

소프트 컴퓨팅 기법을 이용한 개인화된 손동작 인식 시스템

A Personalized Hand Gesture Recognition System using Soft Computing Techniques

전문진*, 도준형**, 이상완***, 박광현****, 변종남***

Moonjin Jeon, Jun-Hyeong Do, Sang-Wan Lee, Kwang-Hyun Park and Zeungnam Bien

* 한국항공우주연구원 다목적3호체계팀

** 한국과학기술원 인간친화 복지로봇시스템 연구센터

*** 한국과학기술원 전자전산학과

**** 광운대학교 정보제어공학과

요약

최근 하지가 불편한 노약자나 장애인이 집 안의 다양한 가전기를 손쉽게 제어하기 위한 비전 기반의 손동작 인식 기술이 발전해 왔다. 다수의 사용자가 하나의 손동작 인식 시스템을 사용할 경우 사용자마다 손동작 특성이 모두 다르기 때문에 특정 사용자의 인식률이 저하되는 문제가 발생한다. 또한 동일한 사용자라 하더라도 시간에 따라 손동작 특성이 변화할 수 있다. 사용자마다 다른 손동작 특성은 모델 학습 및 선택 기법을 사용해 효과적으로 다루어질 수 있다. 시간에 따라 변하는 사용자의 특성은 퍼지 개념을 이용해 효과적으로 다루어질 수 있다. 본 논문에서는 다양량 퍼지 의사결정트리를 이용해 사용자 별 인식모델을 만드는 방법을 제시한다. 또한 새로운 사용자가 시스템을 사용할 경우 가장 적합한 모델을 선택해 인식에 사용하고 인식률을 측정한다.

키워드 : 소프트 리모컨 시스템, 퍼지 의사결정트리, 모델 선택, 손동작 인식

Abstract

Recently, vision-based hand gesture recognition techniques have been developed for assisting elderly and disabled people to control home appliances. Frequently occurred problems which lower the hand gesture recognition rate are due to the inter-person variation and intra-person variation. The recognition difficulty caused by inter-person variation can be handled by using user dependent model and model selection technique. And the recognition difficulty caused by intra-person variation can be handled by using fuzzy logic. In this paper, we propose multivariate fuzzy decision tree learning and classification method for a hand motion recognition system for multiple users. When a user starts to use the system, the most appropriate recognition model is selected and used for the user.

Key Words : Soft remote control system, Fuzzy decision tree, Model selection, Hand gesture recognition

1. 서 론

장애인이나 노약자를 위한 서비스의 수요가 증가하면서 보조공학 기술도 발전하고 있다. 한국과학기술원 인간친화복지로봇시스템 연구센터에서는 장애인이나 노약자가 집안의 다양한 가전기를 간단한 손동작으로 제어할 수 있도록 하는 소프트 리모컨 시스템을 개발해왔다[1]. 이 시스템의 성능을 결정하는 중요한 기술 중에 하나로 손동작 인식 알고리즘을 들 수 있다. 하나의 손동작 인식 시스템을 여러 사용자가 사용할 경우 그림 1과 같이 사용자마다 손동작 특성이 모두 다르기 때문에 특정 사용자의 손동작 인식률이 저하되는 문

제가 발생한다. 또한 동일한 사용자라 하더라도 시간에 따라 손동작 특성이 변화할 수 있기 때문에 고정된 결정 경계를 갖는 인식 알고리즘의 인식률이 저하되는 경우가 발생하게 된다.

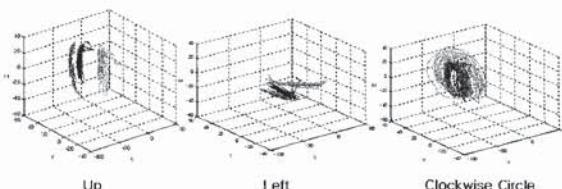


그림 1. 3명의 사용자의 손동작 특성

Fig. 1. Inter-person variation of three hand motions

접수일자 : 2007년 4월 25일

완료일자 : 2007년 11월 21일

감사의 글 : 본 연구는 과학기술부/한국과학재단 우수연구센터 육성사업의 지원으로 수행되었음(R11-1999-008)

인식률 향상을 위해 결정 경계를 늘리면 의도된 손동작에 대한 인식률은 향상되지만 의도되지 않은 동작을 명령으로 인식해 시스템을 오동작 시키는 경우가 자주 발생하게 된다.

이러한 오동작(false positive error)은 손동작 인식 알고리즘을 설계할 때 반드시 고려해야하는 사항이다.

동작 인식을 위한 알고리즘으로 HMM (Hidden Markov Model)이 널리 사용되어 왔다. 수화 인식과 같은 복잡한 동작을 인식하는 경우 HMM이 가장 적합한 인식 알고리즘이라 할 수 있다[2]. 그러나 HMM은 대체로 많은 학습 데이터를 필요로 하고 다른 알고리즘에 비해 분류 능력이 뛰어나지 않다고 알려져 있다[3].

빠른 학습 시간과 뛰어난 인식률의 장점을 갖는 의사결정 트리는 고정된 결정 경계 때문에 손동작과 같은 특성 변화가 심한 데이터에 대해 인식률이 저하되는 단점을 갖는다. 상위 노드에서 오분류된 데이터는 하위 노드의 속성 값이 아무리 정확해도 하위노드에서 분류 결과를 회복할 수 없게 된다. 이러한 점은 학습데이터와 시험데이터의 특성변화가 있는, 즉 손동작과 같이 시간에 따라 변화할 수 있는 데이터일수록 더 심해진다고 할 수 있다. 이러한 단점은 의사결정트리에 폐지 논리를 적용함으로써 해결될 수 있다.

폐지 논리는 수치 데이터의 언어적 모호성을 수학적으로 모델링하고 률 기반의 시스템에서 심볼화된 률을 표현하는데 사용되어 률을 이해하기 쉽도록 한다는 장점이 있다. 의사결정트리에 폐지 논리를 적용한 폐지 의사결정트리[4,5]는 각 노드를 자식 노드로 분기할 때 폐지 소속 함수를 이용한다. 폐지 논리는 불확실성 또는 노이즈가 있는 데이터와 부정확한 데이터를 효과적으로 다룰 수 있기 때문에 손동작과 같은 데이터를 위한 폐지 의사결정트리의 분류능력은 일반 의사결정트리보다 높다고 할 수 있다[6].

폐지 의사결정트리의 학습 과정에서 노드를 분기할 때 하나의 데이터가 두 개 이상의 자식노드로 분기될 수 있는데, 이러한 폐지 의사결정트리의 특성상 전체 트리의 크기가 커지게 된다. 의사결정트리의 크기가 커지면 분류의 일반화 성능에 부정적인 영향을 미치게 된다[7]. 의사결정트리의 구조를 간소화해 일반화 성능을 높이는 방법으로 다변량 분기를 사용해 의사결정트리의 노드를 생성하는 방법이 제시된 바 있다[8]. 본 논문에서는 폐지 의사결정트리의 복잡해지는 노드 구조를 간소화하기 위해 다변량의 개념을 폐지 의사결정트리에 적용한 다변량 폐지 의사결정트리(Multivariate Fuzzy Decision Tree, 이하 MFDT)를 제안하고 MFDT를 이용해 손동작 인식 모델을 만드는 방법을 제시한다.

하나의 인식 모델만을 사용해 다수의 사용자의 손동작을 인식하면 사람마다 손동작 특성이 모두 다르기 때문에 인식률이 낮아질 수 있다. 이러한 점은 사용자별 모델을 만들고 적합한 모델을 선택하는 방법으로 해결할 수 있다. 본 논문에서는 최대 우도 모델 비교 기법을 이용해 새로 입력된 손동작 데이터에 가장 적합한 모델을 선택하고 손동작 인식에 사용한다.

본 논문의 전체적인 구조는 다음과 같다. 2장에서는 개인화된 손동작 인식 시스템에 대해 설명하고, 3장에서는 MFDT의 학습 및 분류방법을 설명한다. 4장에서는 모델 선택에 대해 설명하고, 5장에서는 벤치마크 데이터와 손동작 데이터를 이용한 실험 결과에 대해 설명한다. 6장에서 결론 및 추후과제에 관해 고찰하며 논문을 마치도록 한다.

2. 개인화된 손동작 인식 시스템

다양한 사용자의 손동작을 인식하기 위해 하나의 인식기를 사용할 때 특정 사용자에 대해 손동작 인식률이 낮아지는

점은 사용자 별 인식 모델을 만들고 선택된 모델을 인식기로 사용하는 방법으로 해결될 수 있다. 그림 2에는 개인화된 손동작 인식시스템이 손동작을 인식하는 과정이 나타나있다.

먼저 두 대의 카메라에 입력된 이미지를 이용해 얼굴과 손의 3차원 좌표를 계산하고 얼굴을 중심으로 손의 상대 좌표를 계산해 손의 궤적을 얻는다. 사용자가 손동작을 이용해 명령을 하는 동안의 손의 이동 속도가 평상시의 손의 움직임 보다 빠르다고 가정하면 속도 임계값을 이용해 의미 있는 손동작 구간을 구분할 수 있다. 이 구간의 손 궤적을 이용해 손동작을 잘 표현하는 특징을 추출한다. 추출된 특징 값을 다변량 폐지 의사결정 트리를 이용해 10가지 손동작 또는 일상 동작으로 분류하고 state automata를 이용해 손동작 후 뒤따르는 손동작을 무효화한다. 10가지 손동작 중 하나로 인식되었을 경우 가전기기를 제어하는 신호를 발생시킨다.

사용자가 시스템을 처음 사용하는 경우는 특정 추출 후 최대 우도 모델 비교 기법을 이용한 모델 선택 과정을 통해 적합한 모델을 선택해 선택된 모델을 손동작 인식에 사용한다.

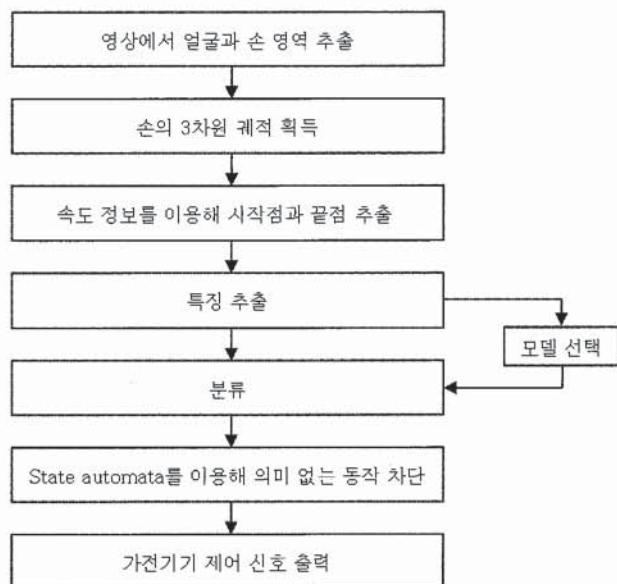
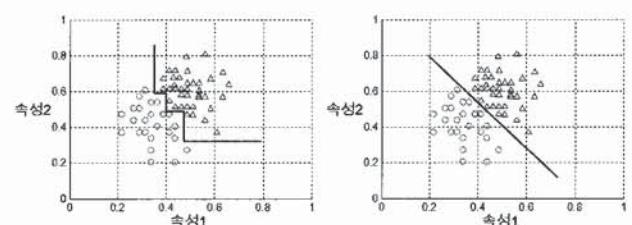


그림 2. 개인화된 손동작 인식 시스템의 구조

Fig. 2. Structure of the personalized hand gesture recognition system

3. 다변량 폐지 의사결정 트리

일반적인 폐지 의사결정 트리가 각 노드를 분기할 때 하나의 속성을 사용하는 반면 MFDT는 속성 벡터를 사용한다. 그림 3에는 단일 속성을 사용한 분기와 속성 벡터를 사용한 분기가 비교되어 있다.



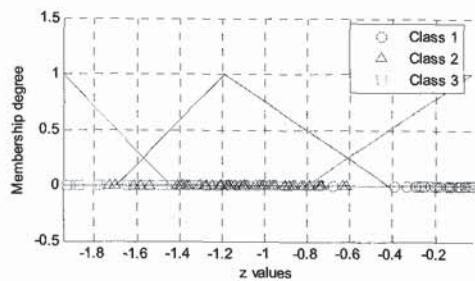


그림 5. 자동으로 생성된 퍼지 소속 함수
Fig. 5. Automatically generated fuzzy membership function

학습 과정을 통해 그림 6과 같은 MFDT 모델이 생성된다. 이 MFDT는 Iris data를 5x2 교차 검증하는 과정에서 생성된 모델이다.

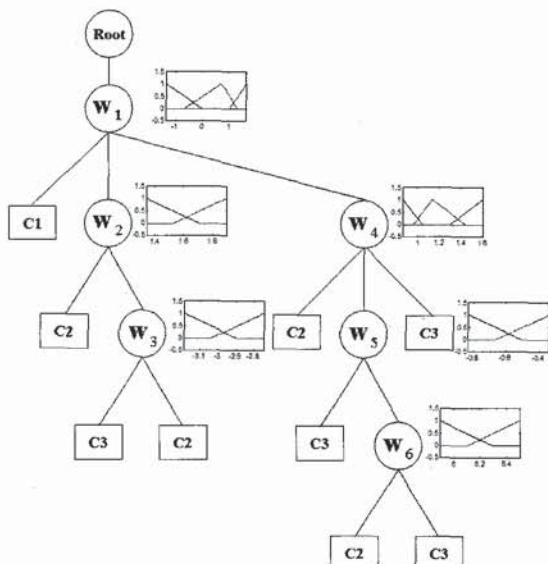


그림 6. 학습된 MFDT
Fig. 6. Trained MFDT

3.2 다변량 퍼지 의사결정트리 분류

생성된 MFDT 모델을 이용해 새로 입력된 데이터를 분류하는 방법은 다음과 같다.

Step 1. 근 노드에서 각 단말노드 까지 거쳐 가는 노드 소속 함수와 속성 벡터를 이용해 T-norm을 구한다.

$$z_i = \mathbf{w}_i^T \mathbf{x}$$

$$T_n = \prod_{i=\text{rootnode}}^{n_{\text{th}} \text{ leaf node}} \mu_{s_{\text{th}}}^i(z_i)$$

\mathbf{x} : 분류할 데이터

Step 2. 가장 큰 T-norm을 갖는 단말노드의 클래스로 입력 데이터를 분류

$$\text{class } C = \arg \max_n T_n$$

4. 모델 선택

새로운 사용자가 손동작 인식 시스템을 사용하려고 할 때 그 사용자에게 가장 적합한 인식 모델을 추천하는 모델 선택 과정이 필요하다. 다수의 사용자로부터 수집한 학습데이터와 다변량 퍼지 의사결정트리 학습 알고리즘을 이용해 각 사용자의 모델을 학습한다. 학습된 모델은 그림 7과 같은 모델 풀에 저장되고 새로운 사용자의 데이터가 입력되면 입력된 데이터와 가장 유사한 모델이 선택된다.

입력된 데이터와 가장 유사한 인식모델을 선택하기 위해 적합한 유사도 측정 방법이 필요하다. 일반적으로 모델과 모델사이의 유사도는 FSM(Fuzzy Similarity Measure)[9] 등의 방법을 이용해 측정할 수 있고, 데이터와 데이터 사이의 유사도는 거리, cosine 등을 이용해 측정할 수 있다. 그러나 모델과 데이터 사이의 유사도는 직접적인 측정이 어렵기 때문에 본 논문에서는 간접적인 방법으로 유사도를 평가한다.

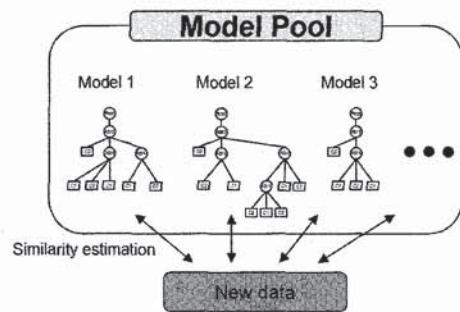


그림 7. 모델 선택
Fig. 7. Model selection

모델 선택을 위한 학습 데이터가 입력되면 모델 풀의 모든 모델을 이용해 분류 테스트를 하고 최대 우도 모델 비교 (Maximum likelihood model comparison)[10]에 근거해 모델을 선택한다.

$$P(m_i | D) = \frac{P(D | m_i)p(m_i)}{p(D)} \propto P(D | m_i)p(m_i)$$

$$\begin{aligned} \text{Selected Model } \hat{m} &= \arg \max_{m_i} P(D | m_i) \\ D: \text{data}, m_i: i_{\text{th}} \text{ model} \end{aligned}$$

5. 실험 및 결과

5.1 벤치마크 데이터를 이용한 인식률 측정

제안된 MFDT 학습 및 분류 알고리즘의 일반적 성능을 측정하기 위해 벤치마크 데이터인 UCI Machine Learning Repository[11]의 Iris data와 Wine data를 사용했다. 데이터의 특징과 MFDT를 이용한 인식률을 표 1과 표 2에 나타내었다. 인식률은 5x2 fold cross validation의 평균 인식률이다.

표 2를 통해 MFDT의 학습 및 분류가 벤치마크 데이터에 높은 성능을 가진다는 것을 알 수 있다.

참 고 문 헌

- [1] Jun-Hyeong Do, Hyoyoung Jang, Sung Hoon Jung, Jinwoo Jung, Zeungnam Bien, "Soft Remote Control System in the Intelligent Sweet Home", *Proc. of IEEE Int. conf. on IROS*, pp. 3984-3989, 2005
- [2] Sung Hoon Jung, *Incremental User Adaptation in Korean Sign Language Recognition Using Motion Similarity and Prediction from Adaptation History*, Master's Thesis, KAIST, 2007.
- [3] Herve Bourlard, "Links Between Markov Models and Multilayer Perceptrons", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.12, No.12, pp. 1167-1178, Dec. 1990.
- [4] Cezary Z. Janikow, "Fuzzy Decision Trees : Issues and Methods", *IEEE Transactions on systems, man, and cybernetics*, Vol. 28, No. 1, pp. 1-14, Feb 1998
- [5] Koen-Myung Lee, Kyung-Mi Lee, Jee-Hyong Lee, Hyung Lee-Kwang, "A Fuzzy Decision Tree Induction Method for Fuzzy Data", *1999 IEEE International Fuzzy Systems Conference Proceedings*, Vol. 1, pp. 16-21, August 22-25, 1999, Seoul, Korea
- [6] Moonjin Jeon, Jun-Hyeng Do, Sang-Wan Lee, Kwang-Hyun Park, Zeungnam Bien, "Hand Motion Recognition using Fuzzy Decision Tree", *Proceedings of the 8th International Workshop on Human-friendly Welfare Robotic Systems*, 21-23 October 2007, pp. 123-127
- [7] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, Wadsworth International Group, Belmont, CA, 1984.
- [8] Olcay Taner Yildiz and Ethem Alpaydin, "Linear Discriminant Trees", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 19, No. 3 pp. 323-353, 2005
- [9] C.T. Lin, *Neural Fuzzy Control Systems with Structure and Parameter Learning*, World Scientific, 1994.
- [10] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, 2nd edition, John Wiley & Sons, 2001.
- [11] C. J. Merz and P. M. Murphy, "UCI repository for machine learning data-bases," <http://www.ics.uci.edu/mlearn/MLRepository.html> Department of Information and Computer Science. University of California, Irvine, 1996.

저 자 소 개



전문진(Moonjin Jeon)

2006년 : 한양대학교 전자전기컴퓨터공학부
학사졸업
2008년 : 한국과학기술원 전자전산학과
석사 졸업
2008년 ~ 현재 : 한국항공우주연구원 다목적
3호체계팀 연구원

관심분야 : 퍼지 이론, 패턴인식, 기계학습
E-mail : mjjeon@kari.re.kr



도준형(Jun-Hyeong Do)

1999년 : 한국과학기술원 전자전산학과
학사 졸업
2001년 : 한국과학기술원 전자전산학과
석사 졸업
2008년 : 한국과학기술원 전자전산학과
박사 졸업
2008년 ~ 현재 : 한국과학기술원 인간친화
복지로봇시스템 연구센터 선임급 위촉연구원

관심분야 : 인간-로봇 상호작용/인터페이스, 컴퓨터 비전, 지
능 로봇
E-mail : jhdo@ctrlsys.kaist.ac.kr



이상완(Sang-Wan Lee)

2003년 : 연세대학교 기계전자공학부
학사 졸업
2005년 : 한국과학기술원 전자전산학과
석사 졸업
2005년 ~ 현재 : 한국과학기술원 전자전산학
과 박사과정

관심분야 : 학습이론, 패턴인식, 특징추출
E-mail : bigbean@ctrlsys.kaist.ac.kr



박광현(Kwang-Hyun Park)

1994년 : 한국과학기술원 전자전산학과
학사 졸업
1997년 : 한국과학기술원 전자전산학과
석사 졸업
2001년 : 한국과학기술원 전자전산학과
박사 졸업
2005년 ~ 2008년 : 한국과학기술원 전자전산
학과 BK초빙교수
2008년 ~ 현재 : 광운대학교 정보제어공학과 조교수

관심분야 : 학습이론, 지능로봇, 인간-로봇 상호작용, 재활공
학
E-mail : akaii@robotian.net



변증남(Zeungnam Bien)

1969년 : 서울대학교 전자공학과 학사 졸업
1972년 : University of Iowa 전자공학과
석사 졸업
1975년 : University of Iowa 수학과 석사
졸업
1975년 : University of Iowa 전자공학과
박사 졸업
1976년 ~ 1977년 : University of Iowa 조교수
1981년 ~ 1982년 : University of Iowa 객원부교수
1987년 ~ 1988년 : Syracuse University 객원 연구원
1988년 : 일본 동경공대 객원교수
1990년 ~ 1995년 : 한국퍼지및 지능시스템학회 회장
2001년 : 대한전자공학회 회장
2003년 ~ 2005년 : 국제퍼지시스템학회(IFSA) 회장
2003년 ~ 2006년 : 한국로봇공학회 회장
1977년 ~ 현재 : 한국과학기술원 전자전산학과 교수
1999년 ~ 현재 : 한국과학기술원 인간친화복지로봇시스템 연구센터 소장
2005년 ~ 현재 : 한국전력 석좌교수

관심분야 : 지능제어, 학습이론, 소프트컴퓨팅, 서비스로봇,
재활공학

E-mail : zbien@ee.kaist.ac.kr