

퍼지 최대-최소 신경망을 이용한 특징 집합 선택에 관한 연구 및 보행 단계인식에의 응용

A Study on Feature selection based the Fuzzy Min-Max Neural Network and Application on Gait Phase recognition using EMG

이태엽¹, 이상완², 변증남³

¹ 한국과학기술원 전자전산학부

E-mail: tyyi@ctrsys.kaist.ac.kr

² 한국과학기술원 전자전산학부

E-mail: bigbean@ctrsys.kaist.ac.kr

³ 한국과학기술원 전자전산학부

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

요 약

본 논문은 패턴 분류 문제에 사용되는 퍼지 최대-최소 신경망 방법을 이용하여 특징 집합으로부터 새로운 특징 집합을 추출해내고 추출된 특징 집합으로부터 의미 있는 특징을 선택해 내는 새로운 방법을 제안한다. 퍼지 최대-최소 신경망은 패턴 분류를 위해 주로 사용이 되어 왔지만, 퍼지 최대-최소 신경망을 이용해 특징 집합의 값들을 패턴 공간내의 초상자의 집합으로 변환하고 변환된 초상자들끼리의 인접성을 척도로 단순한 연산을 통한 빠른 특징 집합을 선택하게 된다. 마지막으로 본 논문의 특징 집합 선택 방법을 하지 근전도 신호를 이용한 보행 패턴 분류에 적용해 보고, 그 결과를 기존 여러 특징 집합 선택 방법들과 비교해 봄으로써 제안한 방법의 타당성 및 적용 가능성을 알아본다.

Key Words : FMMNN, Feature Selection, Gait Phase Recognition

1. 서 론

일반적으로 패턴 분류 과정이라 함은 데이터 취득, 데이터의 전처리 과정, 입력 데이터로부터 특징 집합의 추출 및 선택, 마지막으로 패턴 분류의 과정을 말한다. 이때 추출한 여러 특징 집합들 중에서 의미 있는 특징 집합만을 선택함으로써 패턴 분류 시 보다 빠르고 강인하며 높은 인식을 보이는 패턴 분류기를 설계할 수 있다.

특징 집합의 선택 방법에는 크게 필터(filter) 방법과 랩퍼(wrapper) 방법으로 나누어지며 [1][2], 필터방법은 데이터들끼리의 거리[2][3]나 통계적인 의존성[2][4] 또는 상관성을 척도로 특징 집합들의 순위를 매기고 이로부터 필요한 특징 집합을 선택할 수 있다. 반면 랩퍼 방법의 경우에는 기계학습 방법을 사용하여 특정 특징 집합들에 대한 인식의 정도를 비교하여 높은 인식을 보이는 특징 집합을 선택하게 된다[2]. 보통 필터 방법의 경우 반복적인 계산 과정이 필요 없어 빠르며 일반적인 특징

집합을 선택해 냄으로써 과최적화(over-fitting) 문제를 극복할 수 있다[2]. 이에 반해 랩퍼 방법은 기계학습 방법을 이용해 데이터와 인식기 사이에서 높은 인식률을 보이는 특징 집합들을 선택하기 때문에 필터 방법에 비해 보다 나은 인식 결과를 보인다[2]. 하지만 계산의 복잡성과 특정 인식기에 최적화된 특징 집합을 선택하기 때문에 일반적인 특징 집합을 선택해 내지는 못한다.

본 논문에서는 그 동안 패턴분류에서 많이 사용되어 온 퍼지 최대·최소 신경망[5] 패턴 분류기를 이용한 특징 집합 선택방법을 제안한다. 퍼지 최대·최소 신경망을 통해 특징 집합을 패턴 공간내의 초상자의 집합으로 변환하고 변환된 초상자들끼리의 인접성을 척도로 특징 집합을 선택해 냄으로써 단순한 연산을 통한 빠른 특징 집합을 선택할 수 있는 특징이 있다.

전체 논문의 구성은 먼저 2장에서 퍼지 최대·최소 신경망에 관해 간단히 알아본다. 3장에서는 퍼지 최대·최소 신경망을 이용한 새로운 특징 집합 선택 방법을 소개하고 4장에서

하지 근전도 신호를 이용한 보행 패턴 분류 방법에 본 논문의 특징 집합 선택 방법을 적용해 본다. 그리고 기존의 특징 집합 선택 방법들과 결과를 비교해 봄으로써 제안한 방법의 타당성과 적용 가능성을 알아본 후 5장의 결론을 끝으로 본 논문을 마무리한다.

2. 퍼지 최대·최소 신경망^[5]

퍼지 최대·최소 신경망(Fuzzy Min-Max Neural Network, 이하 FMMNN이라고 함)은 1992년 Simpson에 의해 제안된 패턴 분류 방법이다. 이 방법에서는 입력 패턴 값들을 n차원의 패턴 공간 내에서 정의 되는 초상자(hyper-box)를 사용하여 나타내고 이 초상자들은 최소값, 최대값 그리고 이 최소값과 최대값으로 표현되는 멤버십 함수로 구성된다.

학습과정에서 입력 패턴에 대응하는 초상자들의 확장, 중복, 축소과정을 거쳐 최종 초상자들의 집합을 얻을 수 있다. 그리고 테스트 입력 패턴이 인가되면, 각 초상자들의 멤버십 함수를 이용하여 인가된 테스트 패턴과 초상자의 일치정도를 멤버십 값으로 구할 수 있으며, 가장 높은 멤버십 값을 갖는 초상자를 찾아 그 초상자의 클래스를 선정함으로써 패턴분류를 할 수 있게 된다.

표 1. FMMNN의 학습과정.

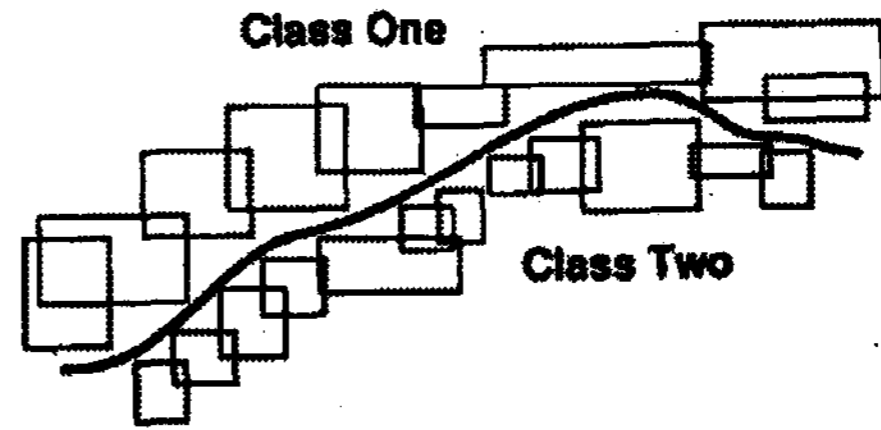
| |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>확장(expansion) - 입력 패턴이 인가되었을 경우 같은 클래스를 가지는 초상자들 중에서 확장기준(expand criterion)을 만족하는 초상자와 확장, 만약 확장할 초상자가 없다면 새로운 초상자 생성</p> <p>중복(overlap) - 확장된 초상자가 다른 클래스의 초상자들과 겹침이 있는지를 검사</p> <p>축소(contraction) - 중복 검사에서 서로 다른 클래스에 속하는 초상자들과 겹침 있을 경우 초상자를 최소한으로 변화시켜 중복을 없앴.</p> |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

퍼지 최대·최소 신경망은 비선형 분리 및 클래스의 중복이 가능하여 비선형 패턴 분류가 가능^[5]하며, 비모수적 방법이기 때문에 입력 데이터의 확률분포를 모르는 경우에도 확률분포에 대한 가정 없이 적용을 할 수 있다^[5]. 게다가 대부분의 비선형 패턴 분류기들이 학습을 위해 많은 시간이 소요되지만, FMMNN의 경우 빠르고 효율적인 특성을 보이기 때문에^[5] 실시간 패턴 분류 시스템에 많이 사용이 되고 있다.

3. FMMNN을 이용한 특징 집합의 선택방법

입력 데이터 패턴 $X = \{X_1, X_2, \dots, X_m\}$, $X_h = \{x_{h1}, x_{h2}, \dots, x_{hn}\} \in I^n$ 는 FMMNN의 학습과정에서 초상자의 최소값 행렬 $V = \{V_1, V_2, \dots, V_p\}$, $V_j = \{v_{j1}, v_{j2}, \dots, v_{jm}\} \in I^n$ 과 최대값 행렬 $W = \{W_1, W_2, \dots, W_p\}$, $W_j = \{w_{j1}, w_{j2}, \dots, w_{jm}\} \in I^n$ (단, $p \leq m$)로 패턴 공간 내에서 나타내어진다. FMMNN은 비모수적인 특성을 가지고 있고 중복성을 인정하기 때문에 입력 데이터 패턴을 초상자들의 집합으로 나타내더라도 초상자들끼리의 경계를 이용해 주어진 데이터의 분포를 반영할 수 있다(그림1). 또한 초상자들은 여러 입력 데이터들을 내포할 수 있기 때문에 데이터의 정보를 최대한 유지하면서 작은 수의 초상자들로 데이터를 표현할 수 있는 장점이 있다.

그림 1. 초상자들의 경계 형성^[5]



일반적으로 데이터의 분리도를 측정함에 있어서 같은 클래스의 데이터끼리는 인접하게 분포하고 다른 클래스 간에는 거리가 크게 분포할수록 큰 분리도를 얻을 수 있다. 이는 서로 다른 클래스의 데이터가 인접하게 분포할수록 데이터를 분류함에 오차가 커짐을 의미하며, 아래의 예를 통해 이를 확인할 수 있다.

그림 2. 두 초상자의 거리가 먼 경우 테스트 데이터에 대한 각 초상자의 멤버십 값 비교

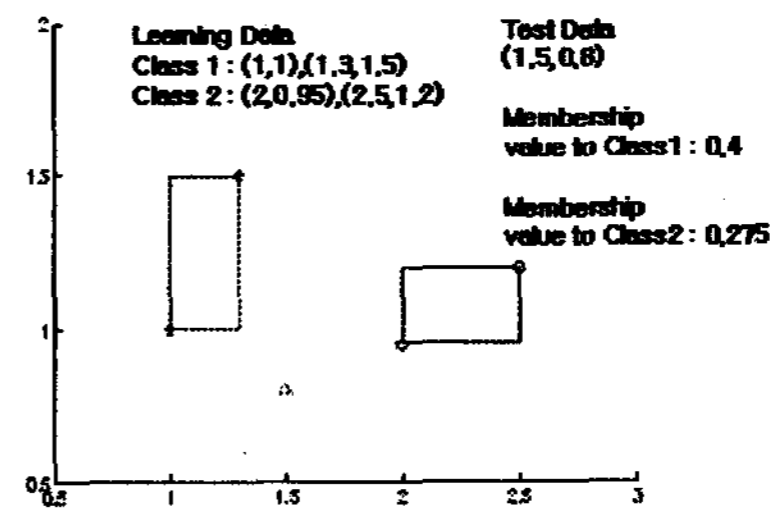
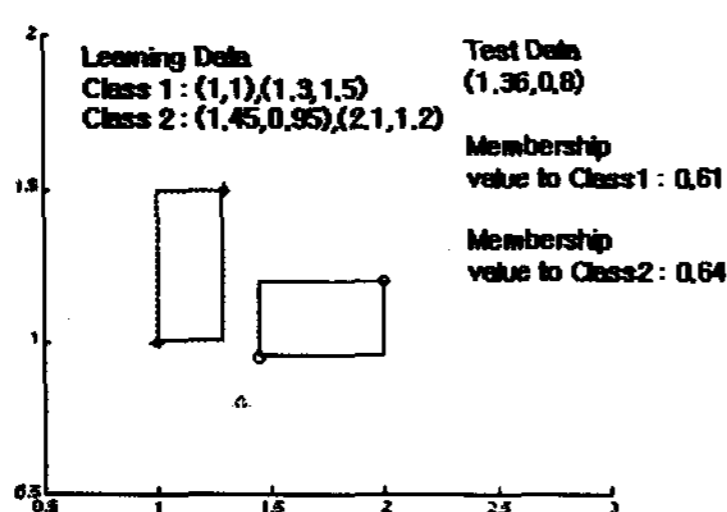


그림 2에서 볼 수 있듯이 2차원의 패턴 공간에서 두 초상자의 거리가 먼 경우, (1.5, 0.8)의 테스트 패턴이 인가되면 클래스1 초상자는 0.4의 멤버십 값을 가지고 반면, 클래스2 초상자는 0.275의 멤버십 값을 가지게 된다. 이 경우 테스트 패턴은 멤버십 값이 큰 클래스1을 선택하게 된다.

그림 3. 두 초상자의 거리가 가까운 경우 테스트 데이터에 대한 각 초상자의 멤버쉽 값 비교



하지만 그림 2에서처럼 테스트 패턴이 x축에 대해 클래스1의 초상자에 더 가깝게 인가되지만 두 초상자가 인접한 경우(그림3) y축의 멤버쉽 값이 더 크게 작용을 하게 된다. 그래서 클래스1 초상자에 대한 테스트 패턴의 멤버쉽 값 $((0.82+1.22)/2 = 0.61)$ 보다 클래스 2 초상자에 대한 멤버쉽 값 $((0.73+1.28)/2 = 0.64)$ 이 더 크게 되어 테스트 패턴은 클래스 2로 분류가 된다.

본 논문에서는 서로 다른 클래스에 속한 초상자들끼리의 인접성을 구하여 인접성이 작은 특징들의 집합을 선택함으로써 데이터의 분류 오차는 줄이고 데이터의 분리도는 높일 수 있는 방향으로 특징 집합을 선택하려 한다.

제안하는 방법은 표2와 같이 요약할 수 있다.

표 2. FMMN을 이용한 특징 집합의 선택과정.

FMMN 학습 결과 초상자의 최소값 행렬 V와 초상자의 최대값 행렬 W를 얻을 수 있다.

Step1. V와 W로부터 평균값 행렬 M_c 를 얻는다.
 $M_c = \text{average}(V, W)$

Step2. i번째 차원에서 j번째 초상자의 클래스와 다른 가장 인접한 초상자와 거리를 구한다.

Repeat i=1,2,...패턴의 차원(N)
 Repeat j=1,2,...초상자의 수(cN)
 j번째 초상자의 클래스와 일치하지 않는 초상자의 집합을 구한다.

$$M_c^i = \{M_{ck}^i \mid M_{ck}^i \in M^i, Y_k \neq Y_j, k=1,2,\dots,sN, k \neq j\}$$

, Y: 초상자의 클래스, sN: 선택된 초상자수

$$D^i = \{(M_j^i - M_1^i), (M_j^i - M_2^i), \dots, (M_j^i - M_k^i)\}$$

, k=1,2,...,sN, k ≠ j

$$d_j^i = \min(D_1^i, D_2^i, \dots, D_{sN}^i)$$

endRepeat

$$r^i = \sum_{j=1}^{cN} d_j^i$$

endRepeat

Step3. r^i 를 큰 값 순으로 순서 짓기

먼저 선택된 특정 차원에 대해서 초상자들의

최대·최소값 행렬로부터 평균 행렬 M_c 를 구한다. 그리고 초상자의 수(cN)만큼 반복을 하면서 선택된 초상자와 다른 클래스의 초상자들에 대해서 두 초상자들 간의 거리(D^i)를 구한다. 이 과정에서 임의의 특정 차원에 대해 각 초상자들 사이의 가장 짧은 거리($d_j^i = \min(D_1^i, D_2^i, \dots, D_{sN}^i)$)를 구할 수 있다. 그리고 각 초상자별로 구해진 거리의 합($r^i = \sum_{j=1}^{cN} d_j^i$)으로 그 차원에서 초상자들끼리의 인접성을 구한다. 이를 전체 차원만큼 반복하게 되면 각 차원별 초상자들끼리의 인접성을 구할 수 있으며 이 거리의 합이 큰 순서별로 특징 집합을 순서 짓게 된다. 그리고 필요한 수만큼의 차원을 선택함으로써 특징 집합을 구성할 수 있다.

초상자들끼리의 단순한 연산을 통해 인접성을 구할 수 있기 때문에 빠르게 특징 집합을 선택할 수 있는 장점이 있다. 이는 실시간 시스템에서의 특징 집합 선택이나 고차원 데이터 패턴에 대한 특징 집합의 선택에 유용할 수 있겠다.

4. 실험 결과 및 분석

본 논문의 타당성 및 적용가능성을 알아보기 위해 하지 근전도 신호를 이용한 보행 패턴 분류에 본 논문에서 제안하는 특징 선택 방법을 적용해보고 기존의 특징 선택방법과 비교를 해보았다.

하지 근전도 신호를 이용한 보행 패턴 분류 방법[7]에서는 그림 4에서 보이는 근육의 근전도 신호로부터 보행의 각 단계를 분류해내게 된다. 이때 보행 단계는 크게 4개 단계로 나누어진다. 그리고 측정된 근전도 신호로부터 근전도 신호에 많이 이용되는 14개의 특징 값 추출 방법[7]을 이용하여(표3) 총 14개, 424 차원의 특징 값들을 얻을 수 있다.

그림 4. 하지 근전도 측정 부위^[7]

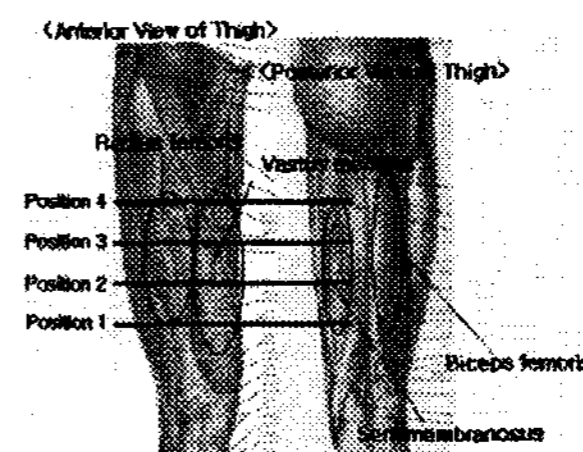


표3의 특징 추출 방법으로 얻은 424차원의 특징 값들에 대하여 본 논문에서 제안한 방법을 이용하여 특징들을 추출하고 추출된 특징들을 이용해 보행 단계 구분을 실시하였다.

표 3. 보행 패턴 분류에 사용된 특징 추출 방법들^[7].

| | 차원 | 특징추출방법 | | 차원 | 특징추출방법 |
|---|----|----------------------------|----|-----|-------------------------------------|
| 1 | 8 | IAV | 8 | 8 | MF (mean Frequency) |
| 2 | 8 | VAR (Variance) | 9 | 168 | HIST |
| 3 | 8 | WAMP (Wilson Amplitude) | 10 | 24 | AR (Auto-regressive coefficient) |
| 4 | 8 | ZC (zero crossing) | 11 | 24 | ARCU |
| 5 | 8 | NT (Number of Turn) | 12 | 40 | EWT |
| 6 | 8 | MA (Mean of Amplitude) | 13 | 64 | EWP |
| 7 | 8 | Wave Length | 14 | 40 | ZCWT |

실험은 MatLab을 이용하여 총 232개의 데이터 샘플을 임의로 학습데이터와 테스트 데이터로 분류하여 30회의 반복 실험을 실시하였고 패턴을 분류하기 위한 인식기로는 FMMNN을 사용하였다.

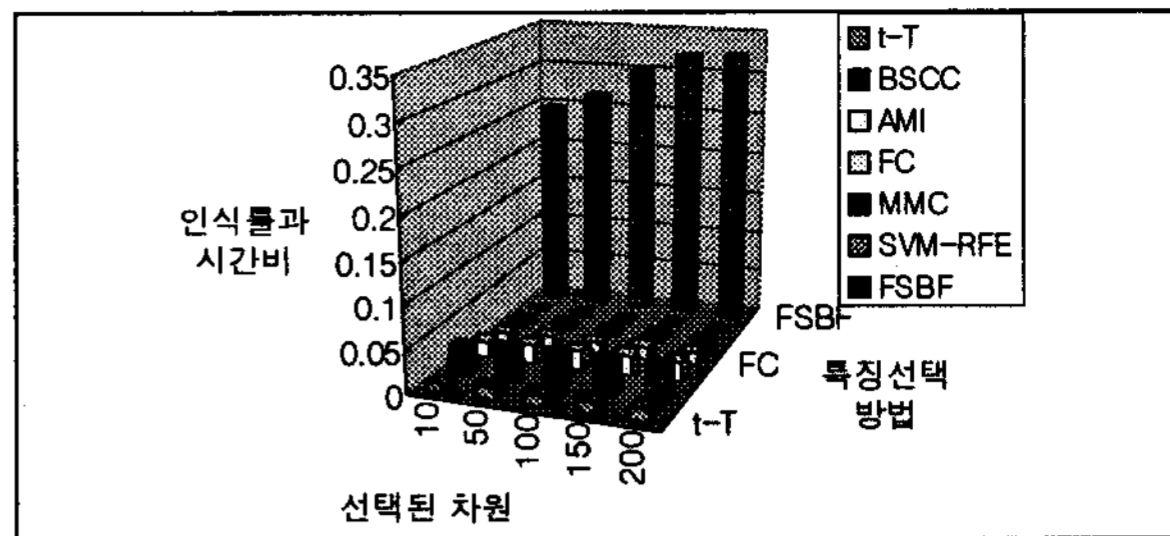
표 4. 여러 특징 집합 선택 방법에 대한 분류 결과

| | 시간 | 선택된 차원 | | | | |
|---------------|-------|--------|-------|-------|-------|-------|
| | | 10 | 50 | 100 | 150 | 200 |
| Random Select | NA | 54.36 | 66.45 | 74.79 | 78.72 | 79.83 |
| t-T | 54.24 | 63.33 | 77.9 | 79.23 | 79.53 | 81.23 |
| BSCC | 21.07 | 67 | 76.11 | 79.53 | 81.19 | 83.76 |
| AMI | 33.24 | 61.7 | 75.59 | 73.24 | 78.41 | 80.08 |
| FC | 60.36 | 59.06 | 74.57 | 75.72 | 79.35 | 81.28 |
| MMC | 94.11 | 63.24 | 65.89 | 70.72 | 75.72 | 80.25 |
| SVM-RFE | 1445 | 62.82 | 73.54 | 77.82 | 80.89 | 82.05 |
| FSBF | 2.54 | 62.65 | 67.52 | 76.28 | 81.25 | 82.44 |

Random Select: 임의의 차원 선택, t-T: 필터 방법의 t-Test^[8] 방법, BSCC: 필터 방법의 bi-serial correlation coefficient^[8], AMI: 필터 방법의 approx. mutual information, FC: 필터 방법의 fisher criterion^[8], MMC: 래퍼 방법 중 Maximum Margin Criterion^[9], SVM-RFE: 래퍼 방법 중 SVM-Recursive Feature Elimination 방법^[10], FSBF: Feature Selection Based on FMMNN: 본 논문에서 제안하는 특징 선택 방법

위의 결과에서 볼 수 있듯이 제안하는 특징 선택 방법(FSBF)은 필터 방법의 평균 수행시간인 42.23초 보다 약 16배 빠르면서도 100 이상의 차원에 대해서 기존의 특징 선택 방법과 비슷하거나 나은 인식률을 보였다. 이를 시간과 인식률의 비($\frac{Recognition Rate}{t} \times \frac{1}{100}$)로 나타내었을 때 그림 5에서처럼 제안하는 방법이 다른 방법에 비해 평균 18배 높은 결과를 보임을 확인할 수 있었다.

그림 5. 인식률과 시간의 비 그래프



5. 결론

본 논문에서는 FMMNN를 이용하여 입력 데이터 패턴을 패턴 공간에서의 초상자들의 집합으로 변환하고 변환된 초상자들의 인접성을 척도로 특징 집합을 선택하는 방법을 제시하였다. 결과에서 확인할 수 있듯이 제안하는 방법은 빠른 수행시간에도 불구하고 100차원 이상에서 기존의 다른 특징 선택 방법과 비슷하거나 나은 인식 결과를 보여 줌을 확인할 수 있었다. 데이터의 차원이 커질수록 특징 집합 선택에 필요한 시간이 커짐을 고려할 때 입력 채널이 많은 근전도 신호나 뇌파 신호를 이용한 패턴 분류나 또는 고차원 데이터의 패턴 분류에 본 방법을 적용함으로써 보다 빠른 특징 집합을 선택할 수 있으리라 본다.

참 고 문 헌

[1] Isabelle Guyon and Andre Elisseeff, "An Introduction to Variable and Feature Selection," Journal of Machine Learning Research 3, 2003.
 [2] Jeong-su Han, "New Feature Subset Selection Method and Its Application for EMG Recognition", Doctoral thesis of KAIST, 2006.
 [3] J. Kittler. "Feature selection and extraction in Handbook of Pattern Recognition and Image Proc.," T.Y.Young and K.S.Fu, Eds. San Diego, CA: Academic, pp. 59-83, 1986
 [4] Liu, Huan., Motoda and Hiroshi., "Feature Selection for Knowledge Discovery and Data Mining," Kluwer academic publishers, Boston, 1998.
 [5] Patrick K. Simpson, "Fuzzy Min-Max Neural Networks - Part 1: Classification," IEEE Trans. on Neural Network, Vol. 3, No. 5, 1992.
 [6] Bogdan Gabry and Andrzej Bargiela, "General Fuzzy Min-Max Neural network for Clustering and Classification," IEEE Trans. on Neural Networks, Vol. 11, No. 3,

2000.

- [7] San Wan Lee, Taeyoub Yi, Jeong-Su Han, Hyoyoung Jang, Heon-Hui Kim, Jin-Woo Jung, Zeungnam Bien, "Walking Phase Recognition for the Lower Limb Disability: Sensing, Feature Extraction, Feature Selection, and Classification," 10th IEEE International Conference on Rehabilitation Robotics (ICORR2007), Noordwijk, Netherlands, June 13-15, 2007
- [8] K.-R. Muller, M. Krauledat, G. Dornhege, G. Curio, B. Blankertz, "Machine Learning Techniques for Brain-Computer interfaces,"
- [9] Satoshi Nijima and Satoru Kuhara, "Recursive gene selection based on maximum margin criterion: a comparison with SVM-RFE," BMC Bioinformatics 2006, 7:543, 2006.
- [10] Te-Ming Huang, Vojislav Kecman, Ivica Kopriva, "Kernel Based Algorithms for Mining Huge Data Sets," Springer