

# 감정 인식을 위한 Interactive Feature Selection(IFS) 알고리즘

## Interactive Feature selection Algorithm for Emotion recognition

양현창 · 김호덕 · 박창현 · 심귀보

Hyun-Chang Yang, Ho-Duck Kim, Chang-Hyun Park, and Kwee-Bo Sim

중앙대학교 전자전기공학부

### 요 약

본 논문은 일반적으로 많은 특징들을 갖고 있는 패턴 분류 문제인 감정 인식을 위한 새로운 특징 선택 방법을 제안한다. '특징 선택'은 패턴 인식 성능의 향상에 기여하고 '차원의 저주'문제에도 좋은 해결책으로 많이 사용되는 방법이다. 그래서, 본 논문에서는 강화학습의 개념을 사용한 상호 작용에 의한 특징 선택 방법인 IFS(Interactive Feature Selection)를 고안하였고 이 알고리즘을 사용하여 선택된 특징들을 감정 인식 시스템에 적용하여 성능이 향상됨을 확인하였다. 또한, 기존의 특징 선택 방법과의 비교를 통하여 본 알고리즘의 우수성을 확인하였다.

### Abstract

This paper presents the novel feature selection method for Emotion Recognition, which may include a lot of original features. Specially, the emotion recognition in this paper treated speech signal with emotion. The feature selection has some benefits on the pattern recognition performance and 'the curse of dimension'. Thus, We implemented a simulator called 'IFS' and those result was applied to a emotion recognition system(ERS), which was also implemented for this research. Our novel feature selection method was basically affected by Reinforcement Learning and since it needs responses from human user, it is called 'Interactive Feature Selection'. From performing the IFS, we could get 3 best features and applied to ERS. Comparing those results with randomly selected feature set, The 3 best features were better than the randomly selected feature set.

Key words : Reinforcement Learning, Feature selection, Emotion Recognition, Speech Signal

## 1. 서 론

감정 인식 연구는 크게 4가지 매체에 대해 시도되어 왔다. 음성, 영상, 생체신호, 몸짓이 그 4가지 매체이고 1990년부터 2005년까지 IEEE 논문으로 출간된 감정 인식 연구 결과로부터 통계를 내본 결과, 음성에 대한 연구가 가장 많았고, 그 다음으로 음성과 영상을 조합한 연구, 그리고 생체 신호와 몸짓에 대한 연구가 소수 보여 짐을 확인할 수 있다. 이러한 분위기는 음성과 영상에 대한 신호의 추출이 생체신호나 몸짓의 추출보다 용이하고 신호의 분류가능성도 더 좋기 때문인 것으로 추정된다.

특히, 생체 신호로는 EEG, ECG, SC 센싱 결과를 주로 사용하는데, 이 신호들을 입력받을 때 피험자의 움직임이나 주변 전자기기에 의한 간섭으로 인해 정확하고 깔끔한 신호를 수집하기가 어렵다는 점이 생체신호를 이용한 감정 인식 연구의 걸림돌로 작용하고, 영상의 경우는 결국 표정인식을 말하는 것인데, 이 경우 영상인식의 일반적인 문제점인 불빛에 따른 인식 성능의 차이와 얼굴 인식에서의 문제점인 안경 등 기타 장신구를 사용했을 때의 인식 성능에 대한 문제가 있다. 몸짓 인식의 문제점도 표정인식에서의 문제점에서 크게

벗어나지 않고 있고 특히 더 큰 문제는 몸짓에 포함된 감성 정보 자체가 충분하지 않다는 것이다.

상기의 4가지 매체 중 3가지 매체의 이러한 문제점과는 달리 음성의 경우 전화상으로도 상대방의 감성을 어느 정도 인식할 수 있다는 점에서 음성정보에 포함된 감성 정보의 유용성을 예측 할 수 있고, 음성을 수집하는 데 필요한 센서로는 값싼 마이크로폰을 이용할 수 있으며 음성과 비교할 때 너무 큰 잡음만 아니면 감성인식을 위한 특징 추출에 큰 영향을 미치지 않는다는 점이 감정 인식 연구에서 음성을 주요 매체로 사용하는 연구진이 많은 이유이다. 이러한 이유로 본 연구에서도 음성으로부터의 감정 인식을 시도한 것이다.

음성 신호를 이용한 감성인식에서 사용하는 특징점은 크게 피치, 에너지, 포먼트 주파수, 말의 빠르기로 구성된다. 연구자에 따라 이 4가지 특징 점을 모두 선택하기도 하고 일부만 선택하기도 하며 각 특징 점들에서도 다양한 통계치를 추출하여 사용하는 것이 일반적이다. 그래서 Dimitrios 와 Constantine은 5개의 감정을 구분하는데 피치, 에너지, 포먼트 주파수에서 87개의 특징점을 추출하여 분류하였고[1], Chul Min 과 Shrikanth는 남성과 여성을 구분하여 특징점을 추출하여 피치, 에너지, 말의 빠르기 등으로부터 17개의 특징점을 추출하였다[2]. 이외에도 Healey는 11개의 특징점, Picard 는 40개의 특징점, Haag et al. 은 13개의 특징점 등으로 각 연구자들 마다 다양한 특징점의 개수를 사용하고 있다[3].

이렇듯 다양한 특징점들을 사용한다는 것은 다른 의미로 보면 아직 감성인식에 사용할 대상 특징점이 명확하게 정립

접수일자 : 2006년 10월 9일

완료일자 : 2006년 11월 30일

이 논문은 서울시 산학연 협력사업(과제번호 : 106876)에 의해 수행되었습니다. 연구비지원에 감사드립니다.

되어 있지 않다는 것이고, 이 부분에 대해서는 획기적인 결과를 보여 줄 수 있는 특징점이 정립되기 전까지는 이 분야의 연구자들이 항상 고민을 해야 하는 부분일 것이다. 이러한 선택의 문제에서 고민을 다소 해소 시켜 줄 수 있는 방법으로 GA를 이용한 방법, Floating search 방법들이 있다[4]. 특히 floating search 방법에는 sequential forward selection 혹은 sequential backward selection method가 자주 사용되고 있는데, Chul Min 과 Shrinkanth는 forward selection(FS) method를 사용하였고[2], Dimitrios와 Constantine은 Sequential floating forward selection algorithm을 사용하여 87개의 feature set으로부터 10개의 best features를 선택하여 실험하였다. Yi-Lin 과 Gang 또한 Sequential forward selection을 사용하여 39개의 후보 Feature set에서 최적의 Feature subset을 선택하였고[5] Fabian et al.은 음악의 장르를 구분하는 문제에서 Genetic Programming을 이용하여 좋은 Feature set을 찾아내었다[6].

이러한 Feature selection 방법들은 '차원의 저주(Curse of dimension)'에 대한 좋은 해결책을 제시해 주었고, 패턴 인식 성능 향상에도 기여를 하였다. 또한, feature selection 방법들은 교사 값을 갖는 경우와 교사 값을 갖지 않는 경우로 나눌 수 있는데 보통의 경우 교사 값을 갖는 경우 들이 많다. 비교사 특징 선택 방법은 인지된 소리 외의 다른 것에 대응하는 결과를 보일 확률이 높기 때문에 교사 값을 갖는 방법을 많이 사용하는 것이다[7]. 그러나 실질적으로 Feature selection method를 사용해야하는 경우는 분명한 교사 값을 갖기 어려운 경우이기 때문에 비교사 값을 사용한 방법이 유용할 것이다. 이러한 두 방법의 문제점을 해결할 수 있는 방법으로 교사와 비교사 학습의 중간지점에 위치한 강화학습을 사용하는 방법을 제안한다.

강화 학습은 sutton 과 Barto 가 제안한 많은 방법들을 통해 다양한 연구들이 진행되어 왔다. Dynamic programming, Monte Carlo method, TD method, Q learning 등의 방법들이 그것인데, 이 방법들은 각각 다른 방법이지만 연속성을 가진 동일한 방법이기도 하다. 이렇듯 다양한 강화학습 방법들이 존재하고 있고, 강화학습의 주요 요소인 State, Action, Reward 의 개념을 High level부터 Low level 까지 개발자가 다양하게 정의하고 구현할 수 있기 때문에 Machine Learning 기법 중 매우 중요한 방법으로 생각되어 진다[8].

본 논문에서는 특정 강화학습 방법을 이용하기 보다는 강화학습의 기본 개념인 어떤 State에서 action을 취했을 때 받게 되는 Reward의 값을 이용하여 Feature set을 선택하는 방법을 제안한다. 특히, 단순히 reward 만을 합산하는 것이 아니라 천이되는 감정상태의 발생 빈도 또한 Feature selection을 위한 평가단계에서 가중치로 사용하여 사용자와의 접촉 빈도가 많아질수록 정밀하게 학습해가는 장점을 가진다.

본 논문의 구성은 다음과 같다. 2장에서 감정 인식 방법에 대한 설명을 하고, 3장에서 제안한 알고리즘에 대한 구조와 예시를 통한 설명을 한다. 4장에서는 제안한 알고리즘을 이용한 시뮬레이션 과 결과를 보이고 마지막 5장에서 결론 및 향후 과제에 대해서 언급한다.

## 2. 감정 인식 방법

본 논문은 감정 인식의 다양한 매체들 중 음성으로부터 Feature를 추출하여 패턴 인식을 한다. 음성에 의한 감정 인

식의 방법에는 크게 Acoustic information을 이용하는 경우와 Language and Discourse Information을 이용하는 경우로 나눌 수 있다. 전자의 경우는 단어 뜻 그대로 음향적 측면에서 피치, 포먼트 주파수, 타이밍, 음질 등을 Feature set으로 이용하는 방법이고, 후자의 경우는 단어의 의미 정보를 이용한다. 즉, 부정적인 단어인지 긍정적인 단어인지에 대한 구분으로부터, 즐거움을 나타내는 단어인지 슬픔을 나타내는 단어인지에 대한 구분 정보를 이용한다.

감정 인식의 방법은 분석과 학습을 위한 감정적 음성 데이터를 취득하여 데이터베이스를 구축하고, Feature set을 추출한다. 이렇게 추출된 Feature를 Pattern Classification 방법을 이용하여 감정별로 학습 및 분류를 하는 과정으로 이루어져 있다.

### 2.1 Database and Preparation

감정적인 음성은 10명의 남성 대학원생들로부터 획득되었다. 피험자들의 연령대는 24~31세이고 4가지 감정(평서, 화, 기쁨, 슬픔)으로 10가지 대사를 연기하도록 요구하였다. 특히 10가지 문장들(대사)은 초기에 준비된 30가지 문장을 모두 연기하도록 한 뒤 피험자 이외의 사람들에게 녹음된 소리를 들려준 뒤 "녹음된 소리가 4가지 감정 중 어떤 감정으로 느껴집니까?" 란 질문에 대해 90%이상의 동의로 주어진 감정과 일치된다고 합의된 것들이다. 준비되었던 문장들은 모두 6~10 음절로 제한된 것들이다. 녹음된 포맷은 11kHz, 16bit, mono 이고, 녹음 시 피험자들은 마이크로폰과의 거리를 10cm로 일정하게 유지하도록 하였다. 마이크로폰과의 거리는 녹음된 음성의 Loudness 혹은 intensity 에 영향을 미치므로 거리 유지가 매우 중요하다. 이렇게 녹음 된 파일들에 대해서 Preprocessing 과정을 거쳐 MS-ACCESS와 연동된 자체 개발한 DB에 저장한다. Preprocessing 과정에서는 FFT를 이용해 스펙트럼 추출, Autocorrelation method를 이용한 Pitch 추출, Pitch의 contour를 나타내는 IR(Increasing Rate), CR(Crossing Rate), VR(Variance) 와 평균, 최대, 최소 등의 통계치 들을 추출한다. 그림 1은 감정 인식기로서 이 프로그램에서 음성의 녹음 및 위와 같은 Preprocessing과 Pattern Classification 기능을 갖고 있고 그림 2와 같은 DB를 포함 하고 있다[9].

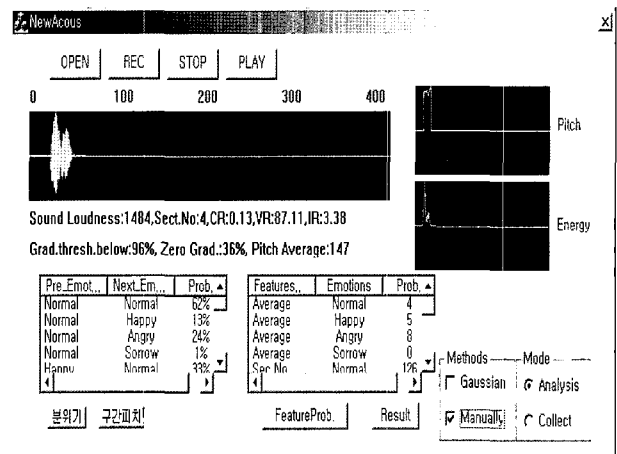


그림 1. 감정 인식기  
Fig 1. Emotion Recognizer

Features DB					
FileName	Pitch	Loudness	IR	VR	CR
E:ME_AM중하...	191	2987	-3.733333	6307.276955	2.666667
E:ME_AM나아...	204	3276	3.583750	3416.602783	0.937500
E:ME_AM나지...	210	3030	-2.525316	3497.506348	0.701754
E:ME_AM모여...	158	2541	-1.975000	3240.971924	1.666667
E:ME_AM이제...	175	2646	-4.915256	4307.966816	0.338983
E:ME_AM이제...	194	2707	0.960000	68.713600	1.200000
E:ME_AM그만...	163	2642	-2.034483	545.543395	0.344828
E:ME_AM이제...	189	2958	1.040000	1457.350464	0.690000
E:ME_AM나아...	231	2991	-2.125000	1212.555542	1.250000
E:ME_AM나아...	234	3167	-0.409836	657.672119	1.311475
E:ME_AM이제...	174	2713	-3.601919	1629.225708	0.454545
E:ME_AM그간...	183	2935	-4.565217	2120.249512	1.304340
E:ME_AM모여...	251	2967	0.000000	7102.243164	1.666667
E:ME_AM나아...	203	2947	0.000001	26497.273438	3.636364
E:ME_AM나아...	257	3000	1.682927	10931.970117	2.439025
E:ME_AM나아...	219	3251	2.277778	1930.432251	1.296296
E:ME_AM그간...	221	3148	-0.946429	2194.571533	1.250000

그림 2. 음성 특징 데이터베이스  
Fig 2. Speech Feature Database

2.2 Pattern Classification Method

패턴 인식 방법은 일반적으로 좋은 인식율을 보여주고 잡음이 낀 신호에도 강인한 성능을 보여주는 인공 신경망을 사용하였다. 1990년 이후로 감성인식 분야에서 가장 많이 사용된 패턴 인식 방법이기도 한 신경망 알고리즘은 실수, 이산값, 벡터 값 등의 입력에 대해서도 사용이 용이한 매우 실용적인 방법이고 보통 Back Propagation Algorithm을 사용하여 네트워크 파라미터를 조정한다. 본 연구에서는 신경망의 파라미터를 다음의 표와 같이 설정하였다.

표 1. 신경망 파라미터 설정  
Table 1. Parameter setting for ANN

Parameter	Value
Input Units	3~5 (상황에 따라 조정)
Hidden Units	11
Output Units	2
Learning Rate	0.003
Tolerance	0.25
Sigmoid function	$\frac{1}{1 + e^{-3x}}$

3. Interactive Feature Selection (IFS) 알고리즘

보통 감정 인식에서는 다양한 Feature set을 사용한다. 과거에는 연구자가 감성과 음향의 분석 결과의 연관성을 발견하여 특정 특징들을 Feature set으로 사용하였으나 그러한 방법은 연구자의 주관성에 따라 더 이상의 성능향상은 기대하기 어렵다. 그러한 이유로 최근의 연구 경향은 30에서 90 개에 이르는 Feature들로부터 Feature selection 방법을 이용하여 좋은 Feature set을 찾아 Pattern classification을 한다. 많은 feature 들을 모두 사용하지 않는 이유는 그것들 모두가 패턴 분류에 적절한 지 확신할 수 없고 그것들이 각각 갖고 있는 잡음이 classifier의 성능을 저하시키기 때문이다. 그래서 Feature Selection(FS)이 classification 분야에서는

효율성을 증가시키기 위해서 많이 사용되고 있는 것이다[7].

3.1 IFS 알고리즘

IFS는 Feature의 상관성과 강화학습을 이용한 알고리즘이다. 강화학습은 에이전트와 환경이 존재하는 구조에서 에이전트를 사용자가 원하는 목적을 이루도록 학습하는 방법이다. 학습하는 방법은 주어진 환경에서 에이전트가 동작을 취하고 취한 동작에 대해 환경이 보상을 취하는 형태로 이루어진다. 이때 각 시간 step은  $t$ , 각 단계에서 에이전트가 받게 되는 환경의 상태는  $s_t \in S$ ,  $S$ 는 가능한 환경 상태의 집합, 으로 표현되고 동작은  $a_t \in A(s_t)$ ,  $A(s_t)$ 는 어떤 상태에서의 '동작들의 집합' 으로 표현 된다. 동작에 대한 보상을  $r_t$  라 하고 이  $r_t$ 는 하나의 에피소드가 끝나면 다음과 같은 식으로 표현된다.

$$R_t = \sum_{k=0}^T \gamma^k r_{t+k+1} \tag{1}$$

위의 식에서  $\gamma$ 는 감쇠계수로써 Continuing task의 경우  $t = \infty$  까지 정의가 되므로 보상 값의 합이 무한대가 되지 않도록 하기 위함이다. 또한 감쇠계수를 0으로 하면 현재 발생한 보상 값만을 인정한다는 의미이므로 감쇠계수에 따라 미래의 값에 대한 가중치를 다르게 줄 수 있다. 결론적으로 강화학습은 식 (3)을 최대화 하는 방향으로 정책을 결정하는 방법이다. 그림 3은 제안한 알고리즘의 흐름도를 나타낸다.

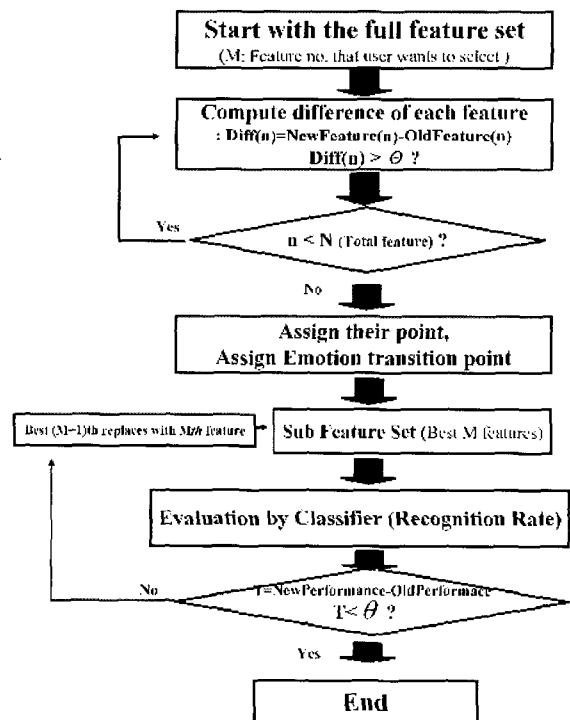


그림 3. IFS 알고리즘 흐름도  
Fig.3. IFS algorithm flowchart

그림에서 보인 것과 같이 처음에는 모든 후보 Feature 집합으로부터 시작한다. 그러나 이 방법은 Sequential Backward Selection(SBS)에서 순차적으로 Feature를 빼가는 것과는 달리 단지 동일 클래스에 대해 높은 상관성을 갖

는지, 혹은 서로 다른 클래스에 대해서는 어느 정도의 낮은 상관성을 갖는지를 측정한다. 동일 클래스인 경우에는 '+' 붙여줌으로써 '보상(reward)'으로써 상관 정도가 평가에 적용 되도록 하고 다른 클래스의 경우에는 '-'를 붙여줌으로써 '벌(penalty)'으로써 적용되도록 한다. 이렇게 1차적으로 전체 후보 집합 중 일부를 선정하고 목적함수에 의해 평가를 받는다. 이전의 적합도와 현재의 적합도의 차이를 비교하여 기준치 이하이면 선정된 후보 집합을 차선의 후보 집합으로 대체하여 평가를 받고 이러한 과정을 종료조건이 만족될 때 까지 반복한다. 다음의 그림 5는 이와 같은 과정에 대한 예를 보여준다[Park & Sim 2006].

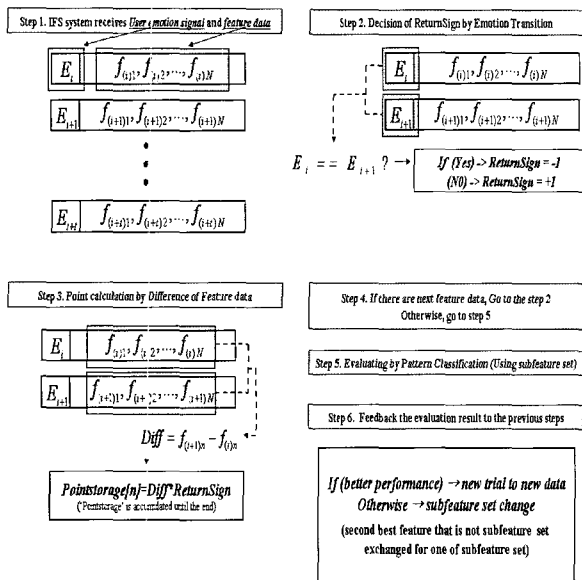


그림 4. IFS 알고리즘 설명도  
Fig.4. Diagram of IFS algorithm

그림 5에서  $E_i$ 는 사용자의 감정 식별자를 나타내고  $i$ 번째 입력으로 들어온 감정을 의미한다.  $f_{ij}$ 는  $i$ 번째 입력에 대한  $j$ 번째 feature를 의미한다. step 1에서는  $i$ 번째 감정과 feature set을 입력 받는다. step 2에서는 다음에 들어온  $E_{i+1}$ 과  $E_i$ 가 동일한지 검사하여 동일하다면 ReturnSign을 -1로 다르다면 +1로 부여한다. step 3에서는  $i$ 번째 feature set과  $i+1$ 번째 feature set의 대응되는 것들끼리 차이를 계산한다. 각각의 차에 step 2에서 구한 ReturnSign을 부호로 붙여 각각 저장한다. step 4에서는 또 들어온 입력에 대해 step 2, 3의 과정을 반복하여 Pointstorage[n]에 각 feature 요소들에 대한 점수를 누적한다. 더 이상 입력이 없게 되면 feature들 중 점수가 높은 것을 선택하여 목적함수에 의한 평가를 받는다. 두 번 이상의 평가이후 성능이 저하되는 경우에는 처음에는 일정 순위에 들지 못했던 것들 중 점수가 높은 feature를 입력하여 다시 평가한다. 이 또한 성능이 향상할 때까지 반복한다.

그림 5는 위와 같은 알고리즘을 이용하여 프로그램 할 때 사용되는 데이터의 구조이다. (1)은 감정과 음성 신호가 입력되었을 때 Preprocessing 과정을 거치고 난 뒤의 데이터 포맷이다. Current State는 현재의 감정이고 Next state는 다음에 받게 되는 감정을 의미하고 이는 정수형으로 기록된다.

$F_1$ 부터  $F_N$ 은 Feature들을 의미하는 것으로써 (2)와 같이 Float 형이고 배열로 되어 있다. 'New'에는 Next state에서 기록된 Feature 값이 저장되고 'Old'에는 Current state에서 기록된 Feature 값이 저장되며 'Point'에는 'New'와 'Old'의 차가 기록된다. 이 값은 차후에 'Weight'에 의해서 갱신된다. (3)은 자주 나타나는 감정에 대해서는 가중치를 주기 위해 존재하는 'Weight'이며 S는 존재할 수 있는 CS(Current State), NS(Next State)의 개수를 의미한다.  $S = \text{Emotion}^{\text{Pair}}$  로 S의 개수를 찾을 수 있다. 임의로 결정된 회수가 지나면 (3)에서 누적된 Point를 이용하여 (2)의 Feature Point에 가중치를 주어 N개의 Feature들로부터  $n$ 개의 Best Feature들을 선정한다.

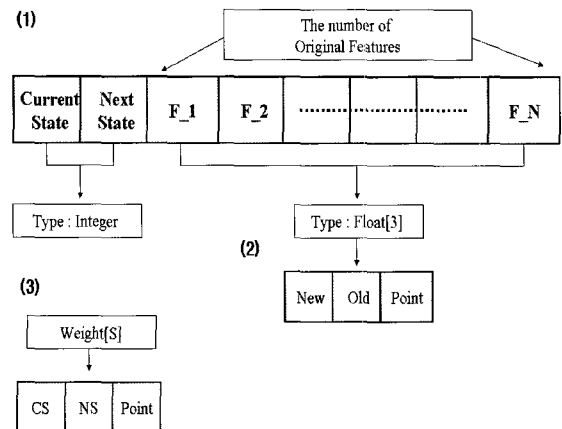


그림 5. 데이터 구조  
Fig 4. Data Format

## 2.2 제안한 알고리즘의 장점

본 연구에서 제안하는 Interactive Feature Selection Algorithm(IFS)은 강화학습에서 모티브를 가져온 알고리즘이다. 이 알고리즘은 감정 인식 시스템이 일반 가정의 로봇이나 가전제품에서 부착되어 사용되어 진다고 가정한다. 그러한 경우라면 사용자가 해당 기기를 사용할 때 가끔 감정적 발화와 그 당시의 감정상태(교사신호로써)를 이 시스템이 입력하는 일이 많이 불편하진 않을 것이기 때문이다. 이러한 점을 감안한다면 본 알고리즘은 사용자별로 감성의 특징점이 다를 수 있다는 점을 훌륭히 해결할 수 있는 사용자 적응형 시스템이며 사용자에게 대해 학습을 하면 할수록 좋은 성능을 보일 수 있는 장점을 갖고 있다.

본 알고리즘은 다음과 같은 특징을 갖고 있다.

- ① 사용자의 감정적 발화 녹음 시 사용자 자신의 감정 상태를 정확히 입력 (교사 신호 입력)
- ② Original Feature의 개수 무제한 입력 가능
- ③ 각 Feature의 포인트는 새로운 감정으로 천이될 때 기준 값 이상의 차이가 나는 대상에만 부여
- ④ 감정 천이 패턴들에 대해 각각 누적 횟수를 기록하여 감정 입력횟수가 일정 수 이상이 되었을 때 Feature의 포인트에 가중치로 사용한다. 이렇게 함으로써 각 Feature의 신뢰도가 향상된다.

위와 같은 특징을 갖고 있기 때문에 SFS 나 GA를 이용한 여타의 selection method 보다 각 사용자에게 더욱 적합한 Feature selection이 가능해진다.

### 4. 시뮬레이션 및 결과

본 절에서는 제안한 Interactive Feature Selection(IFS) 알고리즘을 다음의 함수에 적용하여 성능을 평가한다.

$$3x_0 - 2x_1 + 4x_2 + 10x_3 - 5x_4 + 8x_5 + 7x_6 + 8x_7 - 10x_8 + 6.8x_9 \quad (2)$$

다음의 그림 6은 IFS 알고리즘의 적합도 추이 그래프이다. 그림에서 위의 선은 적합도, 아래 선은 해당 세대에서 Feature 개수를 각각 나타낸다. IFS는 알고리즘의 특성상 사용자가 목표로 하는 Feature의 개수를 미리 정할 수 있기 때문에 그림에서 나타난 세대별 Feature의 개수는 사용자가 미리 결정한 것이고, 위의 선은 그때의 적합도를 나타낸 것이다. 1개에서 2개로 늘렸을 때는 성능이 선형적으로 증가했으나 3개로 늘렸을 때는 3세대 동안 다른 Feature set을 선정하는 작업을 거쳐서 좋은 적합도를 보이는 결과를 찾을 수 있었고 4개로 증가시켰을 때도 3세대 동안 다른 Feature set을 선정하는 작업을 거쳐서 47에 가까운 값으로 증가하였다.

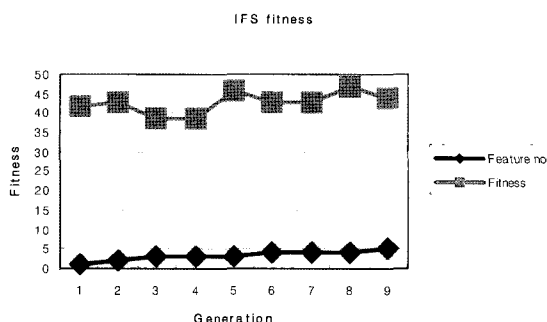


그림 6. IFS 알고리즘의 성능 추이 그래프  
Fig. 6. Performance transition graph of IFS algorithm

한편 제안한 알고리즘의 유효성을 검증하기 위해서 순차 전향 선택(SFS: Sequential Forward Selection) 방법, 유전자 알고리즘 기반 특징 선택(GAFS: Genetic Algorithm Feature Selection) 방법[10], RS(Random Selection) 방법과 비교 평가하였다. 다음의 그림 7은 이들의 방법을 이용해서 Feature 개수(표 1 참조)를 변화시켰을 때 감정 분류 비율을 그래프로 나타낸 것이다. 예상한 바와 같이 RS 방법은 다른 방법들에 비해서 그 성능이 떨어짐을 알 수 있다. 반면, IFS, SFS, GAFS 등은 미묘한 차이를 나타내지만, 본 논문에서 제안한 IFS 방법이 적은 수의 Feature의 개수로서도 우수한 감정 인식 결과를 얻을 수 있음을 알 수 있다.

표 1. 실험에서 사용한 Features  
Table 1. Features used in experiment.

Feature No.	Features
1	Feature mean
2	Feature mean, Speech rate
3	Feature mean, Speech rate, Loudness
4	Feature mean, Speech rate, Loudness, Sect. No.
5	Feature mean, Speech rate, Loudness, Sect. No. Intensity mean

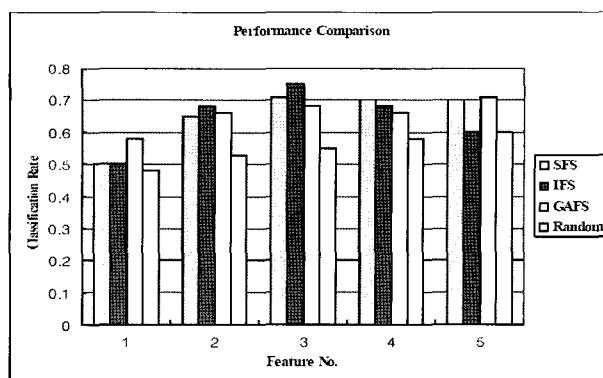


그림 7. IFS, SFS, GAFS, RS의 감정 분류 비교 그래프  
Fig. 7. Emotion Classification Rate Comparison of IFS, SFS, GAFS, and RS.

### 5. 결론

본 논문은 특징점이 많은 패턴 인식의 경우 차원의 저주 문제의 해결책으로 제시될 수 있고 인식 성능의 향상에 도움을 줄 수 있는 Feature selection algorithm으로써 강화학습의 개념을 기반으로 제안한 Interactive Feature Selection Method에 관한 논문이다. 본 연구에서 구현한 IFS를 통해 Best Features를 찾아내어 일반적으로 좋은 성능을 보여주고 있는 ANN으로 감정 인식을 한 결과 임의로 선택한 Feature 보다 좋은 결과를 보임을 확인하였다. 차후에는 더욱 다양한 경우에 대한 결과를 보여 알고리즘의 우수성을 확인하도록 할 것이다.

### 참고 문헌

- [1] D. Ververidis and C. Kotropoulos, "Emotional speech classification using Gaussian mixture models," *Proceedings of ISCAS*, vol. 3, pp. 2871-2874, May, 2005.
- [2] C. M. Lee and S. S. Narayanan, "Toward detecting emotions in spoken dialogs," *IEEE Transactions on Speech and Audio Processing*, vol. 13, pp. 293-303, March, 2005.
- [3] J. Wagner, J. H. Kim and E. Andre, "From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification," *Proceedings of ICME*, pp. 940-943, July, 2005.
- [4] P. Pudil and J. Novovicova, "Novel Methods for Subset Selection with Respect to Problem knowledge," *IEEE Intelligent Systems*, pp. 66-74, March, 1998.
- [5] Y. L. Lin and W. Gang, "Speech Emotion Recognition based on HMM and SVM," *Proceedings of Machine Learning and Cybernetics*, vol. 8, pp. 4898-4901, Aug., 2005.
- [6] F. Morchen, A. Ultsch, M. Thies and I. Lohken,

"Modeling Timbre Distance With Temporal Statistics From Polyphonic Music," *IEEE transaction on Audio, Speech and Language Processing*, vol. 14, issue 1, pp. 81-90, Jan., 2006.

[7] E. F. Combarro, E. Montanes, I. Diaz, J. Ranilla, and R.Mones, "Introducing a Family of Linear Measures for Feature Selection in Text Categorization," *IEEE transactions on Knowledge and Data Engineering*, vol. 17, no. 9, pp. 1223-1232, Sept., 2005.

[8] R. S. Sutton and A. G. Barto, *Reinforcement Learning : An Introduction*, A Bradford book, London, 1998.

[9] C. H. Park and K. B. Sim "The Implementation of the Emotion Recognition from Speech and Facial Expression System," *Lecture Notes in Computer Science(LNCS)*, vol. 3611, pp. 85-88, 2005.

[10] 박창현, 김호덕, 영현창, 심귀보, "패턴 인식문제를 위한 유전자 알고리즘 기반 특징 선택 방법 개발", *한국퍼지 및 지능시스템학회 논문지*, 제16권, 제4호, pp. 466-471, 2006.



김호덕(Ho-Duck Kim)  
 2005년 : 중앙대학교 전자전기공학부 졸업  
 2006년 : 동대학원 전기전자공학부 석사과정 재학 중

관심분야 : 지능시스템, 임베디드 시스템, 유비쿼터스 로봇 등



박창현(Chang-Hyun Park)  
 2001년 : 중앙대학교 전자전기공학부 졸업  
 2003년 : 동 대학원 전자전기공학부 공학석사.  
 2006년 : 동 대학원 전자전기공학부 공학박사

2006년 ~ 현재 : 한국전자통신연구원(ETRI) 선임연구원

관심분야 : 감정인식, 패턴인식, 기계학습, 진화 연산 등

저 자 소 개



양현창(Hyun-Chang Yang)  
 2006년 : 중앙대학교 대학원 전자전기공학부 박사과정 재학 중

관심분야 : 지능로봇, 감정인식, 로봇정책, 스마트 홈, 유비쿼터스 센서네트워크 등



심귀보(Kwee-Bo Sim)  
 1990년 : The University of Tokyo 전자공학과 공학박사

[제16권 4호(2006년 8월호) 참조]

1991년 ~ 현재 : 중앙대학교 전자전기공학부 교수  
 2006년 ~ 현재 : 한국퍼지 및 지능시스템학회 회장

E-mail : kbsim@cau.ac.kr  
 Homepage URL : <http://alife.cau.ac.kr>