

퍼지 로직을 이용한 감정인식 모델설계

김 이 곤, 배 영 철
여수대학교 전기공학과

Design of Emotion Recognition Model Using Fuzzy Logic

Yigon Kim, Yong-Chel Bae
Dept. of Electrical Engineering, Yosu National University
Yosu 554-749, Korea

<Abstract>

Speech is one of the most efficient communication media and it includes several kinds of factors about speaker, context emotion and so on. Human emotion is expressed in the speech, the gesture, the physiological phenomena(the breath, the beating of the pulse, etc). In this paper, the method to have cognizance of emotion from anyone's voice signals is presented and simulated by using neuro-fuzzy model.

I. 서론

최근 인공지능은 보다 인간중심의 문제해결 방법의 추구로 인하여 사회 모든 분야에서 관심이 집중되고 있다. 공학적인 분야는 물론 인문사회 분야에서 그 응용분야를 넓히고 있다. 더 나아가 지금까지의 인간 지능에 대한 계산모델뿐만 아니라, 인간의 비 지적인 감정, 무의식에 대한 계산모델에 대한 많은 요구가 있다. 인간의 감정에 관한 연구는 여러 분야에 이용될 수 있는데, 대표적인 것은 능동적인 인터페이스다.

이 능동적인 인터페이스는 사용자에게 메뉴를 제공하고 사용자의 명령에 반응하는 형식이 아닌, 사용자의 행동, 감정, 심리 상태를 파악하고 그에 따라서 대응하는 형식을 갖는다. 따라서 이러한 능동적인 양방향성 인터페이스 기술은 기본적으로 사용자의 감정상태 등을 파악, 예측에 대한 것은 필수적이며, 이를 체계적으로 접근하기 위해서는 사용자의 감정상태와 그 변화에 대한 계산모델이 필요하다. 따라서, 좀 더 인간 중심적인 제품을 위하여 사용하는 인간의 심리상태, 감정상태까지 파악하고자하는 요구가 더욱 증가하고 있다.

감정이란, 외부의 물리적 자극에 의한 감각, 지각으로부터 인간의 내부에 야기되는 심리적 체험으로, 쾌적함, 고급감, 불쾌감, 불편함 등의 복합적인 감정을 말한다. 따라서 이와 같은 인간의 감정을 정성, 정량적으로 측정 평가하고 이를 제품이나 환경 설계에 응용하여 보다 편리하고 안락하며, 안전하게 하고 더 나아가 인간의 삶을 쾌적하게 하고자 하는 기술이 감성공학이다. 국내, 외의 감성공학 기술개발 동향은, 국내는 1992년 G7 후보과제의 하나로 선정되어 된 후 삼성, LG, 대우 등의 가전사에서 제품의 감성설계를 위한 감성공학 연구조직을 설치하여 연구를 수행중이고, 국외의 동향으로, 일본은 1990년부터 98년까지 200억엔 규모의 통산성 대형국책프로젝트의 하나로 '인간감각계측응용기술개발 프로젝트'를 수행중이다. 그리고 영국 러트루우 대학의 CES, 네덜란드 아이트호벤의 인간감각연구소(IPO)등에서는 인간의 감각연구결과를 제품에 응용하기 위한 연구가 활발하다.

II. 음성신호로부터 감정인식

감정 인식은 여러 방법으로 접근할 수 있으나 본 연구에서는 특정인의 음성인식을 통해 현재의 감정상태를 인식하는 것으로, 일상 인간의 대화속에서 느끼는 감정의 변화를 감지하는 음성신호를 이용한 감정인식 모델을 제시하고자 한다.

감정 모델링에서 가장 어려운 부분중의 하나는 사람으로부터 실제로 감정을 나타내는 정보를 추출하는 것이다. 이에 관련하여, Yanaru는 사람의 감정 모델링에 대한 수학적 모델을 제시한 바 있다. Yanaru의 논문은 P.Plutchik과 P.T.Young의 심리학 연구 결과를 토대로 하고 있는데 Yanaru가 가정한 R.Plutchick과 P.T.Young의 이론은 인간은 적은 수의 기본 감정을 갖고 있으며, 다른 감정들은 이 감정의 복합적인 조합으로 이루어지고, 이 기본감정들은 복잡하고 다양한 형태로 표현될 수 있는데, 어떤 특정의 기본감정에 대하여 강도의 크기는 같으나 방향이 다른 기본감정이 존재한다. 인간의 감정변화는 감정요소의 형태, 강도의 크기, 감정요소가 제시된 시간에 따라서 영향을 받는다.

Yanaru는 인간의 감정상태와 감정요소를 Plutchik의 가설에 따라서 8개의 기본감정으로 나누어 표현하고, 각 기본감정은 모두 자신의 감정추론기관이 있어서, 8개의 기본 감정으로 나누어진 현재의 감정상태와 감정요소 중 자신의 강도를 입력으로 하여 감정변화를 추론하게 하였다.

본 연구에서는 인간의 감정을 Yanaru가 제안한 8개의 기본 감정을 기준으로 표현하였다. 그러나 상황에 따라 인간의 감정은 이들 복합된 형태로 나타나는 것이 일반적이지만 연구의 목적을 간략하게 하기 위해 독립된 형태로 표현하였다. 감정을 나타내는 특징으로는 우리나라 국악의 창에서 인간의 희로애락을 표현하는 음의 고저와 장단을 기본으로 하여 분석하였다.

III. 실험장치구성 및 인식모델설계

실험장치 구성

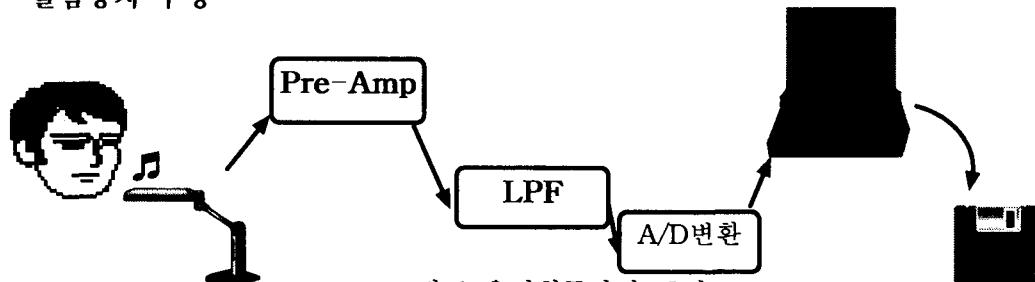


그림 1 음성취득장치 구성도

데이터 취득은 그림1과 같이 마이크로부터 전기적 신호로변환된신호를 전단증폭기를 통해 증폭하여 340KHz 이하의 음성신호만을 필터링하여 A/D변환기입력에 인가한다. A/D변환된 신호는 PC를 통해 취득되어 분석알고리즘을 통해 분석한다. 본 연구에서 얻어진 음성신호는 특정인의 감정에 대한 음성특징을 얻기 위하여 특정구문"아! 그렇습니까?"을 8개의 감정을 제시하여 말하게 하였으며, 각각 3회 반복하게 하여 데이터를 취득하였다. Yanaru가 사용하였던 기본감정 joy, sadness, expectancy, surprise, anger, fear, hate, acceptance으로 분류하여 특정 사람에 대한 실험을 행하였다.

Neuro-Fuzzy 인식모델설계

음성 신호로부터 특징 정보를 추출하여 입력 변수로 하여 상태 진단 모델을 설계한다. 입력 변수가 x_n 로 ($n=2$ 인 경우). 출력 변수가 y^i 인 퍼지 규칙은 다음과 같이 표현된다.

$$R^i: \text{if } x_1 \text{ is } A_1^i \text{ and } x_2 \text{ is } A_2^i, \text{ then } y^i = b^i \quad (3-1)$$

($i= 1, 2, 3, \dots, c$)

여기서, R^i 는 i 번째의 규칙, c 는 규칙 수, A_1, A_2 는 삼각형의 퍼지 집합이고, y^i 는 i 번째 룰의 후건부 출력이며, b^i 는 i 번째 룰의 싱글톤 출력값이다. 퍼지 모델의 추론법으로서 맥스-프로덕트(Max-product) 합성법을 이용하며, 무게 중심법을 이용하여 비 퍼지화한 값을 모델의 출력으로 한다. 무게중심법을 이용한 비 퍼지화값은 모델식(3-1)의 입력이 (x_1, x_2) 인 경우에 식(3-2)과 식(3-3)으로 표현된다.

$$w^i = \mu_{A_1}^i(x_1) \times \mu_{A_2}^i(x_2) \quad (3-2)$$

$$y = \frac{\sum_{i=0}^m w^i \times y^i}{\sum_{i=0}^m w^i} \quad (3-3)$$

식(3-2)의 $\mu_{A_1}^i(x_1)$ 는 퍼지 집합 A_1 에 대한 멤버쉽함수, $\mu_{A_2}^i(x_2)$ 는 퍼지 집합 A_2 에 대한 멤버쉽함수, y^i 는 i 번째 룰의 출력, w^i 를 입력 (x_1, x_2) 의 규칙 R^i 에 대한 적합도라 부른다. 입력 퍼지 집합의 멤버쉽 함수는 종형 멤버쉽 함수를 이용하였으며, 각 변수의 멤버쉽 함수를 변수 값으로 하는 1개의 규칙이 구성된다. 모델링 수행 순서는 Fig 3-1과 같이 수행한다.

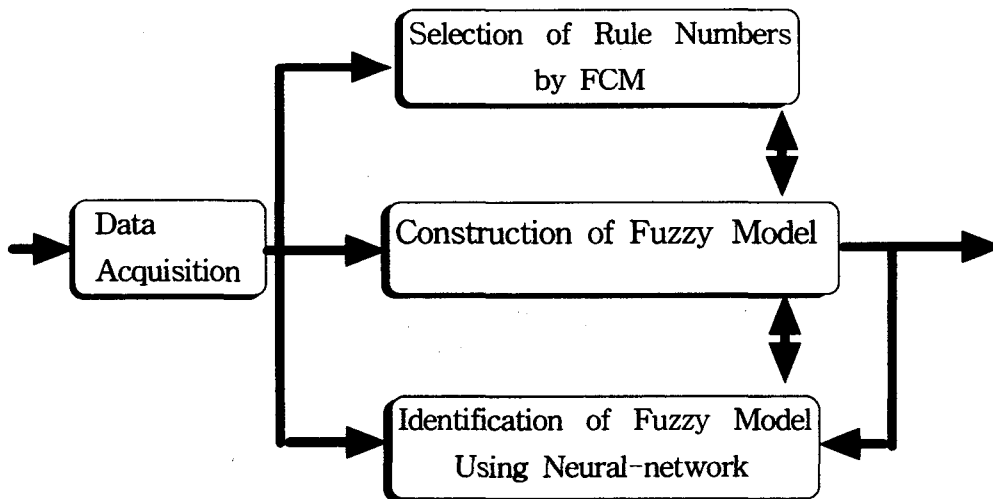


Fig.3-1 Diagram of Modeling Process.

Fig 3-1과 같이 구성된 모델링 수순에 의해 Fig 3-2와 같이 c 개의 규칙을 갖는 모델로 표현할 수 있다. Fig 3-2에서 (A)층은 입력층(Input layer), (E)층은 출력층(Output layer)이고 나머지 층들은 중간층(Hidden layer)이며, 각 층들을 서로 연결하고 있는 연결 가중치는 퍼지 규칙의 파라미터를 나타내는 값을 갖는다.

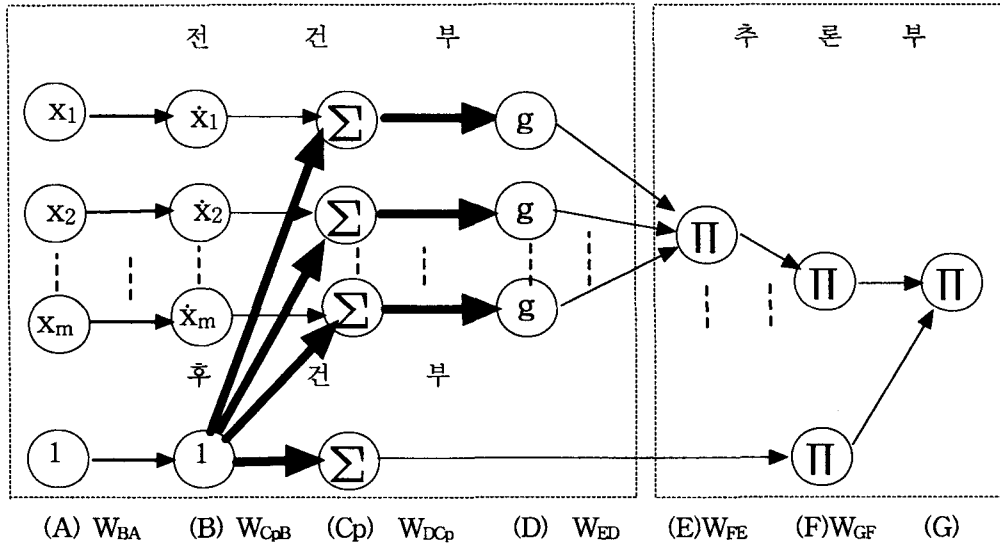


Fig.3-2 Structure of Neural Network for a fuzzy rule

(A)층과 (B)의 관계는 입력을 0과 1사이의 값으로 정규화하는 과정으로 이때 \dot{x}_i 는 식(3-4)과 같다. (B)층과 (Cp)의 사이의 하중치 w_{CpB} 는 1값을 갖고 (D)와 (E)사이의 하중치 w_{ED} 는 시그모이드값을 퍼지멤버십값으로 환산하기위해 4값을 갖도록하였다. 그리고 중간층들 중 (B),(C),(D),(E)층은 전건부를 나타내고 있어 이들 층과 연결되는 가중치 $w_{BA}, w_{CpB}, w_{DCp}, w_{ED}$ 는 전건부의 멤버쉽 함수를 구성하는 파라미터이고 Base층과 연결된 w_{CcB} 는 후건부 싱글톤 b 를 나타낸다. 여기에서 뉴런의 Σ 기호는 입력 치의 합을 나타내고 Π 기호는 입력들의 곱을 나타낸다. 식(3-6)의 $g(x)$ 는 시그모이드함수인 식(3-5)의 $f(x)$ 를 미분한 식으로서 종형 멤버쉽 함수를 만들고 이 식으로부터 입력 변수의 소속 정도를 나타내는 멤버쉽함수 값을 생성한다.

$$\dot{x}_i = \frac{1}{m} \sum_{j=1}^m x_{ij} \quad (i=1,2,3,\dots,m(\text{데이터 개수})) \quad (3-4)$$

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (3-5)$$

$$g(x) = f(x)' = a \left(\frac{1}{1 + e^{-ax}} \right) \left(1 - \frac{1}{1 + e^{-ax}} \right) \quad (3-6)$$

단, x : 뉴런의 입력
 a : sigmoid 함수의 기울기

(B)층과 (A)층을 연결하는 w_{BA} 는 (B)층의 값이 -1에서 1사이의 값이 되도록 하는 정규화를 위

한 값이고, D층의 입력은 식(3-6)에서 $w_{CbB}(x + w_{DCb})$ 이다. 여기서 w_{CbB} 와 w_{DCb} 값은 시그모이드 함수의 기울기와 위치를 나타내며 삼각형 멤버쉽함수를 결정하는 요인이 된다. 위 식에서 w_{DCb} 값은 멤버쉽함수의 기울기를 결정하고, w_{CbB} 값은 함수 모양의 중심을 결정한다.

$$f(x) = \frac{1}{1 + e^{-w_{CbB}(x + w_{DCb})}} \quad (3-7)$$

(D)층에서 출력되는 $g(x)$ 값은 멤버쉽함수로서 0에서 0.25까지 출력됨에 따라 연결 가중치를 4로 정하여 멤버쉽함수의 최대치가 1이 되도록 한다. 이렇게 각 입력층에 입력된 변수들은 중간층을 거치는 과정에서 $A_{ji}(x_j)$ 의 의미를 반영하고 (E)층에서 출력되는 값들을 각각 곱하여 i 번째 룰의 적합도 w^i 를 계산한다. 후건부 선형 방정식을 구성하는 회로는 Fig 3-2에서 1로 된 입력층 (input layer) (B)와 (Cc)층과 연결되도록 하면 연결 가중치 w_{CcB} 값이 후건부의 싱글톤값 b 를 나타낸다. 따라서 (Cc)층의 출력은 다음 식과 같이 i 번째 룰의 후건부는 $y^i = b^i$ 와 같다. (F)층은 지금까지 퍼지화된 전건부의 적합도 w^i 와 후건부의 출력 y^i 로부터 추론을 행하는 층으로서 적합도 w^i 와 출력 y^i 를 곱한 값을 출력한다. (F)층에서 추론된 값은 연결치 w_{GF} 를 통하여 (G)층에 입력되는데 이때 w_{GF} 는 모든 룰의 (E)층 뉴런에서 계산된 w^i 을 합한 값의 역수로서 (G)층에 입력되는 추론 값은 식(3-8)과 같다.

$$w_{GF} = \frac{1}{\sum_{K=1}^n w^K} \quad (3-8)$$

$$y^{i*} = \frac{w^i * y^i}{\sum_{K=1}^n w^K} \quad (3-9)$$

w_{ED} 번째 룰에서 계산된 식(3-8) 출력 y^{i*} 로부터 식(3-9)과 같이 합하면 비 퍼지화 과정을 행한 것과 같다.

$$y^* = \sum_{i=1}^n y^{i*} \quad (3-10)$$

FCM(Fuzzy C-Mean)으로 결정된 파라미터들을 전건부 멤버쉽함수와 후건부의 출력치로 설정한다. 그러나 이들의 값은 근사화된 값이므로 최적으로 조정되어야만 더 정확한 출력을 얻을 수 있다. 만일 연결 가중치가 부적합한 값으로 설정되어 있다면 그 에러의 변화 분에 대한 값을 각 층에 역전파시켜 조정 시킴으로써 정확한 연결 가중치를 동정하도록 한다. 전건부 파라미터 동정의 경우 각 규칙에 따른 멤버쉽 함수의 모양과 관련된 가중치 w_{CbB} 와 w_{DCb} 들을 변화시켜 적절한 모양이 되도록 동정하고, 후건부 동정으로는 싱글톤값을 역전파(back propagation) 학습을 이용하여 동정한다. 다층 신경회로망의 역전파 학습 알고리즘은 일반화된 델타 룰(Generalized delta rule)을 적용함으로써 이루어지는데 임의의 (n)번째 층을 표현한 모델에 따르면 Fig 3-3과 같다

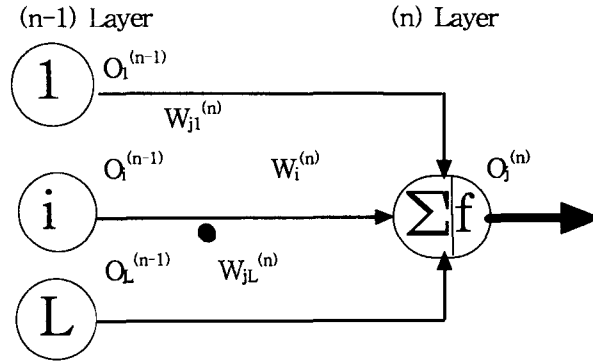


Fig.3-3 Neuron Model of (n)th Layer

($n-1$)층에 존재하는 L 개의 뉴런에서 출력되는 $O_i^{(n-1)}$ 는 각각의 연결 가중치와 곱해지고 합산되어 $I_j^{(n)}$ 라는 값으로 (n)층에 입력된 다음 (n)층의 $f(I_j^{(n)})$ 의 함수를 통하여 출력 $O_j^{(n)}$ 이 얻어진다.

$$I_j^{(n)} = \sum_{i=1}^L w_{ji}^{(n)} * O_i^{(n-1)} \quad (3-11)$$

$$O_j^{(n)} = f(I_j^{(n)}) \quad (3-12)$$

이때 ($n-1$)층에 입력된 P 번째 학습데이터로부터 (n)층에서 원하는 출력이 $D_{pj}^{(n)}$ 이고 실제로 추론된 출력이 $O_j^{(n)}$ 이라면 그 에러 값은 $\epsilon^{(n)} = D_{pj}^{(n)} - O_j^{(n)}$ 이고, 이때 에너지 평가 함수는 식(3-13)과 같이 정의된다.

$$E_{pj} = \frac{1}{2} (D_{pj}^{(n)} - O_j^{(n)})^2 \quad (3-13)$$

에너지 평가 함수 E_{pj} 를 최소로 하기 위해 P 번째 데이터의 오차에 대한 i 번째의 결합 하중치 $w_{ji}^{(n)}$ 의 변화량 $\Delta w_{ji}^{(n)}$ 는 다음과 같이 유도된다.

$$\frac{\partial E_{pj}}{\partial w_{ji}} = -(D_{pj}^{(n)} - O_j^{(n)}) \frac{\partial f(I_j^{(n)})}{\partial I_j^{(n)}} \frac{\partial I_j^{(n)}}{\partial w_{ji}} \quad (3-14)$$

(n)층 입력의 i 번째 입력 가중치에 대한 미분을 구하면 식(3-13)과 같고 이를 식(3-14)에 대입하면 식(3-15)과 같다.

$$\frac{\partial I_j^{(n)}}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} (\sum_{k=1}^L w_{jk}^{(n)} * D_k^{(n-1)}) = O_i^{(n-1)} \quad (3-15)$$

$$\frac{\partial E_{pj}}{\partial w_{ji}} = -(D_{pj}^{(n)} - O_j^{(n)})f(I_j^{(n)})O_i^{(n-1)} \quad (3-16)$$

다음 학습되기 전의 상태를 (t) 라면 $(t+1)$ 상태의 가중치를 갱신하는 변화량은 에러의 변화분에 역으로 작용한다.

$$w_{ji}^{(n)}(t+1) = w_{ji}^{(n)}(t) + \Delta w_{ji}^{(n)}(t) \quad (3-17)$$

$$\Delta w_{ji}^{(n)} = \eta(D_{pj}^{(n)} - O_j^{(n)})f(I_j^{(n)})O_i^{(n-1)} \quad (3-18)$$

여기서 η 는 학습률(learning-rate)이다. 간략화시키기 위해 식(3-19)라 놓으면 식(3-20)과 같이 된다.

$$\delta^{(n)} = (D_p^{(n)} - O^{(n)})f(I^{(n)}) \quad (3-19)$$

$$w_i^{(n)}(t+1) = w_i^{(n)}(t) + \eta\delta^{(n)}O_i^{(n-1)} \quad (3-20)$$

위 역전파 학습 알고리즘에 따라 뉴로-퍼지의 전진부 삼각형 멤버쉽 함수와 후진부 싱글톤 값이 다음의 학습 과정을 통해 동정된다. 단 첨자 k 는 k 번째 퍼지 규칙을 나타내며 첨자 p 는 p 번째 데이터를 표시한다.

$$\varepsilon_p = (D_p - O^{(G)}) \quad (3-21)$$

$$\delta_p^{(G)} = \varepsilon_p \quad (3-22)$$

$$\delta_p^{(F)k} = w_{GF}^k \delta_p^{(G)} \quad (3-23)$$

$$\delta_p^{(D)k} = \frac{1}{\alpha} \delta_p^{(E)k} \quad (3-24)$$

$$\delta_p^{(Cp)k} = g'(I^{(D)k}) \delta_p^{(D)k} w_{DCp}^k \quad (3-25)$$

$$\begin{aligned} &= \frac{\partial(O^{(D)k})}{\partial(I^{(D)k}) \delta_p^{(D)k} w_{DCp}^k} \\ &= \frac{\partial}{\partial(I^{(D)k})} \frac{\alpha e^{-I^{(D)k}}}{(1 + e^{-I^{(D)k}})} \delta_p^{(D)k} w_{DCp}^k \end{aligned} \quad (3-26)$$

$$= \alpha O_p^{(D)} \left(1 - \frac{2}{(1 + e^{-I^{(D)}})}\right) \delta_p^{(D)} w_{DCp}^k \quad (3-27)$$

(F)층의 곱셈 계산에 따라 (E)층과 (C_C)층 뉴런에 대한 델타 식은 다음과 같다.

$$\delta_p^{(E)k} = \delta_p^{(F)} \frac{\partial O^{(F)}}{\partial O^{(E)k}} \quad (3-28)$$

$$\delta_p^{(Cc)k} = \delta_p^{(F)} \frac{\partial O^{(F)}}{\partial O^{(Cc)k}} \quad (3-29)$$

위 식으로부터 $(t+1)$ 상태의 연결 가중치는 식(3-30)과 같이 수정된다.

$$w_i^k(t+1) = w_i^k(t) + \eta \delta_i^k O_i \quad (3-30)$$

이와같이 역전파 학습 알고리즘에 따라 전건부 멤버쉽 함수와 후건부 계수(싱글톤 값)가 학습 데이터에 맞도록 동정되며, Fig 3-4와 같이 이루어진다. 퍼지 모델링은 일반 모델링과 달라서 전건부와 후건부로 된 규칙 기준 모델(rule based model)의 동정 문제이다. 규칙의 구조 동정은 전건부와 후건부의 동정으로 나눌 수 있으며, 전건부와 후건부 동정은 구조 동정과 파라미터 동정으로 구분한다. 후건부 동정에서는 후건부 싱글톤을 동정하고 전건부 동정에서는 멤버쉽함수의 기울기와 중심을 결정하는 w_{CpB} 와 w_{DCp} 를 동정한다. 본 연구에서 퍼지 모델의 규칙 수는 클러스터의 변동을 이용한 평가 지표가 국소 최소가 되는 후건부의 최적 분할 수로 선택하였다.

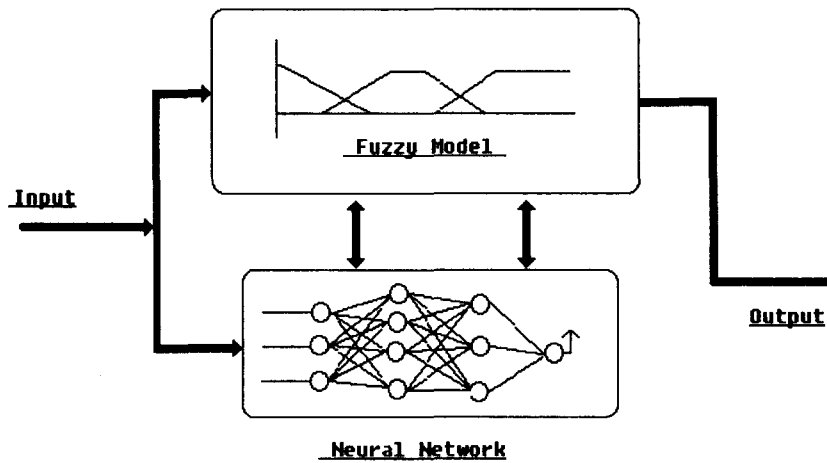


Fig 3-4 Block Diagram of Recognition Model

학습데이터는 각 8개의 기본감정에 대해서 각각 작성되었다. 본 연구에서 목표로 하는 퍼지-신경망은 특정인의 음성에서 기본감정에 대하여 학습하도록 하였다.

IV. 실험 결과 및 고찰

실험구성도 및 하드웨어

그림2.1과 같이 구성하였으며, 마이크로폰은 독일 Brüel Kjæl Type 2671이고, LPF는 NF Electronic Instrument FV-664(fc는 15KHz)이다. A/D는 AD2838로 삼흥데이터 시스템제품을 사용하였으며, 이때 Sampling Time은 40KHz로 하여 데이터획득을 시도하였다.

한 사람의 화자를 이용하여 4명의 청취자로 하여금 감정을 인식할 수 있도록 반복하여 훈련한 후에 같은 문장을 3회 반복하였으며, 3회 반복된 문장을 대상으로 분석하였다. 특징데이터 획득방법은 문장의 길이(장단), 에너지분포(음의 강도,음의 고저)로 표현되는 국악의 원리를 분석 기초로 하였다.

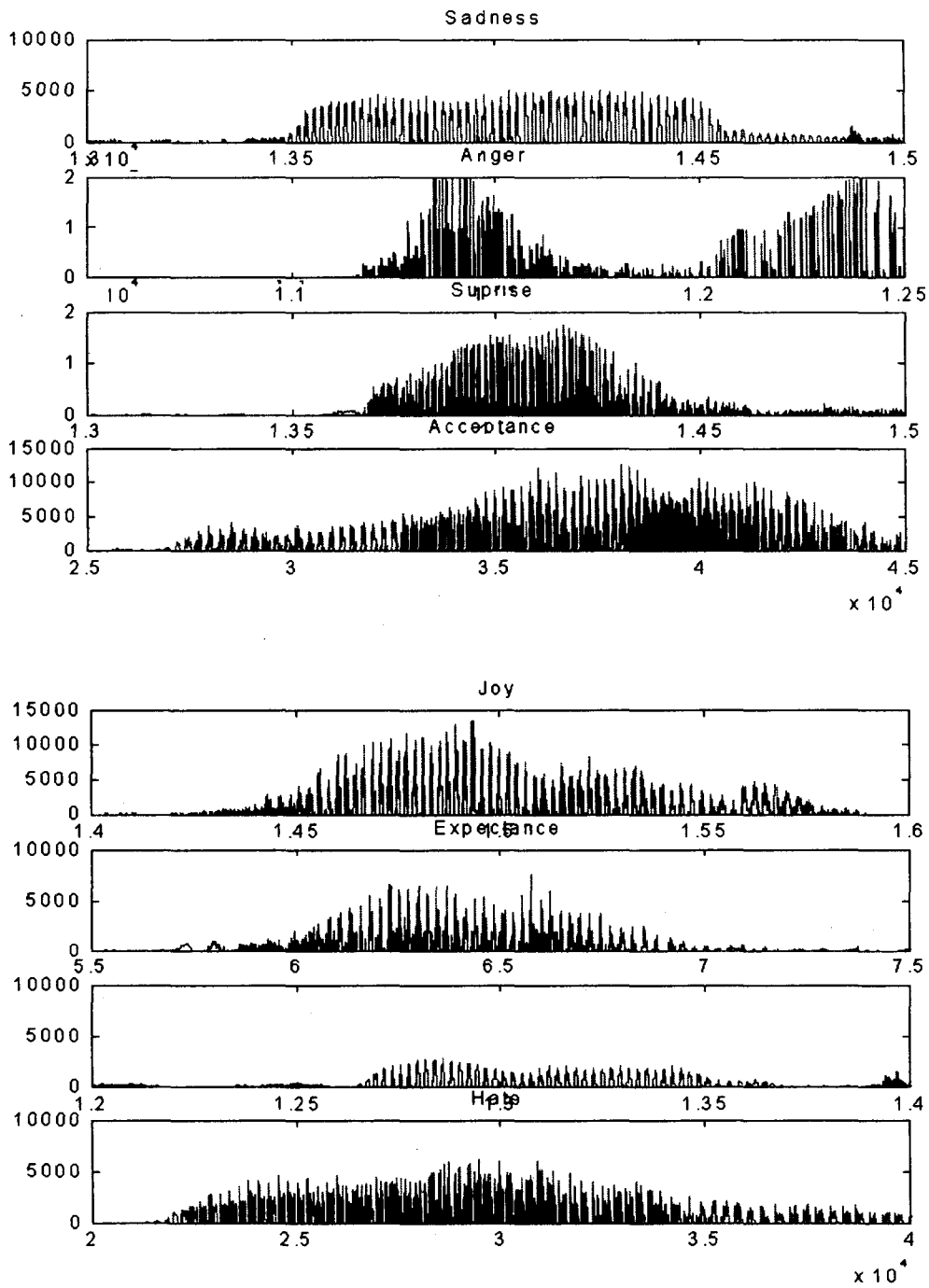


Fig. 4.1 Voice Signals

따라서 본 연구에서는 그림 4.1과 같이 시간영역 소스신호를 FFT를 이용하여 분석한 결과 그림 4.2와 같이 결과를 얻었다.

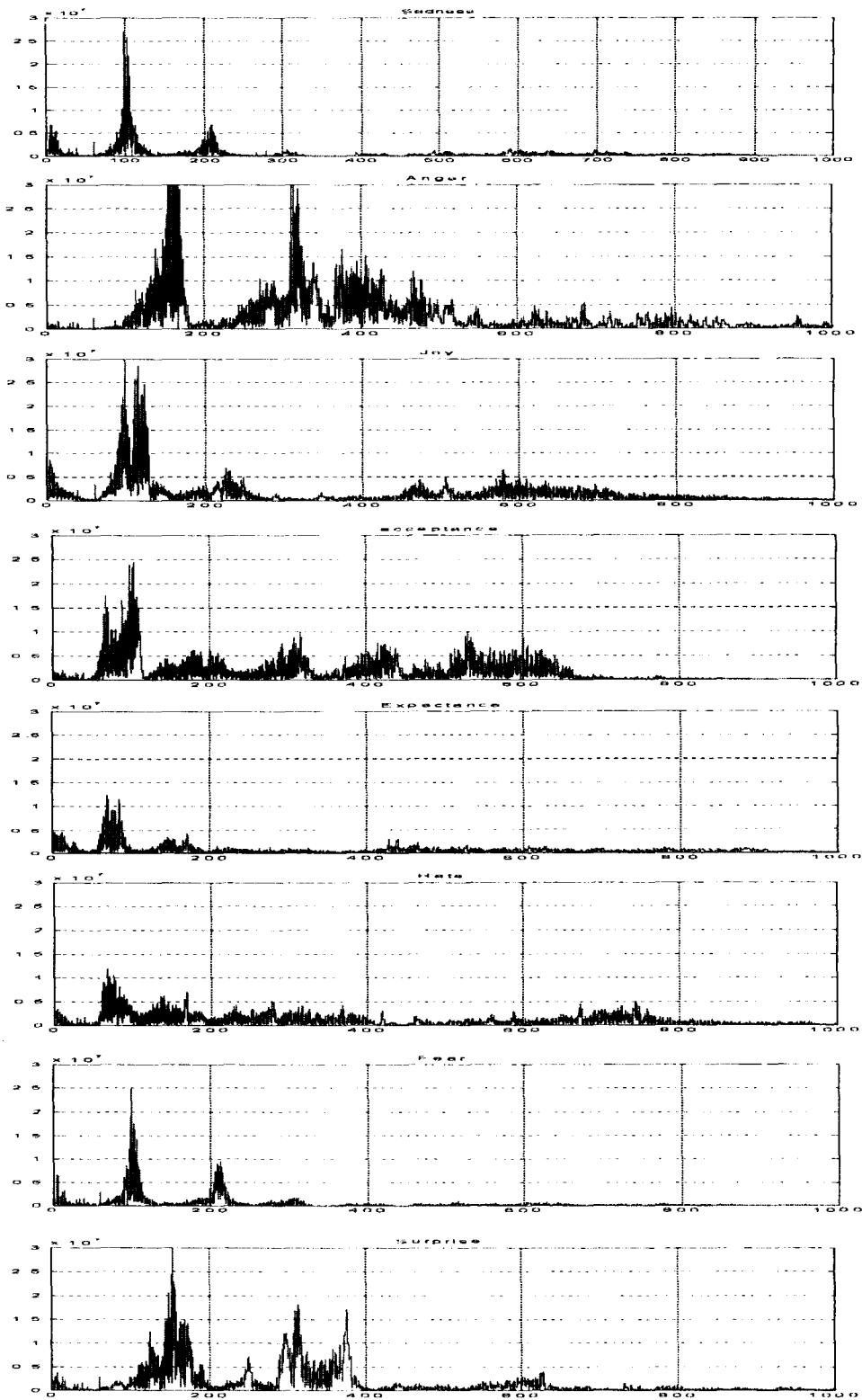
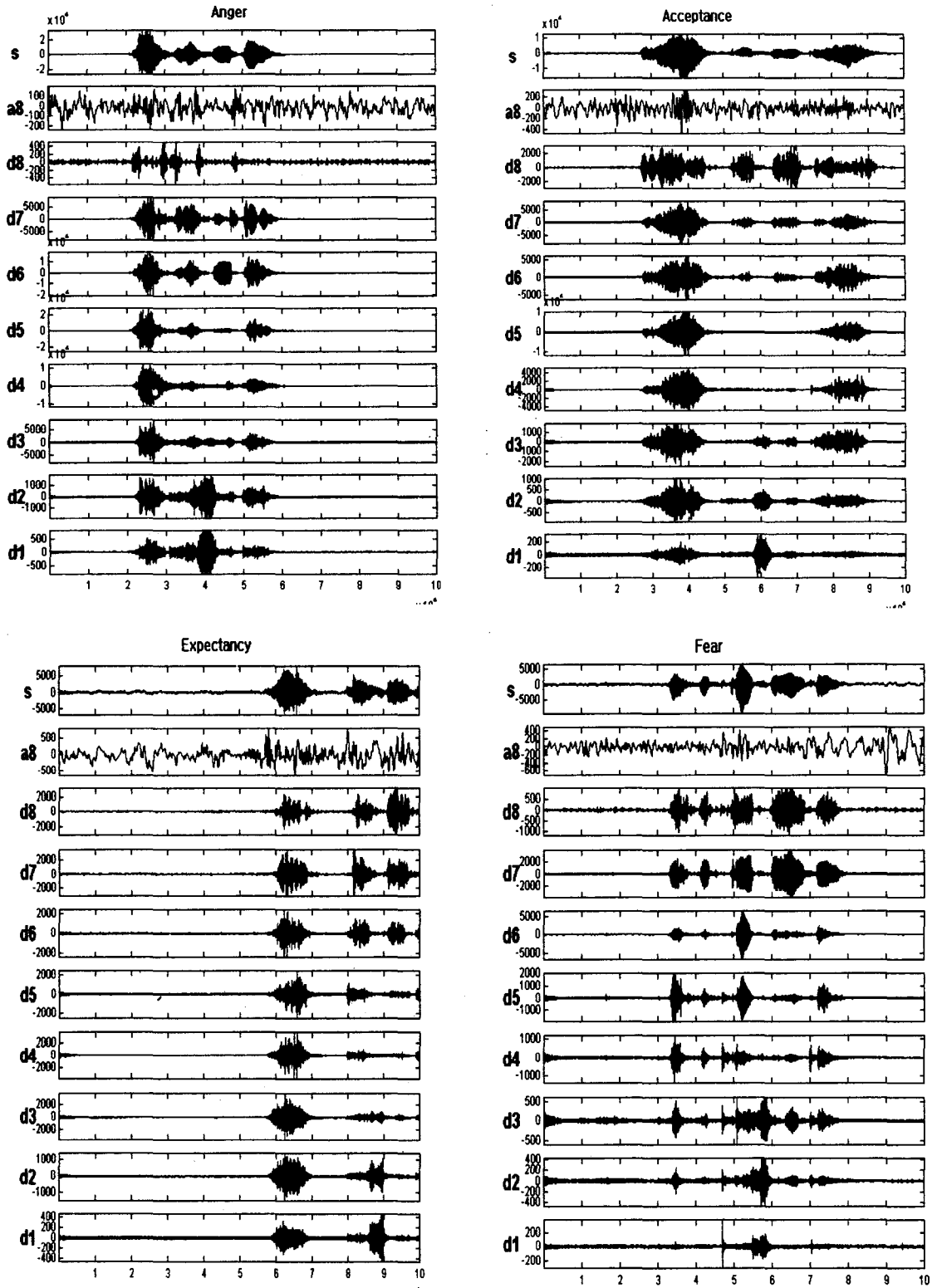


Fig 4-2 Spectrums of Voice signals.

다시 신호를 Wavelet 변환(db.5, Level 8)을 이용하여 분석한 결과 그림 4.3과 같은 결과를 얻었다.



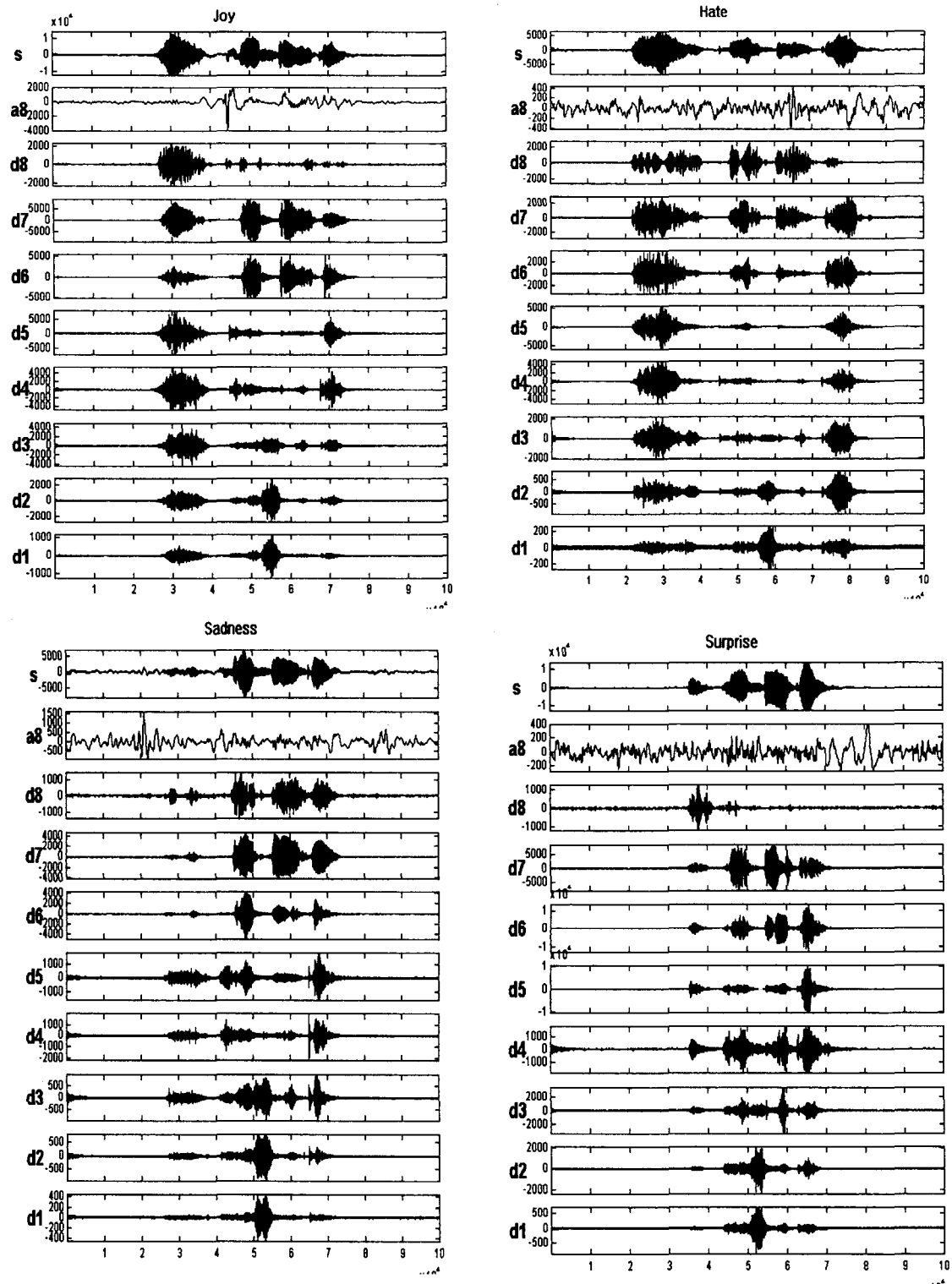


Fig. 4-3 Result signal of Wavelet Transform

그림4.3에서와 같이 감정의 종류에 따라 각 주파수대역에 분포되는 음성신호의 크기가 다르게 나타남을 알 수 있다.

본 실험에서는 퍼지-신경망을 이용하여 각 개인의 음성에 대하여 감정인식이 가능한 모델설계에 대한 연구로, 특징 데이터를 이용하여 설계된 모델을 학습하고, 인식이 가능한지의 여부를 알기 위하여 학습데이터와 실험데이터를 분리 실험하여 결과를 고찰하였다. 그림4.2의 FFT를 이용한 주파수 스펙트럼 분석결과와 같이 저주파성분과 고주파성분의 분포비와 에너지 분포를 통해 8개의 기본감정으로 분석한 결과는 아래 표4.1과 같다.

Table 4.1 Characteristic data of voice signals(in FFT)

Characteristics Emotions	Intensity	Period of tones (Samples,40KHz)	Frequency(Hz)	
Surprise	284	7000	170	300
Anger	442	6000	160	320
Sadness	203	11000	100	210
Expectancy	166	10000	70	160
Acceptance	308	17000	80	520
Joy	312	13000	110	220
Hate	200	12000	70	720
Fear	162	8000	100	220

발음 중에서 음의 길이는 “아! 그렇습니까?”에서 “아!” 부분의 길이를 샘플링 수로 결정하였다.

Table 4.2 Characteristic data of voice signals(in Wavelet)

Emotions Level	Surp.	Anger	Sad.	Expect.	Accept.	Joy	Hate	Fear
	3000	22000	900	5000	10000	10000	5000	2200
d8	1000	300	300	3000	2100	2000	1000	600
d7	1000	7000	500	3000	6000	6500	2000	1800
d6	2000	18000	400	2000	6000	5000	2100	2000
d5	2500	21000	300	1900	10000	4000	4500	1800
d4	2000	10000	200	2100	4000	4000	4000	800
d3	800	6000	300	2100	1200	2000	1800	260
d2	100	1000	100	1100	1000	1000	400	180
d1	20	600	50	400	200	200	40	40

이 결과를 이용하여 인식모델을 설계하고 결과를 분석하여 그림4.4에 나타낸 것과 같이 89%의 인식률을 보여주고 있다.

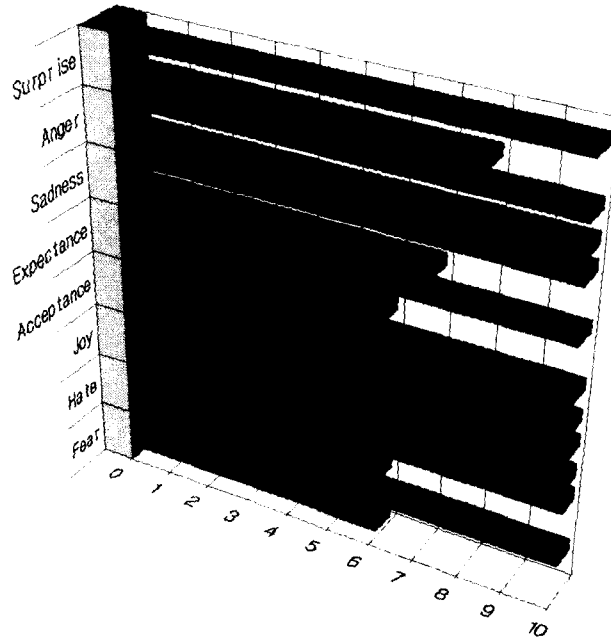


Fig 4.4 Result of Simulation

V. 결론

이상과 같이 본 연구에서는 개인의 감정을 인식하는 감정인식 모델에 대하여 제안하였다. 제안한 모델은 입력으로 주어지는 감정요소를 갖는 개인의 음성신호를 입력하여 특정인의 감정인식에 적용하였다. 연구 결과 복잡한 감정인 Hate, Acceptance, Expectance 등의 감정은 인식률이 낮음을 확인할 수 있었지만 89%의 인식율을 보임으로써 가능성을 확인 할 수 있었다. 서론에서도 언급했다시피 현재 인터페이스 기술의 발전은 인간의 감정적인 부분까지도 커버할 수 있기를 기대하고 있다.

집안에서 편하게 각종 기기들을 컨트롤하고 집안의 여러 프로세서들이 알아서 유저를 최적의 환경상태에 놓이도록 환경을 조절해주고 편의를 도모해줌에 더불어 감정적인 면에서의 환경까지도 사람에게 맞추어 조절해 줄 수 있는 컴퓨터의 개발을 위해서도 감정인식에 대한 연구는 큰 의의를 가지고 있으나 데이터 수집의 어려움, 토대로하고 있는 음성의 심리학 연구의 합리성, 타당성 여부 검증등 많은 어려움이 산재해 있다. 이러한 현 시점에서 본 연구는 음성에 의한 감정인식 시도자체에 큰 의의를 부여할 수 있겠다. 데이터 수집에 따라 다양한 계층, 다양한 집단의 특징을 반영할 수 있는 감정인식모델의 개발이 가능하다. 나아가 본 연구의 내용은 능동적인 인터페이스 구축의 기술개발을 위한 기초연구로서 심분 활용될 수 있을 것으로 사료된다.

참고문헌

Jennifer Healey, Rosalind Picard, "Digital Processing of affective signals", picard@
media.mite.edu

L.W.Cambell,D.A.Becker,A.Azarbayjani,A.F.Bobick,and A.Pentland,"Invariant feature for
3-D gesture recognition", in proceedings, International Conference on Automatic face
and Gesture Recognition, Pp 157-162,Killington,VT,1996,IEEE

Elias Vyzas and Rosalind W. Picard," Offline and Online Recognition of Emotion
Expression from Physiological Data", MIT Media Laboratory,
e-mail:Picard@Media.mit.edu