

신경회로망을 이용한 다중 전극 와우각 이식 시스템용 음성처리 알고리즘

최진영·조진호*·이건일

A Neural Speech Processing Algorithm for Multielectrode Cochlear Implant System

Jin-Young Choi, Jin-Ho Cho*, and Kuhn-Il Lee

- Abstract -

A New speech processing algorithm using neural networks is proposed. We transform input data into frequency domain and process them by neural networks of 22 output neurons which have Bark scale on the ground that the Bark scale is similiar with that of the characteristics of human cochlea. An utilized neural network is multilayer perceptron, and the characteristics of coehlea have it trained by error back propagation learning algorithm. The trained neural networks suffices functions of human cochlea including the effects of automatic gain control, compression and equalization. Simulation results show that the proposed speech processing algorithm has good performance in automatic gain control, compression and equalization.

1. 서론

인간의 내이에 존재하는 와우각(cochlea)은 일종

〈접수 : 1990년 4월 30일〉

경북대학교 전자공학과

* 경북대학교병원 의공학과 및 경북대학교 전자공학과

Dept. of Electronics, Kyungpook National University

* Dept. of Biomedical Eng., Kyungpook University Hospital and Dept. of Electronics, Kyungpook National University

의 주파수 분석기의 역할을 한다. 음성정보는 와우각에서 제각기의 주파수 성분별로 각 청신경섬유를 통하여 뇌에 전달됨으로 음성인식이 된다. 사람의 귀가 그 부피에 비하여 감도 및 다이내믹 레인지에 서나 소리의 세기나 높이에 대하여 인공장치에 비할 수 없는 정확한 식별력을 가질 수 있는 것은 인간의 와우각 및 신경의 전달체계 그리고 뇌의 학습에 의한 판단기능의 복합적 결과로 볼 수 있다. 최근에 전자공학의 비약적 발전으로 인하여 전자장치를 이용한 인조기관의 개발에 대한 연구가 활발해 지고 있다. 이 중에서도 특히 와우각 이식 시스템(cochlear implant system)은 이를 청각이 불구가 된 환자에게 시술할 경우 청력의 소생을 기대할 수 있어서 그 이용이 매우 기대되는 것이다. 와우각 이식 시스템은 입

려된 음성을 귀의 특성에 맞도록 전기적 변위로 바꾸어주는 음성처리 부분, 청신경 말단부위를 자극시키는 부분 그리고 생체조직과 접합시키는 전극 부분으로 나누어진다. 1985년 White 등¹⁾은 다중전극 아날로그 음성처리를 구현하여 와우각 이식 시스템에 응용하여 시험하였으나 음성처리에서의 음성정보의 손실로 인하여 그 인식율이 낮았다. 1988년 Mead 등²⁾은 CMOS VLSI 기술로 선형 2차 여파기 480개를 직렬 연결하여 아날로그 전자 와우각(analog electronic cochlea)을 구현하였다. 그러나 이 시스템은 선형 시불변 시스템으로서 와우각의 중요한 기능중 하나인 자동이득조정 기능이 없는 단점이 있다. 1989년 Morris 등³⁾은 고계산율 및 저전력소비 DSP(digital signal processing) 칩을 사용한 다중전극 시스템에 대한 연구결과를 보고하였다. 그러나 이는 기존의 음성처리기들에서와 같이 자동이득조정, 압축 및 등화(equalization)등을 외부에서 조정해 주어야하는 문제점을 갖는다. 한편 1986년 Wilson 등⁴⁾은 음성처리기가 스펙트럼의 해상도가 높을수록 인식율도 높아짐을 발표하였다. 따라서 최근에는 가청대역 스펙트럼을 가지는 음성처리기가 필요하게 되었다.

본 논문에서는 기존의 음성처리기가 가지고 있는 문제점을 해결하기 위하여 최근에 연구가 활성화되고 있는 신경회로망을 사용하여 새로운 음성처리 알고리즘을 제안하였다. 제안된 음성처리 알고리즘은 신경회로망의 학습기능을 이용하여 기저막(basilar membrane)의 간격과 비슷한 바크 스케일(Bark scale)⁵⁾로 10(kHz)까지의 대역을 22개의 전극으로 할당하였으며 시뮬레이션을 통하여 자동이득조정, 압축 및 등화의 특성에 있어서 양호한 결과를 얻을 수 있음을 보였다.

2. 이 론

귀의 구조는 그림 1과 같이 외이와 중이 및 내이의 3부분으로 나누어진다. 귓바퀴에서 집음된 소리는 귓구멍을 통과한 후 고막을 진동시키게 된다. 중이는 공기의 진동을 액체진동으로 변환시키는 음향변환기로서 고막과 난원창(oval window)를 연결하며 큰소

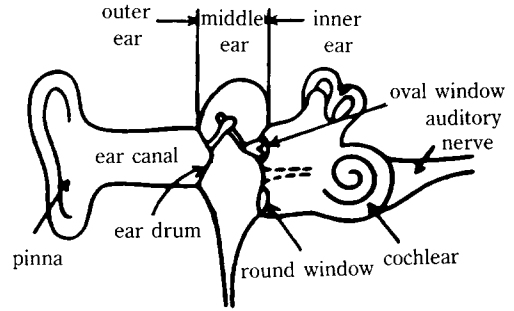


그림 1 사람 귀의 구조
Fig. 1 Structure of human ear.

리가 들어올 경우 근육이 수축하여 소리의 크기를 축소하는 보호기능이 있다. 내이는 중이로 부터 전달된 압력을 신경신호로 바꾸어주는 역할을 하며 음파가 2와 5/8바퀴 꼬인 와우각을 진행함에 따라 고주파가 감쇄하는 여파기의 특성을 보이며 내부의 유동액의 움직임에 의하여 헤어 셀(hair cell)이 이러한 기계적 움직임에 따른 전기적 저항이 변함에 따라 각 세포간에 120~140 mV 정도의 전위차가 생기게 된다. 이때 다중성분의 입력신호가 들어오에 따른 헤어 셀의 스위치와 같은 특성으로 인하여 매스킹 효과가 나타나게 된다.

와우각 이식 시스템은 헤어 셀은 파괴되고 없으나 청신경이 대부분 살아 있을 경우에 시술하여 청각기능을 회복시켜 줄 수 있다. 와우각 이식 시스템은 단일 전극 시스템과 다중 전극 시스템으로 나눌 수 있다. 단일 전극 시스템은 음의 높이보다는 음의 크기의 변화를 인식하므로 낮은 진동음(buzz) 혹은 땡그랑(clang)같은 정도의 음을 인식하고 300 Hz 이상에서는 음의 크기가 감소하여 1000 Hz 이상의 음은 인식하기 어려우므로 소리의 유무와 시간에 따른 전체적인 세기의 변화정도만 인식하는 시스템⁶⁾이다. 그러나 청신경 시스템의 손상이 크거나 청각이 발달되지 못한 사람에게는 다중 전극보다는 단일 전극 시스템을 사용하게 된다. 다중 전극 시스템은 대역폭이 300(Hz)를 넘을 때 사용하게 되며 각 전극마다 특정한 주파수 범위를 가진다.

그림 2는 Patric 등⁷⁾이 제안한 음성처리기이다. 마이크를 통하여 입력된 음성은 청취자의 상태에 따라

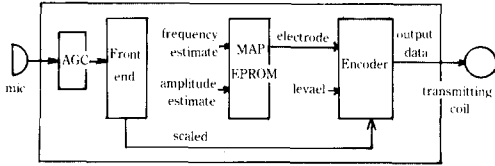


그림 2 멜버른형 음성처리기의 블록 다이어그램
Fig. 2 Melbourne type speech processor block diagram.

외부에서 이득이 조정된 후 신호의 피크가 적당한 크기가 되도록 엔코더에서 신호의 레벨이 조정된다. 이 신호는 기억장치(EPROM)에 미리 저장된 적당한 자극 매개변수로 바뀌어 전송 코일을 통하여 자극기에 인가된다. 즉 기존의 방식들은 외부에서 여러가지 요소들을 제어하여야 하는 불편함이 있는데 이는 신경회로망을 사용할 경우 쉽게 해결할 수 있다. 본 논문에서는 신경회로망의 활성화함수(activation function)⁸⁾인 출력 $f(z)$ 는

$$f(z) = \frac{1}{1 + e^{-z}}$$

와 같고 그림 3에서와 같은 sigmoid함수를 단일 뉴런의 전달함수로 사용하였다. 이로써 과도한 입력이 들어올 경우 출력이 포화되게하여 광범위한 범위의 입력에 대한 출력값이 0과 1사이의 값으로 압축하는 기능을 갖게하였다. 또한 신경회로망의 학습시 주파수에 따른 학습의 크기를 변화시킬 수 있으므로 등화의 기능을 가지는 음성처리기를 구성할 수가 있다. 본 논문에서 사용된 신경회로망은 그림 4와 같이 입

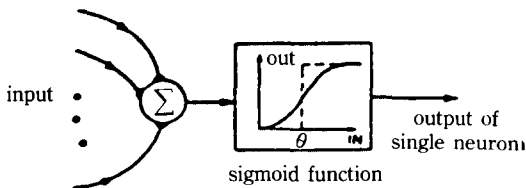


그림 3 sigmoid 함수를 사용한 단일 뉴런의 모형
Fig. 3 Model of single neuron using sigmoid function.

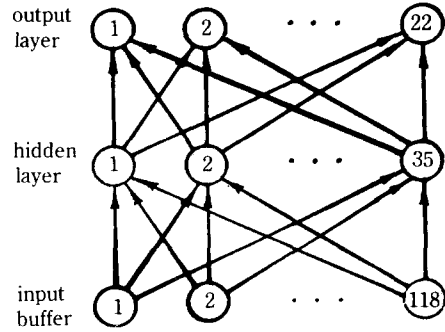


그림 4 역전달 네트워크
Fig. 4 Back propagation network.

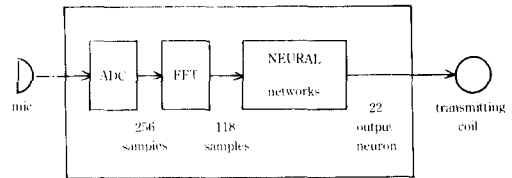


그림 5 신경회로망을 사용한 제안된 음성처리기
Fig. 5 Proposed speech processor using neural networks.

력비퍼, 은닉층(hidden layer) 및 출력층(output layer)으로 구성되며 신경회로망의 학습은 오차 역전달(error back propagation) 알고리즘을 사용하여 귀의 특성을 고려한 신경회로망을 구성하게 된다.

제안된 음성처리 알고리즘에서는 그림 5에서와 같이 마이크를 통하여 입력된 음성신호를 A/D 변환하고 푸리에 변환한 다음 118개의 입력비퍼에 인가하여 신경회로망의 22개의 출력이 각각의 주파수 범위를 가지게 된다.

이때 신경회로망의 각 출력뉴런이 가지는 주파수 범위는 기저막의 간격과 비슷한 바크 스케일로 할당하고 오차 역전달 학습방법을 이용하여 신경회로망을 학습시켰다. 여기서 바크 스케일 B는

$$B = 13 \times \arctan(0.76f) + 3.5 \times \arctan\left(\frac{f}{3.5}\right)$$

이고 f는 신호의 주파수(kHz)이다. 바크 스케일로 각 전극에 할당된 주파수 범위는

표 1 Bark scale로 각 전극에 할당된 주파수 범위

Table 1 Frequency range of Bark scaled electrodes.

elec-trode	Bark scaled frequency range(Hz)	Bark scaled input buffer of neural networks
1	78-156	1
2	157-234	2
3	235-312	3
4	313-391	4
5	392-469	5
6	470-625	6-7
7	626-781	8-9
8	782-938	10-11
9	939-1094	12-13
10	1095-1250	14-15
11	1251-1484	16-18
12	1485-1719	19-21
13	1720-1953	22-24
14	1954-2344	25-29
15	2345-2734	30-34
16	2735-3203	35-40
17	3204-3828	41-48
18	3829-4531	49-57
19	4532-5391	58-68
20	5392-6406	69-81
21	6407-7578	82-96
22	7579-9141	97-118

Table 1과 같다.

3. 시뮬레이션 결과 및 고찰

본 논문에서 제안한 음성처리를 사용하여 입력 주파수에 대한 출력의 상태를 시뮬레이션을 통하여 나타내었다. A/D 변환된 256개의 입력 음성자료는 푸리에 변환되어 118개의 신경회로망 입력버퍼로 인가되고 은닉층을 지나서 22개의 출력뉴런을 통하여 결과를 내보내게 된다. 1400~2400(Hz)의 입력신호에 대한 13번째 전극의 출력을 그림 6에 나타내었다. 13번째 전극의 바크 스케일에 의한 할당 주파수 범위의 입력신호가 인가될 경우에 출력뉴런이 바르게 반응함을 확인할 수 있다. 입력신호의 크기가 달라

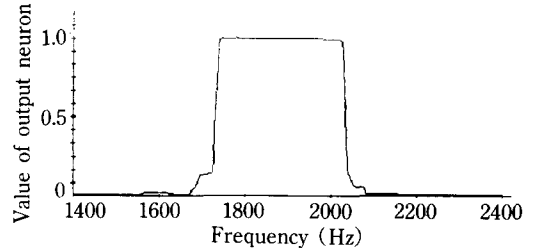
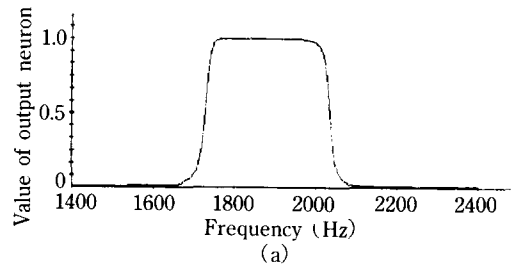
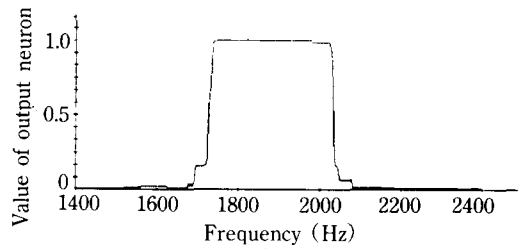


그림 6 1400~2400(Hz)의 입력 신호에 대한 13번째 뉴런의 출력

Fig. 6 Output of 13th output neuron when 1400~2400(Hz) input signal are transferred.



(a)



(b)

그림 7 입력 신호의 크기가 달라짐에 따른 출력 뉴런의 특성 (a) 입력신호의 크기를 1/10배로 축소하였을 때의 13번째 출력 뉴런의 값, (b) 입력 신호의 크기를 4배로 크게 하였을 때의 13번째 출력 뉴런의 값

Fig. 7 Characteristics of output neuron when input signals are varying. (a) Output value of 13th output neuron when amplitude of input signal is 1/10th reduced. (b) Output value of 13th output neuron when amplitude of input signal is 4th amplified.

짐에 따른 제안된 음성처리의 성능을 확인하기 위하여 입력신호의 크기에 따른 출력의 변화특성을 그

림 7에 나타내었다. 입력신호가 1/10로 축소되었을 경우와 4배로 증폭되었을 경우에 대한 음성처리 알고리즘의 결과는 그림7(a), (b)과 같았다. 즉 기존의 음성처리기의 문제점인 자동이득조정, 입력신호의 압축 및 등화의 처리에 있어서 신경회로망을 이용한 제안한 음성처리 알고리즘은 신경 회로망의 학습기능 및 활성화함수의 특성을 이용함으로 외부 조정의 필요가 없이 음성처리 알고리즘 자체가 모든 기능을 수용할 수 있음을 시뮬레이션을 통하여 확인하였다.

4. 결 론

본 논문에서는 와우각 이식 시스템의 음성처리부에 기존의 대역 통과 여파기나 EPROM 등과 같은 것을 사용하지 않고 최근 연구가 활성화되고 있는 신경회로망을 이용하여 음성처리부를 구성하였다. 인간의 귀가 주파수 성분으로 음성을 인식하는 것과 같이 입력음성을 푸리에 변환한 후 신경회로망으로 입력시켰으며, 출력뉴런을 통하여 그 결과를 검토하였다. 출력 전극은 기저막의 스케일과 비슷한 마크 스케일로 22개의 전극에 10(kHz) 까지의 주파수 범위를 할당하였다. 제안된 음성처리알고리즘으로 입력신호를 분석한 결과 기존의 음성처리기가 가지고 있었던 자동이득조정, 압축 및 등화의 문제점을 개선할 수 있음을 시뮬레이션을 통하여 확인하였다. 신경회로망을 이용한 음성처리 알고리즘으로 와우각 이식 시스템을 구현할 경우 기존의 음성처리기보다 높은 인식율을 얻을 수 있을 것으로 기대된다.

참 고 문 헌

1) R. L. White et al., "Development of a portable

speech processor for a multiple-channel auditory prosthesis," 7th Quart. Prog. Rep., NIH Contract NO1-NS-3-2357, June 1985.

- 2) R. F. Lyon, Caver Mead, "An analog electronic cochlea," IEEE Trans. Acoust., Speech, Signal Processing, vol. 36, pp. 1119-1134, 1988.
- 3) L. R. Morris, Peter Barszczewski, "Algorithms, hardware, and software for a digital signal processor microcomputer-based speech processor in a multielectrode cochlear implant system," IEEE Trans. Biomed. Eng., vol BME-36, pp. 573-584, No. 6, June 1989.
- 4) B. Wilson, C. Finley, and D. Lawson, "Speech processors for auditory prostheses," Fourth Quart. Rep., NIH Contract NO1-NS-5-2396, Sept. 1986.
- 5) Sadoki Furui, "Digital speech processing synthesis, and recognition," Marcel Dekker Inc., pp. 232, 1989.
- 6) G. E. Loeb, "Single and multichannel cochlear prostheses: rationale, strategies, and potential," Cochlear Implants, Raven Press, pp. 17-28, 1985.
- 7) J. F. Patrick, P. A. Croby, M. S. Hirshorn, J. A. Kuzma, D. K. Money J. Ridler, and P. M. Seligman, "Australian multichannel implantable hearing prosthesis," Cochlear Implants, Raven Press, pp. 93-100, 1985.
- 8) 최진영, 조진호, 이건일, "역전달 신경회로망에 대한 새로운 활성화 함수의 제안" 대한전자공학회 1989년도 추계종합학술대회 논문집, 제 12권, 제2호, pp. 165-168, 1989.