리가 최대로 벌어졌을 경우의 값, 두 다리가 최대값일 때 공과 선수의 거리, 공과 선수가 마지막으로 만난 후 공과 선수의 최대거리, 공과 선수가 만난 횟수 등이다. 인식을 위한 방법으로는 구간별 포아송 난-제로 확률 분포를 통한 방법, 퍼지와 신경망을 결합한 방법으로 실험하였다. 제한한 방법으로 테스트한 결과 각각 87.1%, 93.8%의 인식결과를 얻었다. 퍼지와 신경망을 이용한 방법이 걷기에서 통계적 가중치보다 많은 오류가 발생했지만, 킥과 드리블 후 킥의 인식에서 통계적 가중치의 오류를 현저히 감소시켰으며, 오류가 전체적으로 특정 동작에 치우치지 않고 고르게 나타났으며, 인식율이 높았다. 테스트는 축구 경기에서 비디오로 녹화된 입력을 가지고 테스트했다. 향후 구현 및 연구과제는 동적 동작 영역의 자동 인텍싱, 객체의 자동탐색 및 특징의 자동추출, 유전자 알고리즘(genetic algorithms)을 이용하여 인식문제를 해결하는 방법과 신경망과 유전자 알고리즘을 혼합해 동작을 인식하는 방법의 연구가 필요하다.

## 참 고 문 헌

- [1] Stephen S. Intille, "Tracking using a local closed-world assumption: Tracking in the football domain," M.I.T. Media Lab Perceptual Computing Group Technical Report No. 296, 1994.
- [2] Elisabeth Andre, Gerd Herzog, and Thomas Rist, "On the simultaneous interpretation of real world image sequences and their natural language description: The system SOCCER," Proc. of the 8th ECAI, pp.449-454, Munich, 1988.
- [3] Chueh-Wei Chang and Suh-Yin Lee, "A video information system for sport motion analysis," Journal of Visual Languages and Computing, pp. 265-287, 1998.8.
- [4] Dennis Yow, Boon-lock Yeo, Minerva Yeung and Bede Liu, "Analysis and presentation of soccer highlights from digital video," ACCV '95, 1995.
- [5] Toshio Kawashima, Kazuyoshi Yoshino, and Yoshinao Aoki, "Qualitative image analysis of group behaviour," Dept. of Information Engineering, Hokkaido University.
- [6] Tsuyoshi Taki, Jun-ichi Hasegawa and Teruo Fudumura, "Development of motion analysis system for quantitative evaluation of teamwork in soccer games," School of Computer and Cognitive Sciences, Chukyo University.
- [7] Koh Kakusho, Noboru Babaguchi and Tadahiro Kitahashi, "Recongnition of social dancing from auditory and visual information," International Conference on Automatic Face and Gesture Recognition, IEEE, October 14-16, 1996.
- [8] Lee Campbell and Aaron Bobick, "Using phase space constraints to represent human body motion," International Workshop on Automatic Face and Gesture Recognition, Zurich, 1995.
- [9] Jia-Ching Cheng, and Jos M.F.Moura, "Model-based Recognition of Human Walking in Dynamic Scenes," IEEE Signal Processing Society 1997 Workshop on Multimedia Signal Processing, Princeton, New Jersey, USA, Electronic Proceedings, June 23-25, 1997.
- [10] Yaser Yacoob, "Parameterized Modeling and Recognition of Activities," Computer Vision and Image Understanding, Vol. 73, No. 2, February, pp. 232-247, 1999.
- [11] Sunghoon Choi, "Soccer game analysis using mosaicking and color-based tracking," Department of Electrical and Electronic Engineering, Pohang Univ, 1997.
- [12] Taeone Kim, Yongduek Seo, and Kisang Hong, "Physics-based 3D position analysis of a soccer ball from monocular image sequences," Department of Electrical and Electronic Engineering, Pohang Univ, 1997.
- [13] 김현숙, 김광용, 남성현, 황종선, 양영규, "뉴럴네트워크를 이용한 축구경기에서 공격패턴 자동 분류 기법", 정보과학회논문지: 소프트웨어 및 응용 제 27권 제 7호, 2000.7.
- [14] Hyeon-Kyu Lee and Jin H. Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, No.10, October 1999.
- [15] 이성환, "패턴인식의 원리", 홍릉과학출판사, 1994.9.

남 성 현

정보과학회논문지: 소프트웨어 및 응용  
제 27 권 제 7 호 참조

김 광 용

정보과학회논문지: 소프트웨어 및 응용  
제 27 권 제 7 호 참조

김 현 숙

정보과학회논문지: 소프트웨어 및 응용  
제 27 권 제 7 호 참조

황 종 선

정보과학회논문지: 소프트웨어 및 응용  
제 27 권 제 7 호 참조

양 영 규

정보과학회논문지: 소프트웨어 및 응용  
제 27 권 제 7 호 참조