

# 청각 장애인을 위한 음성 신호의 자극패턴 추출에 관한 연구

채대곤<sup>o</sup> · 변정근 · 최두일 · 백승화<sup>\*</sup> · 박상희  
연세대학교 전기공학과 · \* 명지대학교 제어계측공학과

A Research on Speech Processing and Coding Strategy for Cochlear Implants

D. Chae<sup>o</sup> · J. Byun · D. Choi · S. Baek<sup>\*</sup> · S. Park  
Dept. of Electrical Eng. Yonsei Univ.  
\* Dept. of control & Instrumentation Eng. Myongji Univ.

## Abstract

A Study on the speech processing and coding strategy for cochlear implants have been developed to create a speech signal processing system which extracts stimulus parameter including formants, pitch, amplitude information. In this study we have presented the method which extracts characteristic information of speech signal and adapt patients with hearing handicap.

## 1. 서론

청각 보철은 청각 신경을 전기적으로 자극함으로써 소리를 듣는 효과를 내는 시스템을 말한다. 청각 보철 시스템은 매우 다양한 방식들이 개발되어 왔으나 최근에는 다중 채널 방식의 무선 통신 방식이 주로 쓰이고 있다. 후주의 Nucleus-22 시스템은 대표적인 다중 채널 청각 보철 시스템으로써 전 세계적으로 가장 호평받고 있으며, 국내에서도 세브란스 병원등에서 이 시스템을 구입하여 청각 장애인에게 이식하고 있다.

Nucleus-22 시스템은 음성의 특징 정보를 추출하는 음성 신호 처리부와 이 정보를 적절하게 코딩하여 전달하는 전달부, 내이에 이식된 receiver-stimulus 부로 나누어 진다. 본 연구에서는 Nucleus-22 시스템의 receiver-stimulus 부와 호환될 수 있는 음성 신호 처리부와 환자 적용 프로그램을 구현하였다.

## 2. 패턴 분석

Nucleus-22 시스템을 시술받은 환자의 귀에 고정되어 있는 외부 코일을 통해서 내이의 수신 자극기 ( receiver-stimulator ) 에 전달되는 정보를 분석해 보면 그림 1 과 같다. 전체 반송 주파수 ( carrier frequency ) 는 2.5 MHz 이며 포트 하나당 6개의 패턴을 보여 주고 있다.

그림 1 의 정보는 바이트 단위로 전송되어 내이의 수신 자극기에서 해독된다. 소프트웨어부분에서 추출된 음성의 특징 정

보는 일정한 함수 관계에 의하여 그림 2 와 같은 패턴으로 전송이 되어야 한다.

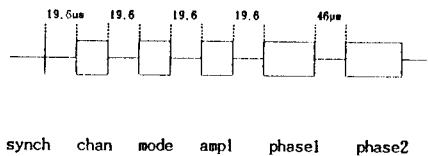


그림1 Nucleus-22 시스템 송신부의 신호 패턴

### 2.1 동기 신호

동기 신호는 4개의 cycle로 구성되며, 여러 신호 패턴의 동기를 맞추는데 사용된다.

### 2.2 채널 정보

1번에서 22번까지의 전극중 자극 전극을 선택한다. 선택된 전극에 해당되는 패턴의 cycle수는 아래의 공식에 따른다.

$$\text{The number of cycles} = 8 * \text{channel number} + 3$$

### 2.3 모드 정보(mode)

모드 정보란 환자의 내이에 이식된 전극중 자극 전극과 기준 전극의 간격을 나타내는 정보로서, 이에는 BP, BP+1, BP+2 모드가 있다. BP모드는 자극 전극과 기준 전극이 바로 이웃하는 경우이며, BP+1모드는 자극 전극과 기준 전극 사이에 전극이 하나 더 있는 경우이며, BP+2모드는 자극 전극과 기준 전극 사이에 전극이 2개 있는 모드이다. 그래서, BP모드에서는 19개의 전극을 사용할 수 있으며, BP+1모드에서는 20개, BP+2모드에서는 21개의 전극을 사용할 수 있다. 일반적으로 대부분의 환자는 BP+1모드를 사용한다.

### 2.4 amplitude 정보

음성 신호의 크기는 65이하의 크기와 65에서 229의 크기, 마지막으로 229이상의 크기로 나뉜다. 이중 정상적인 소리 크기인 경우에 ( 65-229 ) 네번째의 패턴은 고정어 되고 다섯번째와 여섯번째 패턴을 변형시킴으로써 음성의 크기를 환자가 인

식하게 되며, 아주 작거나 큰 소리인 경우 (65이하, 229이상)는 다섯번째와 여섯번째 패턴이 고정되어 되고, 네번째 패턴이 변하게 된다. 이와 같이 하는 이유는 인공 와우의 하드웨어적인 특성에 기인하는 것이므로 크게 고려할 바는 못 된다. 음성 신호의 크기는 다음과 같이 계산된다.

$$\text{음성 신호의 크기} = 25 * \log(\text{rms value})$$

### 2.5 width 정보

다섯번째와 여섯번째 패턴은 width 정보를 나타낸다. 여기서 중요한 것은 다섯번째 패턴과 여섯번째 패턴의 길이 바이트 수가 같아야 한다는 것이다. 그 이유는 환자의 내이에 시술된 전극은 커패시터가 달려 있는데 다섯번째 패턴 신호가 해독되면 내이의 수신 자극기는 지정된 전극과 기준 전극에 전류를 보내게 되고, 여섯번째 패턴 신호가 수신 자극기에 의해서 해독되어 지정된 전극의 극성을 바꾸어서 전류를 보내므로 다섯번째 패턴 신호에 의해서 충전된 커패시터가 여섯번째 패턴 신호의 전송과 동시에 방전되는 것이다. 이는 환자의 안정성을 고려한 것이므로 매우 중요하다고 볼 수 있다.

## 3. 코딩 전략(coding strategy)

코딩 전략은 포먼트를 몇 개 추출하여 전송할 것인가, 또는 피치 정보를 어떻게 반영할 것인가에 관한 테크닉으로 볼 수 있다.

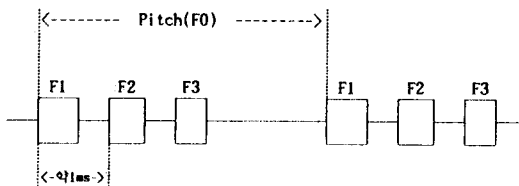


그림 2 본 연구의 코딩 전략

그림 2에 본 연구의 코딩 전략을 나타냈다. 본 연구에서는 포먼트를 3개 추출하므로 각 포먼트에 해당되는 세 개의 자극 패턴을 전송하도록 되어 있으며, 3개의 포먼트 F1, F2, F3를 한 묶음으로 해서 피치 간격으로 신호를 보내도록 설계되어 있다.

## 4. 환자 적응 프로그램

### 4.1 환자 적응의 필요성

인공 와우를 이식한 환자의 경우 환자마다 내이의 구조, 청각적 특성, 생리적인 특징이 다르므로, 22개의 전극을 사용하는 방식이 다르다. 즉, 어떤 환자의 들리는 전극이 다른 환자에게는 전혀 안 들리는 전극이 된다든가, 아니면 어떤 환자의 들리는 소리의 크기가 다른 환자에게는 못 알아듣는 소리가 될 수가 있다. 이런 결점을 보완하기 위해서 만들어진 것이 환자 적응 프로그램이다.

환자 적응 프로그램의 핵심은 환자의 내이 특성에 맞는 map 화일 작성 및 이를 이용한 mapping 전략에 있다. mapping에는 주파수 매핑 (frequency mapping) 및 T/C 매핑이 있다.

### 4.2 Mapping 전략

실제 환자의 map file은 다섯 가지 항목으로 구성되어 있다. 다섯 항목에는 T level, C level, Range, Mode, Active가 있다. T레벨은 소리가 들리기 시작하는 음성의 크기이며, C레벨은 듣기에 가장 편안한 상태의 소리 크기이며, Range는 T와 C의 차이이다. Mode는 사용 가능한 전극 갯수를 설정해주며, Active는 사용 가능한 전극을 표시해 준다. 표 1은 실제 환자의 map 화일을 보여 준다.

channel	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
T	99	122	122	131	130	136	153	147	143	141	130	125	115	108	116	102	100	89	73	
C	99	133	133	124	136	125	142	159	153	149	147	140	134	126	118	125	114	111	98	73
Range	0	11	11	2	5	5	6	6	6	6	10	5	11	10	9	12	11	5	10	
Mode	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1	SP-1
Active	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No

표 1. 환자 map

표 1의 환자의 경우에는 20개의 전극중 1번, 2번, 3번, 4번, 5번, 19번, 20번의 전극이 사용이 불가능하다. 그래서 사용할 수 없는 7개의 전극에는 자극 정보를 보내지 않고, 사용 가능한 13개의 전극에만 자극을 가하는 적응 프로그램이 필요하다.

### 4.2.1 Frequency mapping

전술한 바와 같이 주파수 매핑은 환자의 맵정보를 이용하여 사용 가능한 전극에만 자극을 주는 것뿐만 아니라, QMF 필터에서 추출되는 32개의 채널을 환자의 내이에 맞는 20개의 전극으로 바꾸어 주는 과정도 포함한다. 표 2에 이 과정이 잘 나타나 있다.

filter 의 channel	lower - upper	Elec	filter 의 channel	lower - upper	Elec
1	0 - 125	20	17	2000 - 2125	8
2	125 - 250	20	18	2125 - 2250	7
3	250 - 375	20	19	2250 - 2375	6
4	375 - 500	19	20	2375 - 2500	6
5	500 - 625	18	21	2500 - 2650	5
6	625 - 750	17	22	2650 - 2750	5
7	750 - 875	16	23	2750 - 2875	4
8	875 - 1000	16	24	2875 - 3000	4
9	1000 - 1125	15	25	3000 - 3125	3
10	1125 - 1250	14	26	3125 - 3250	3
11	1250 - 1375	13	27	3250 - 3375	2
12	1375 - 1500	12	28	3375 - 3500	2
13	1500 - 1625	11	29	3500 - 3625	2
14	1625 - 1750	10	30	3625 - 3750	1
15	1750 - 1875	9	31	3750 - 3875	1
16	1875 - 2000	9	32	3875 - 4000	1

표 2. 주파수 맵핑 전략

표 2의 방법으로 32개의 QMF 필터 채널을 20개의 전극으로 매핑을 한 후 환자의 map 을 조사하여 만일 사용하지 않는 전극이 있을 경우엔 그 전극의 앞, 또는 뒤에 있는 최단 거리의 전극이 자극될 수 있도록 하였다.

4.2.2 T/C mapping

환자의 map에는 사용 가능 전극 및 불가능 전극이 표기되어 있을 뿐 아니라, T, C로 표시된 값들이 배열되어 있다. 실제 음성은 그 크기가 1에서 239사이의 어느 값에 해당되도록 조절이 되어 있다. 그러나, 환자에 따라서 어떤 특정 채널에 대해서 들을 수 있는 소리의 범위가 다르므로 이를 조정할 필요가 생긴다. 예를 들어 표 1의 map을 갖는 환자의 경우 채널 7에서 T값이 136이므로 136보다 작은 크기의 소리가 입력이 되더라도 음성을 인식할 수 있게 하기 위해서는 이 소리의 크기를 T값으로 변환해 주어야 한다. 반대로 채널7의 C값은 142로 되어 있는데 이보다 큰소리가 입력이 될 경우에 환자는 괴로움을 느끼게 되므로 이 경우에는 소리의 크기를 C값으로 변환해 주어야 한다.

5. 자극 정보 추출

5.1 피치 추출 ( Pitch extraction )

음성 데이터를 10KHz 샘플링하고 약 20msec 구간마다 창함수를 리워서 자기 상관 함수 ( Autocorrelation function )를 구한다. 원음성의 주기와 자기 상관 함수의 주기는 일치하므로 자기 상관 함수의 주기를 구하여 이를 음성 데이터의 피치로 한다.

5.2 포먼트 추출 ( Formant extraction )

포먼트는 환자가 음성을 인식하는데 제일 중요한 역할을 한다. 본 실험에서는 세 개의 전극을 자극해야 하므로 포먼트를 세 개 ( F1, F2, F3 ) 구해야 한다. 포먼트 추출용 알고리즘으로는 일종의 서브밴드 코딩 ( subband coding )인 QMF 필터 ( Quadrature mirror filter )를 이용하여 32개의 채널로 나누는 기법을 이용했다. 또한, 일종의 전처리부로서 적응 알고리즘 ( LMS )을 이용하여 입력되는 음성 신호의 잡음 레벨을 추정하여 이를 스펙트럼 차감법 ( Spectral Subtraction Method )으로 제거하는 적응 스펙트럼 차감법을 제안했다

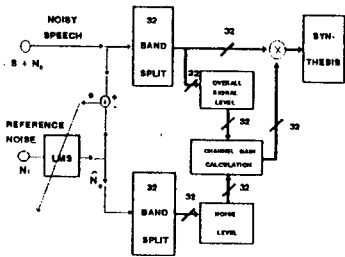


그림 3. 적응 스펙트럼 차감의 블록선도

필터의 각 채널을 통과한 배경 잡음이 제거된 음성 데이터의 에너지를 구하고 이 중 가장 큰 순서로 세 개를 구하여, 각 순서에 해당하는 채널 번호를 큰 것부터 제 1 포먼트, 제 2 포먼트, 제 3 포먼트로 한다. 이 때 이웃하는 채널 번호가 포먼트로 추출되지 않도록 조정해야 하고, 추출된 32개의 채널 번호를 내이의 전극 갯수인 22개로 적당히 대응시켜야 한다.

6. 실험 결과

전체 시스템의 블록선도는 그림 4와 같다. 필터뱅크에 입력되기 전에 음성 신호는 LMS 알고리즘부를 통과한다. 여기서 추정된 잡음은 QMF 필터뱅크를 거치면서 제거되며, 동시에 포먼트 및 에너지를 구하게 된다. 이 특징 정보를 환자 적응 프로그램에 적용하면 최종의 자극 패턴을 얻을 수 있다.

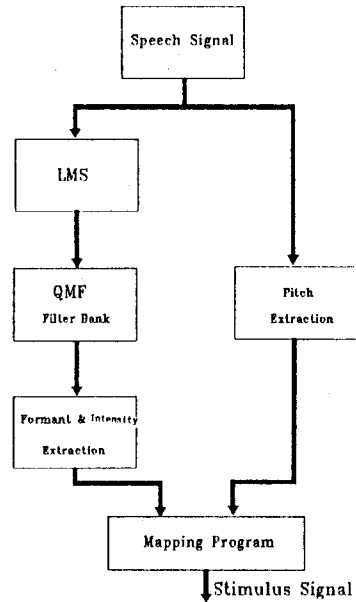


그림 4. 전체 시스템의 블록선도

그림 5, 그림 6에 실험에 사용된 모음 데이터 및 자음 데이터가 나와 있다.

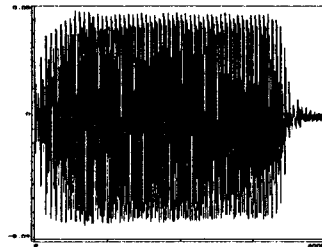


그림 5. 모음 /아/의 파형

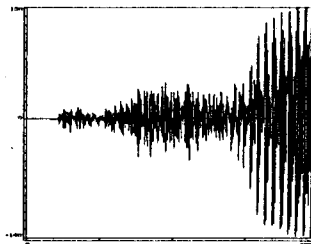
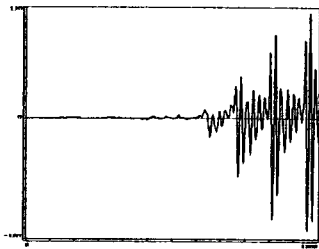
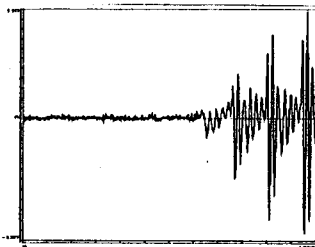


그림 6. 자음 /카/의 파형

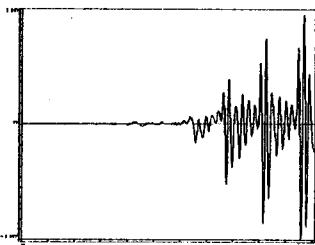
그림 7 은 적응 스펙트럼 차감법을 이용한 소음 제거 상태를 보여 주고 있다.



(a) 원신호



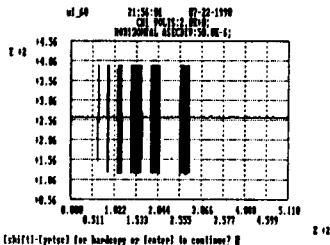
(b) 잡음이 첨가된 신호



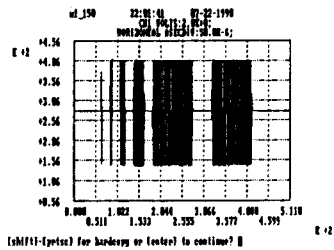
(c) 스펙트럼 차감법에 의해 개선된 신호

그림 7. 손상된 자음 /바/ 에 대한 원신호와 개선된 신호

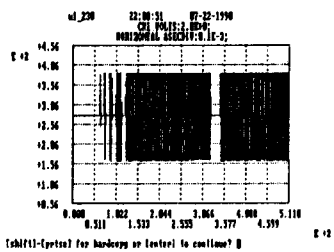
본 연구에 의한 음성 신호 처리기로 자극 신호 패턴을 추출한 결과는 아래의 그림 8, 그림 9 에 나타나 있다.



(a) 채널 1, 크기 60인 자극 패턴



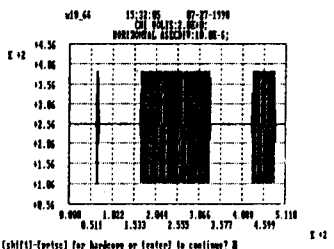
(b) 채널 1, 크기 150인 자극 패턴



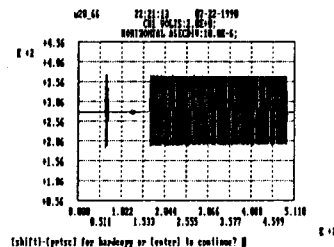
(c) 채널 1, 크기 230인 자극 패턴

그림 8. 채널1, 크기60,150,230인 자극 패턴

그림 8 은 채널이 같고 음성의 크기만 다른 자극 패턴들이다. 즉, 채널을 나타내는 두번째 패턴의 길이는 변하지 않고, 크기를 나타내는 다섯번째와 여섯번째의 길이만 변하게 된다. 그림 9 에는 반대로 크기가 고정되고 채널이 변하는 패턴을 표시했는데 여기서는 두번째 패턴인 채널의 길이만 변한다.



(a) 채널 10, 크기 66인 자극 패턴



(b) 채널 20, 크기 66인 자극 패턴

그림 9. 채널 10,20, 크기 66인 자극 패턴

그림 8 과 그림 9 는 일반적인 자극 패턴은 되지만, 특정 환자에 대한 자극 패턴은 못된다. 따라서, 표 3 과 표 4 에 표 1 의 map을 갖는 환자의 자극 파라미터를 나타냈다.

frame	elec1	amp1	elec2	amp2	elec3	amp3
1	17	102	16	116	18	100
2	18	100	16	116	18	100
3	18	111	16	116	18	100
4	17	114	16	116	18	100
5	18	111	16	116	18	100
6	17	104	16	116	18	100
7	17	102	16	116	18	100
8	17	102	16	116	18	100
9	17	105	16	116	18	100
10	17	102	16	116	18	100
11	17	114	16	116	18	100
12	17	102	16	116	18	100
13	17	102	16	116	18	100
14	17	102	16	116	18	100
15	17	102	16	116	18	100
16	17	102	16	116	18	100
17	16	116	18	100	15	108
18	16	116	18	100	15	108
19	6	130	6	130	6	130
20	17	102	18	100	16	116

표 3 모음 /아/ 에 대한 frame마다의 자극 파라미터

frame	elec1	amp1	elec2	amp2	elec3	amp3
1	6	130	6	130	7	136
2	6	130	9	147	6	130
3	7	136	6	130	6	130
4	6	130	9	147	6	130
5	6	130	7	136	9	147
6	6	130	9	147	6	130
7	6	130	6	130	7	136
8	6	130	6	130	6	130
9	6	130	7	136	6	130
10	6	130	6	130	6	130
11	6	130	6	130	7	136
12	6	130	6	130	6	130
13	6	130	6	130	7	136
14	6	130	6	130	6	130
15	6	130	6	130	9	147
16	6	130	9	147	6	130
17	18	100	18	100	17	102
18	18	100	18	100	18	100
19	18	111	18	100	16	116
20	18	111	18	100	18	100

표 4 자음 /카/ 에 대한 frame마다의 자극 파라미터

표 3 에는 모음 /아/ 에 대한 자극 파라미터가 기록되어 있다. 여기서 elec1은 제 1 포트에 대한 전극을 나타내며, amp1은 elec1에 가해지는 음성의 크기를 나타낸다. 19번 프레임만 제외하고는 모두 15 채널 이상을 자극하므로 모음 /아/는 내이에 시술된 코클리어의 1000 Hz 이하의 저주파 부분을 자극하게 된다.

표 4 에는 자음 /카/ 에 대한 자극 파라미터가 기록되어 있다. 16번 프레임까지는 자극 전극이 10 이하이므로 코클리어의 고주파 부분을 자극하며, 17번 프레임부터 20번 프레임까지는 모음의 영역인 저주파 부분을 자극하게 된다.

## 6. 결론

본 연구에서는 청각 보철을 위한 음성 신호 처리기의 구현을 하였다. 그 결과 호주제 Nucleus-22 시스템의 Receiver - Stimulator와 호환 가능한 자극 신호 패턴을 얻을 수 있었다. 앞으로 임상 실험을 통하여 환자에게 적합한 채널 및 크기의 매핑 함수를 찾고, 청각 장애인의 언어 변별에 미치는 영향을 조사할 계획이다.

## 참고문헌

1. Paul B. Ostergaard, Technical notes and research briefs, J. Acoust. Soc. Am 87(2), Feb. 1990
2. L. Robert. Morris Algorithms, Hardware, and Software for a Digital Signal Processor Microcomputer-Based Speech Processor in a Multielectrode Cochlear Implant System IEEE Trans. Biomed. Eng. Vol. 36, No6, June 1989
3. H. Mcdermott, B App Sc (Hon) Elec. Cochlear implant for simultaneous multichannel stimulation. grant R01-NS-21027 National Institutes of Health.
4. M. G. Bellanger, Digital filtering by polyphase network: Application to sample-rate alteration and filter banks IEEE, Trans. Acoust. Speech, Signal Processing, Vol. ASSP-24, pp 109-114, Apr. 1976
5. Steven, F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Trans. ASSP-27, NO. 2, pp113-120, 1979
6. 황성배, 최두일, 채대근, 김영선, 백승화, 박상희 " QMF를 이용한 청각 보철용 음성 신호 처리기의 실시간 처리 구현 " 대한 의용 생체 공학회, 1992. 11.