

구조적 자기구성 지도를 이용한 인간 행동의 성별 분류

류중원, 조성배

연세대학교 컴퓨터과학과

rjungwon@candy.yonsei.ac.kr, sbcho@csai.yonsei.ac.kr

Gender Classification of Human Behaviors Using Structure Adaptive Self-organizing Map

Jungwon Ryu^o

Sung-Bae Cho

Dept. of Computer Science, Yonsei University

요 약

본 논문에서는 구조적 자기구성 지도 모델을 사용하여 인간 행동의 성별을 분류하는 인식기를 제안하였다. 26명의 사람이 '화난 상태' 혹은 '보통 상태'의 두가지 정서 하에서 '문 두드리기,' '손 흔들기,' '물건 들어올리기'의 세가지 동작을 수행하는 동안, 행위자 관절점의 속도나 위치 정보로부터 성별을 분류하였다. 또한 SASOM의 성능 비교 분석을 위하여 전통적인 SOM, 다층 퍼셉트론과 그의 두 가지 결합 모델, SASOM와 의사결정 트리 결합 모델, 단일 의사 결정트리, k -최근접 이웃 등의 인식기를 구현하여 성능을 비교분석 하였다. 실험 결과 SASOM 분류기가 가장높은 인식률을 보였으며 분류기로서 유용함을 알 수 있었다.

1. 서론

인간과 컴퓨터가 효과적으로 상호작용하기 위해서는 인간 행동의 인식에 관한 학제적 연구가 필요하다. 본 논문에서는 점광화면으로 표현된 행위자의 행동으로부터 성별을 분류하는데 구조적 자기구성 지도(SASOM, Structure Adaptive Self-organizing Map)를 분류기로 사용하였다.

Teuvo Kohonen이 처음으로 제안한 자기구성 지도(Self-organizing Map)는 비교사 학습 방식의 하나로, 데이터의 위상을 보존하면서 스스로 학습하여 노드들 간의 가중치가 업데이트 된다[1]. 자기 구성지도는 학습 단계에서 전방 전달(feedforward) 방법을 사용하므로 빠른 학습을 수행할 수 있고, 연속적인 학습이 가능하다는 장점이 있다. 그러나 자기구성 지도는 학습되기 전에 위상을 미리 고정시켜 놓아야 하고, 클러스터링에 비해 분류 능력이 낮다. 구조적 자기구성 지도는 위상 구조가 학습 중에 적절히 변하기 때문에 이러한 문제를 해결할 수 있고, 코호넨의 비교사 학습방법에 LVQ 알고리즘을 결합한 교사 학습방법을 통하여 분류기로서 향상된 결과를 얻을 수 있다[2].

본 연구에서는 구조 적응형 자기구성지도를 행위자의 성별을 인식하는 분류기로 사용하여 그 사용 가능성 및 타당성을 검증하고자 한다. 또한 전통적인 인공지능 방법중의 하나인 신경망, SOM 또는 결합 신경망, SASOM과 의사결정트리 결합, 단일 의사결정 트리, k -최근접 이웃 등의 분류기와 성능을 비교해 보고자 한다.

2. 인간 행동의 인식

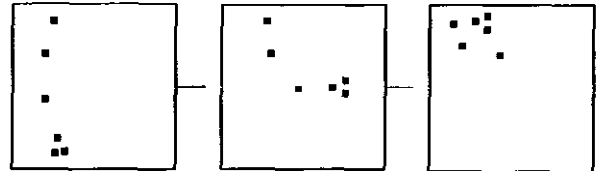


그림 1. 점광화면의 예

최근 심리학 분야에서 점광화면(point light display)은 생물학적 운동의 인식에 관한 연구에 있어서 가장 널리 사용되는 방법이다. 이는 행위자의 움직임을 점으로 표현하여 이들의 움직임을 관찰하는 방법이다. 이 방법을 처음으로 창안한 G. Johansson은 사람의 몸에 10개의 반짝이는 유리구슬을 부착하고 이들의 움직임을 관찰하였는데, 유리구슬들이 움직이기 시작하면 어느정도 사람들이 행위자를 인식할 수 있다는 사실을 밝혀 내었다. 특히 점광화면을 통해서 보행자의 성별을 판별하는 연구는 1977년 Cutting과 Kozlowski가 최초로 수행한 이래로 계속되어오고 있다[2, 3].

한편, 인공 지능 분야의 기계학습 이론은 주어진 데이터들로부터 귀납적으로 특정 규칙을 얻어내는데 사용되는데, 우리는 이들 패턴인식 방법론의 하나인 신경망 분류기법을 사용하여 행위자의 성별을 인식하는 실험을 하였다.

3. 실험 방법

3.1 전처리

행위자의 동작이 일어나는 동안 Optotrak (Northern Digital Co.)이라는 3D 좌표 분석 시스템을 통해 머리, 어깨, 팔꿈치, 손목, 손가락1, 손가락2의 6개의 관절점의 좌표 데이터를 남녀 각 13명으로부터 수집하였다. 이 좌표 데이터를 통해 성별 인식 실험에 사용될 특성값인 x, y, z 좌표 (Pos_{xyz}), y, z 좌표 (Pos_{yz}), yz 평면상의 관절점의 속도 (Vel_{yz}), xyz 3차원 공간상의 관절점의 속도값 (Vel_{xyz})의 네가지 값을 추출하였다.

행위자는 '보통상태', '화난상태'의 두 가지 정서 하에서 '문 두드리기', '손흔들기', '물건 들어올리기'의 세가지 동작을 수행하였다. 각 동작은 10회씩 반복되었고, 1,248개가 학습 데이터, 312개가 테스트 데이터로 사용되었다. 모든 입력 벡터는 150 차원으로 정규화 되었다.

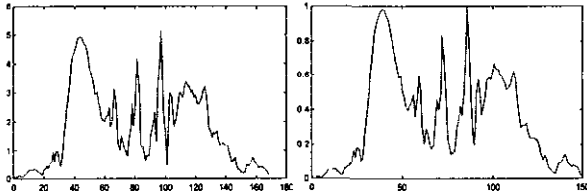


그림 2. 패턴 정규화 전·후의 모습

3.2 분류기

(가) 구조적용 자기구성 지도

구조적용 자기구성 지도는 노드 분화 방법과 LVQ 알고리즘을 적용하여 자기구성 지도의 구조 적응능력을 향상시킨다. 자기구성 지도의 4×4 크기로 초기화한 후, 주어진 학습 데이터를 통해 전통적인 코호넨 알고리즘으로 SOM 지도를 학습시킨다. 시간 $t+1$ 에서의 노드 i 의 가중치는 다음과 같다.

$$m_i(t+1) = m_i(t) + a(t) \times n_{ci}(t) \times \{x(t) - m_i(t)\}$$

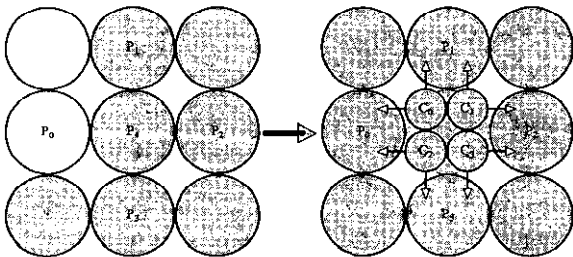


그림 3. 분화된 노드의 가중치 결정

$a(t)$ 는 시간에 따른 학습률을 나타내는 함수, $n_{ci}(t)$ 는 이웃 함수, $m_i(t)$ 는 i 번째 노드의 가중치, $x(t)$ 는 입력 벡터값이다. $n_{ci}(t)$ 에서 c 는 승리자 노드의 인덱스이다. 승리자 노드는 다음의 수식으로 얻을 수 있다[1].

$$\|x - m_c\| = \min(\|x - m_i\|)$$

지도의 노드들 중 여러 클래스의 데이터가 섞인 노드가 있으면 이것을 2×2 크기의 노드로 분화한다. 분화된 노드

들을 LVQ 알고리즘으로 학습하며, 종료 조건이 만족할 때까지, 분화 및 LVQ 학습의 과정을 반복하여 수행한다(그림 3). 분화된 노드들의 가중치는 자신의 부모 노드의 가중치 값을 기반으로 산출된다(그림 3)[4].

$$C_i = \frac{P_4 \times 2 + P_i + P_{i+1}}{S}$$

($i = i \bmod 4$, S 는 분화에 참여한 노드의 수)

(나) 다층 퍼셉트론

하나의 신경망 분류기는 900~2700개의 입력 노드와 50개의 은닉 노드, 2개의 출력 노드를 갖는다. 구성된 단일 신경망 분류기(S-MLP) 여러개를 결합하면 좀더 나은 성능의 분류기를 얻을 수 있는데[5, 6], 본 논문에서 제안한 신경망 분류기는 그림 4과 같다. 신경망 a~c는 '보통상태'의 동작들에 의해 학습된 신경망이고, 신경망 d~f는 '화난상태'의 동작들에 의해 학습된 신경망이다. 또 신경망 a와 d는 '문 두드리기' 동작, 신경망 b와 e는 '손 흔들기' 동작, d와 f는 '물건 들어올리기' 동작만으로 학습되었다.

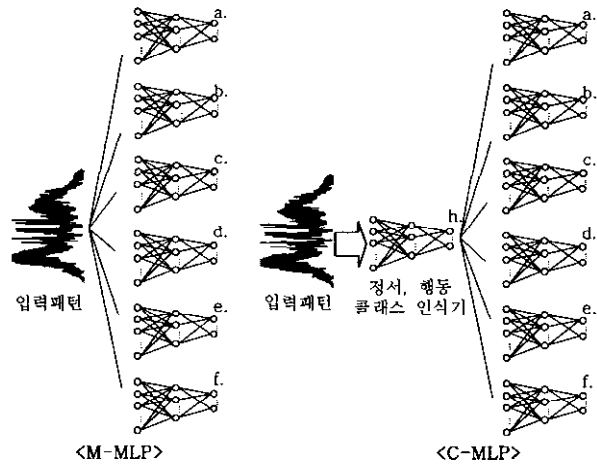


그림 4. 두가지 결합 신경망

이렇게 학습된 전문가 신경망들은 다음의 두가지 방법에 의해 결합 되었다. M-MLP는 6개의 전문가 신경망의 출력값중 최대값을 최종 분류 결과로 정하는 방법이고, C-MLP는 여섯 개의 소분류기 앞에서 입력패턴의 소속 클래스를 맞추는 대분류기 h를 두어 결합하는 방법이다. 신경망 h에 의해 선택된 소분류기의 출력값이 전체 신경망 분류기의 최종 출력값이 된다.

(다) 기타 분류기

이밖에도 분류기의 성능을 비교·분석을 위하여 코호넨의 자기구성 지도(10×10), 단일 의사결정 트리, k -최근점 이웃 등의 분류기를 사용하여 실험을 수행하였다. k -최근점 이웃에서는 $k = 10$ 부터 25까지 변화해 가면서 실험하였는데 가장높은 인식률을 보인 20값을 사용하였다.

4. 실험 결과

그림 5는 SASOM의 학습 초기단계와 학습이 끝나고 난 후의 자기구성지도의 모습이다. 남자는 1, 여자는 0으로 표시되어 있다. 클래스가 다른 데이터가 섞여있는 노드가 분화되어 여러차례의 노드 분화 후에 오른쪽과 같은 형태로 자기구성 지도가 만들어 졌다.

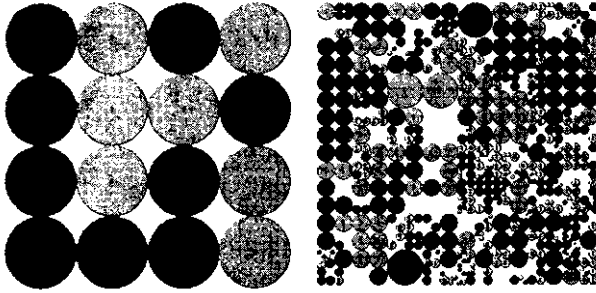


그림 5. 노드 분화 전·후의 구조 적응형 자기구성 지도

분류기들의 최종 인식률 결과는 표 1과 같다. 분류기별 성능을 보면 SASOM을 사용한 경우가 평균 90.8%로 가장 높은 성능을 보였다. 이는 실험에 사용된 여러 방법들 중에서 평균 90% 이상의 인식률을 보인 유일한 분류기이며 여섯 지점의 x, y, z 좌표값을 사용한 경우에는 97.1%까지의 인식률을 보였다. SOM의 결과와 비교해 볼때 구조 적응형 자기구성지도가 패턴 분류기로서의 성능 향상을 가져왔음을 알 수 있다.

신경망을 분류기로 사용한 경우, 하나의 신경망보다는 여러개의 소신경망을 결합한 C-MLP의 경우가 더 높은 인식률을 보였다. 이는 적절한 결합 방법을 사용하면 신경망 하나를 분류기로 사용하는 것보다 여러개의 전문화된 신경망을 결합했을 때 분류기의 성능을 높일 수 있다는 것을 보여준다.

네 가지 정보 표현 형태 중에서 성별 인식에서 가장 높은 인식률을 보인 것은 Pos_{xyz} 로 나타났다. Pos_{xyz} 에 대해서는 대부분의 분류기들이 인식률 향상을 보였다(그림 6).

표 1. 분류기별 성능 비교

분류기	데이터				인식률 평균
	Vel_{yz}	Vel_{xyz}	Pos_{yz}	Pos_{xyz}	
SOM	65.7%	60.6%	76.6%	76.9%	70.0%
SASOM	86.5%	85.9%	93.6%	97.1%	90.8%
S-MLP	75.2%	81.4%	59.6%	84.6%	75.2%
M-MLP	60.2%	59.0%	61.2%	61.5%	60.5%
C-MLP	80.0%	80.8%	84.6%	86.5%	82.9%
DT	70.2%	67.0%	69.2%	72.8%	69.8%
kNN	82.4%	81.4%	84.9%	92.9%	85.4%

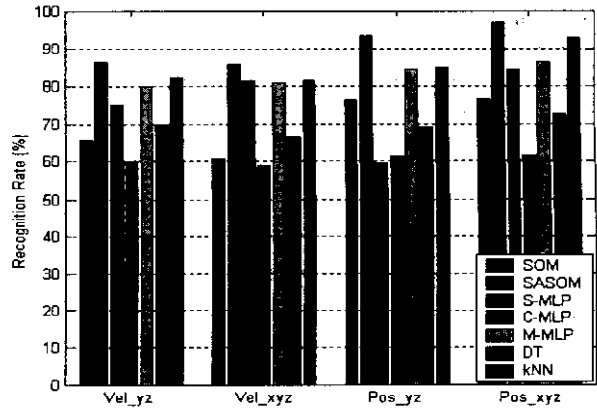


그림 6. 데이터 종류에 따른 분류기의 인식률 비교

5. 결론 및 토의

코호넨의 자기구성 지도 알고리즘을 개선한 구조 적응형 자기구성 지도 모델을 사용하여 인간 행동의 성별인식 실험을 인식하였다. 실험을 통해 코호넨 학습 알고리즘과 LVQ를 결합한 교사 학습 방식인 구조 적응형 자기구성 지도가 인간 행동의 성별인식 문제에 분류기로서 좋은 성능을 보인다는 것을 입증하였다. 전통적인 SOM을 이용한 경우보다 약 20%의 인식률의 향상을 가져왔다. 역전파 신경망을 결합하여 구현한 C-MLP의 평균 인식률이 82.9%, k -최근접 이웃 분류기의 결과가 85.4%였는데, SASOM은 이보다 높은 90.8%의 인식률을 보였다.

참고 문헌

- [1] T. Kohonen, *Self-organizing Maps*, Springer, Berlin Heidelberg, 1995.
- [2] G. Johansson, "Visual motion perception," *Scientific American*, vol. 232, p. 76~89, 1975.
- [3] J. Ryu and S. B. Cho, "Searching for optimal features for gender recognition with neural network classifier," *Proc. of 6th Intl. Conf. on Soft Computing*, IIZUKA2000, October, 2001.
- [4] S. B. Cho, "Self-organizing map with dynamical node splitting: Application to handwritten digit recognition," *Neural Computation*, vol. 9, p. 1345~1355, 1997.
- [5] A. J. C. Sharkey, "On combining artificial neural nets," *Connection Science*, vol. 8, p. 299~313, 1996.
- [6] S. Gutta, J. Huang, P. J. Phillips, and H. Wechsler, "Mixture of experts for classification of gender, ethnic origin, and pose of human faces," *IEEE Transaction on Neural Networks*, vol. 11, no. 4, p. 948~960, July, 2000.