

## 피치동기 다중 스펙트럼을 이용한 청각보철장치의 음성신호처리 및 DSP 시스템 설계

신중인·박석준·신대규·이재혁\*·박상희

연세대학교 공과대학 전기공학과, \*경동대학교 정보통신공학부  
(1999년 2월 18일 접수, 1999년 7월 10일 채택)

### Speech Signal Processing using Pitch Synchronous Multi-Spectra and DSP System Design in Cochlear Implant

J.I.Shin, S.J.Park, D.K.Shin, J.H.Lee\*, S.H.Park

Dept. of Electrical Engineering, Yonsei Univ.,  
\*School of Information Technology Eng., Kyungdong Univ.  
(Received February 18, 1999, Accepted July 10, 1999)

**요 약 :** 본 연구에서는 내이의 손상에 의한 감각성 난청환자들의 청력회복을 위한 청각보철장치내의 가장 중요한 부분인 어음발생기의 음성신호처리 알고리즘 및 하드웨어를 개발하였다. 증폭, 저역통과 필터, AGC의 역할을 수행하는 외이 및 중이는 아날로그 시스템으로 모델링 하였고, 시간 지연된 다중 필터 및 변환기의 역할을 수행하는 내이는 실시간 처리가 가능한 고속 DSP 회로로 구현하였다. 특히 내이의 기지막특성은 미선형 다중 필터뱅크로 모델링한 후, 피치와 동기화된 다중 스펙트럼을 출력할 수 있는(pitch-synchronous multi-spectra : PSMS) 전략을 이용함으로써 청각계의 tonotopy와 periodicity를 만족시킬 수 있었다. 또한, 주요 음성신호처리의 대부분이 S/W로 수행되므로 다양한 실험을 위한 시스템 수정이 용이하며, C 언어로 프로그램이 개발되었기 때문에 다른 프로세서를 사용하는 H/W에도 쉽게 이식될 수 있다는 장점을 갖는다.

**Abstract :** We propose efficient speech signal processing algorithms and a system for cochlear implant in this paper. The outer and the middle ear which perform amplifying, lowpass filtering and AGC, are modeled by an analog system, and the inner ear acting as a time-delayed multi filter and the transducer is implemented by the DSP circuit which enables real-time processing. Especially, the basilar membrane characteristic of the inner ear is modeled by a nonlinear filter bank, and then tonotopy and periodicity of the auditory system is satisfied by using a pitch-synchronous multi-spectra(PSMS) method. Moreover, most of the speech processing is performed by S/W so the system can be easily modified. And as our program is written in C-language, it can be easily transplanted to the system using other processors.

**Key words :** Cochlear implant, Speech processor, Pitch-synchronous multi-spectra(PSMS) method, DSP system

## 서 론

청각계는 크게 외이, 중이 그리고 내이로 나뉜다. 선천적 또는 후천적으로 발생하는 청각계의 손상에 의한 청각장애의 유형은 외이와 중이의 이상으로 인한 전도성 난청(conductive disorder)

der)과 내이의 유모세포 손상에 의한 감각성 난청(sensorineural disorder)으로 분류할 수 있다. 전도성 난청인 경우에는 보청기나 인공이소골 등으로 상당한 회복이 가능하지만, 감각성 난청인 경우에는 기술적이고 실제적인 많은 어려움이 발생하지만 청각보철장치(그림 1)를 이용하여 사람의 음성을 어느 정도 구별하는 것은 가능하다[1-3].

일반적으로, 음성신호처리시의 변수들은 음원(sound source)의 정보(periodic, aperiodic, mixed)와 성도(vocal tract)의 전

통신저사 : 신중인, (120-749) 서울시 서대문구 신촌동 134 연세대학교 대학원 전기공학과 생체정보공학연구소  
Tel. (02)361-268 Fax. (02)312-7735  
E-mail. shinji@cyber.yonsei.ac.kr

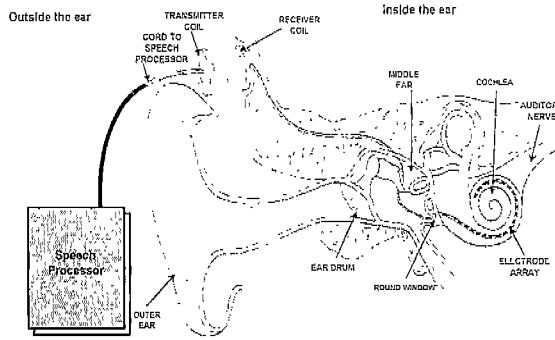


그림 1. 청각보철장치  
Fig. 1. Cochlear Implant

달함수에 관한 정보를 전달해야 하는데, 유성음(voiced speech sound)은 처음 2, 3개의 형성음(formants) 정도면 그 전달함수가 적절히 정의될 수 있으며, 피치(pitch) 정보를 함께 사용하면 음성의 분석이 가능하고, 무성음(unvoiced speech sound)은 전체적인 스펙트럼 형상의 정보면 충분히 인식 가능하다. 또한 그러한 변수들을 수~수십 msec 마다 갱신하면 인식 가능한 음성의 발생이 가능한 것으로 알려져 있다[4-5]. 이렇게 음성의 대략적 정보만으로도 정상 청각의 청취자들도 적절히 음성을 인식할 수 있기 때문에, 청각보철장치를 통하여 장애자들이 인식할 수 있는 음성의 전달이 가능하다. 감각성 난청환자들은 남아 있는 청각신경섬유(auditory nerve fibers)들에, 이러한 정보들을 전기적 신호로 변환시켜 전기적인 자극을 가함으로써 어느 정도의 청력이 회복될 수 있다. 이를 위하여 변환과정의 결과로 발생하는 임펄스 패턴들(impulse pattern)이, 음성신호의 중요한 정보들을 포함한 형태로 청각신경 내에서 전기적인 자극의 형태로 발생되어야 한다[6]. 외이 및 중이를 통하여 내이의 전정창에 유입된 음성신호는 달팽이관의 경계인 기저막(basilar membrane)을 따라 진행하는데, 기저막은 위치에 따라서 물리적 성질이 다르기 때문에 입력신호의 주파수에 따라 공진하는 위치가 달라진다. 기저막이 최대 공진을 일으키는 위치를 특성 위치(characteristic place), 이때의 주파수를 특성주파수(characteristic frequency)라 하며, 이 주파수 간격을 Bark scale이라고 한다. 1 Bark의 간격은 기저막 상의 약 1.5mm의 간격에 해당한다. 따라서 기저막의 한 점은 시간지연특성과 대역통과 필터 특성을 갖는데, 이러한 특성을 선기적으로는 필터뱅크 특성으로 모델링할 수 있다. 즉 기저막의 특성 위치에 해당하는 부분에 특성 주파수를 갖는 자극신호를 가하면 소리의 지각이 가능하다[7-8]. 현재 사용되고 있는 청각보철장치는 다음의 두 가지로 분류된다. 하나는 단일 채널(single-channel) 시스템으로서 음성신호의 시간적인 코드(temporal code)만을 전송하기 때문에, 음성과 같이 비정상적인(non-stationary) 신호의 정보를 충분히 전달하기 힘들다. 음성신호의 전달을 위해서 시간-주파수에 관련된 특성 정보들이 모두 요구되는데, 주파수

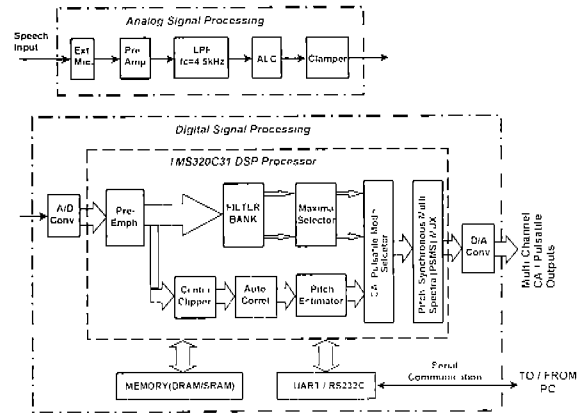


그림 2. 제안된 시스템의 전체 블록도  
Fig. 2. Block diagram of proposed system

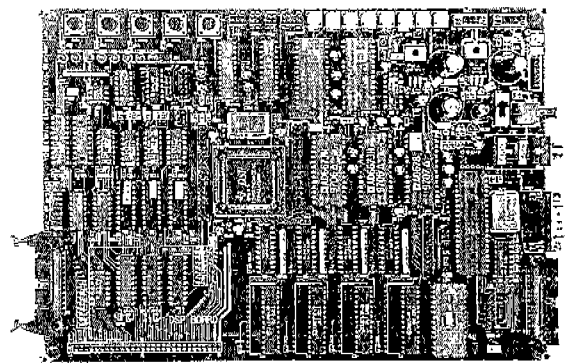


그림 3. 개발된 H/W의 외관 사진  
Fig. 3. Picture of developed H/W

에 대한 정보가 제대로 전달되지 못하기 때문이다. 다른 하나는 다중 채널(multi channel) 시스템으로서 달팽이관의 기저막의 위치와 관련된 특정한 주파수로 자극함으로써 tonotopy를 실현하고, 적어도 낮은 주파수에서는 시간적인 코드도 전송하는 시스템이다. 이러한 청각보철장치가 최적의 성능을 내기 위해서는, 전기적인 자극이 어떤 청각신경섬유의 집단에 선택적으로 가해져야 하며(place principle-tonotopy), 또한 적어도 4~5KHz까지의 음성신호의 시간적인 구조(periodicity)가 신경패턴에서 유지되어야 한다. 따라서 자극 코드는 place principle과 periodicity detection을 둘 다 사용해야 한다[9-10]. 이에 적합한 시스템이 다중 채널 시스템이며, 현존하는 많은 청각보철장치들이 이러한 방식을 이용하고 있다. 청각보철장치는 음성신호처리의 측면에서는 단일채널에서 다중채널로 발전되어 왔으며 자극 채널의 수는 8~22개등으로 각 연구 기관마다 상이한 양상을 보여왔으나, 환자의 개개의 특성을 고려한 전략의 선택, 시술후의 재활훈련등 선택적인 복합요인들이 잘 조화되어야 비교적 성공적일 수 있다는 것이 임상적으로 증명되고 있기 때문에 어떤

방식이 이상적이라고 결론을 내리기는 어려운 상황이다[2]. 세계적으로는 Nucleus, UCSF/Storz등에서 청각보철장치를 설계하여 이식하고 있으며[11-12], DSP를 이용한 시스템으로는 Motorola DSP56001을 이용한 시스템[3], AT&T의 TXP32CP DSP 개발도구를 이용한 실현적인 시스템[10]등의 연구가 있었다. 각 시스템들은 모두 독특한 방식으로 설계되어 왔으나 기술적이고 임상적인 어려움이 많기 때문에, 현재 사용되고 있는 청각보철장치와 음성신호처리의 대부분은 자연적인 신경 코드의 제한적인 양상만을 반영하고 있다. 따라서 가능하면 이론적 근거에 맞추고, 간단한 구조를 가지면서도 최대의 효과를 얻을 수 있는 전략의 수립이 필요한 것이다.

본 연구는 감각성 난청환자들의 재활을 위한 장치인 청각보철장치 내의 음성신호처리의 음성신호처리 알고리즘(algorithm) 및 그 H/W 구현에 관한 연구이다. 증폭, 저역통과 필터, AGC (automatic gain control, 자동이득조정)의 역할을 수행하는 외이 및 중이는 아날로그 회로로 구현하며, 시간지연된 다중 필터 및 변환기의 역할을 수행하는 내이는 DSP(digital signal processing) 회로로 구현한다. 특히 청각계의 tonotopy와 periodicity를 만족시키기 위하여, 음성신호를 Bark scale의 주파수 대역별로 분할하여 코딩하는 방식인 다중 필터 방식으로 모델링한 후 입력음성신호의 피치와 동기되는 자극 패턴을 출력하는 알고리즘을 개발함으로써, 청각계통의 생리학적 특성을 최대한 반영하고자 한다. 본 연구에 의하여 개발된 음성신호처리의 전체 블록도와 H/W는 각각 그림 2 및 3과 같다. 이 시스템은 전처리부인 아날로그 신호처리부와 주 신호처리부인 디지털 신호처리부로 구성되어 있다.

**음성신호처리 알고리즘**

내이의 특성을 만족시키기 위한 음성신호처리의 대부분은 DSP 프로세서 내에서 S/W로 제어된다(그림 2). 아날로그 회로에서 전처리된 음성신호를 입력받아 pre-emphasis 및 center-clipping으로 파형을 정형한 후, 그 신호에 자기상관을 취하여 이 신호를 근거로 피치를 검출한다. 한편, pre-emphasis된 신호를 다중필터로 통과시켜 비선형 구간의 주파수 대역별로 신호를 추출한 뒤, 상위 3개의 에너지를 갖는 대역만을 선택함으로써 place principle을 실현한다. 최종적으로 이 신호를 피치와 동기화 시킴으로써 periodicity를 만족시키는 자극패턴의 출력이 가능하다.

**1. 입력 음성신호의 전처리 및 피치 추정**

음성신호의 비정상성을 고려하여 신호처리의 기본 구간인 1 프레임(frame)을 25msec로 설정한다. 먼저 입력음성신호  $s(n)$ 에 대하여 식 (1)과 같이 pre-emphasis를 취하여 고주파성분을 강조함으로써 성대에서 성도를 거쳐 입술을 통해 음성신호가 방사될 때 발생하는 고주파 손실을 보상한다.

$$y(n) = s(n) - \alpha s(n-1) \tag{1}$$

여기서,  $\alpha$ 는 pre-emphasis의 정도를 조절하는 상수이다( $0.9 \leq \alpha \leq 1$ ). 이 pre-emphasis된 신호에 대하여 식 (2)의 center-clipping으로 spectrum flattening을 함으로서 입력신호의 주기성을 더욱 강조한다. 그리고 이 신호를 이용하여 구간간인 1 프레임 내의 피치를 검출한다.  $x(n)$ 을 center-clipper의 입력신호라 하면 center-clipper의 출력은 다음식과 같다 :

$$y(n) = C[x(n)] \tag{2}$$

$$C[x(n)] = \begin{cases} 0 & \text{for } C_L < x(n) < -C_L \\ x(n) & \text{for } else \end{cases}$$

또한,  $C_L$ 은 clipping 기준 레벨이다. 또한 자기상관은 다음의 식 (3)에 의해 수행된다 :

$$R_x(k) = \sum_{m=0}^{N-k-1} y(n+m) y(n+m+k) \tag{3}$$

여기서,  $N$ 은 1 프레임내의 샘플수 이다.

**2. 다중 스펙트럼 출력**

한편, 20Hz~4.5kHz구간에 대하여 Bark 단위에 기초한 8채널 다중 필터 뱅크(multi-filter bank)를 설계하여 pre-emphasis된 동일 프레임 구간의 신호에 대하여 각 주파수 대역별로 신호를 추출한다. Bark 단위는 기저막을 근사적인 등간격으로 구분한 것이며, 정신물리학에서 임계대역을 나타낼 때 사용하는 주파수 단위로서 E.Zwicker에 의하여 제안되었는데, 1 Bark는 청각계통의 정신물리학적 동조곡선에서 -3dB 점간의 임계대역폭을 나타낸다. E.Zwicker와 E.Terhardt[8]가 제안한 식(4), (5)에 의하여 다중 필터 뱅크의 중심주파수  $f_c[Hz]$ 를 환산하여 8채널의 다중 필터를 구현하였다.

$$f_c [kHz] = \frac{\tan\left[\frac{Z_c}{13}\right]}{0.76}, \quad f_c < 1kHz \tag{4}$$

$$(5) f_c [kHz] = 10^{\left[\frac{Z_c}{14.2} - \frac{8.7}{14.2}\right]}, \quad f_c < 1kHz$$

여기서  $Z_c$ 는 Bark 단위의 중심주파수이다. 8채널의 전극을 사용하는 것을 전제로 20Hz~4.5kHz사이의 구간에서 2.2 Bark의 등간격으로 중심주파수  $f_c$ 를 구하면 표 1과 같은 청각계통의 비선형 특성과 유사한 특성을 갖는 8개의 대역통과 필터를 구현할 수 있다.

그리고 maxima selector에서는, 다중 필터를 통하여 추출된 각 대역별 신호들에 대하여 대역별 에너지 계산을 하여 큰 값을 갖는 상위 3개의 대역만을 선택한다. 즉, 다중 필터 뱅크의 출력을  $FB_{j,k}(n)$ 이라 하면, maxima selector의 출력은,

$$S_{j,m} = MAX\{E[FB'_{j,k}(n)]\} \tag{6}$$

표 1. Bark scale을 갖는 다중 필터 특성

Table 1. The characteristic of the Bark-scaled multi-filter

Bark scale	2.2	4.4	6.6	8.8	11	13.2	15.4	17.6
$f_c(\text{Hz})$	225	463	732	1016	1452	2074	2964	4234
BW/2(Hz)	119	134	142	218	311	445	635	635

여기서,  $f$ 는  $f$ 번째의 프레임 ( $f \geq 1$ ),  $k$ 는  $k$ 번째의 채널 ( $1 \leq k \leq 8$ ),  $n$ 은 1 프레임내의 샘플수 ( $0 \leq n < 250$ ),  $m$  ( $1 \leq m \leq 8$ )은 상위 3개의 최대값을 갖는 채널번호를 의미한다. 여기서 선택된 3개의 채널은 에너지가 큰 주파수 부분을 의미하므로 음성신호의 주파수 특성을 고려할 때, 형성음 정보를 포함한다고 할 수 있다. 또한, 피치 추정기의 출력은  $P_f = \text{constant}$  이므로 최종 multiplexer의 출력  $M_{f,i}(n)$ 은,

$$M_{f,i}(n) = \begin{cases} FB_{f,i}(n) & \text{for } n = P_f(n) \times i, \text{ if } k = m \\ 0 & \text{if } k \neq m \end{cases} \quad (7)$$

라고 할 수 있다. 여기서,  $i$ 는 정수이다. 이렇게, 선택된 3개의 대역에 대해서만 앞서 계산된 피치정보와 동기화된 신호를 출력하도록 하면, 형성음 정보와 피치 정보가 포함된 다중 스펙트럼 정보를 얻을 수 있다(pitch-synchronous multi-spectra(PSMS) method). 입력음성의 intensity 정보는 환자의 특성에 따라 수용수준이 다를 수 있으므로 H/W 적인 조정작업으로 수행해야 한다. 한편, 달팽이관의 기저막은 시간지연이 된 필터뱅크의 특성을 갖고 있기 때문에 자극신호들은 채널에 따른 시간지연을 갖고 있어야 한다. 각 채널내의 동기자극 신호들에 대한 시간지연된 출력은 식 (8)에 의하여 계산된다.

$$f_{f,i}(n) = M_{f,i}(n - d_i) \quad (8)$$

여기서,  $d_i$ 는 채널별 지연요소( $\mu\text{sec}$ )이다. 최종 자극신호를 출력하는 방법은 analog(또는 CA, compressed analog) coding, pulsatile(또는 IP, interleaved pulsatile) coding 방식(그림 5)을 사용하거나 두가지의 복합적인 방법을 사용하는데, 어느 방식이 이상적이라고 결론을 내리기는 어려우며, 환자의 상태에 따라서 그 효과는 많은 차이가 난다[2]. 본 연구에서는 필요에 따라서 CA/Pulsatile mode selector에서 필요한 방식을 선택적으로 출력할 수 있도록 하여 시스템의 유연성을 증가시켰다.

### DSP 시스템 설계

본 연구에 의해 개발된 H/W는 그림 3과 같다. 마이크로 입력된 음성신호는 전치증폭기(pre-amplifier)에서 증폭된 후, LPF에서 신호처리에 불필요한 대역이 제거된다. 그리고 입력 신호의 wide dynamic range를 narrow dynamic range로 압축하는 ALC(automatic level control)회로를 거친 뒤에 clamp 회로에서 DC offset을 조정한다. 아날로그 신호처리된 음성신호

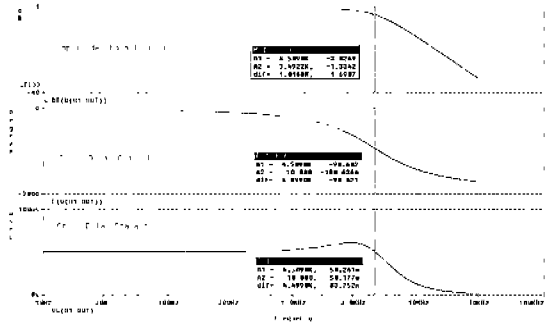


그림 4. LPF 특성곡선  
Fig. 4. Characteristic curve of LPF

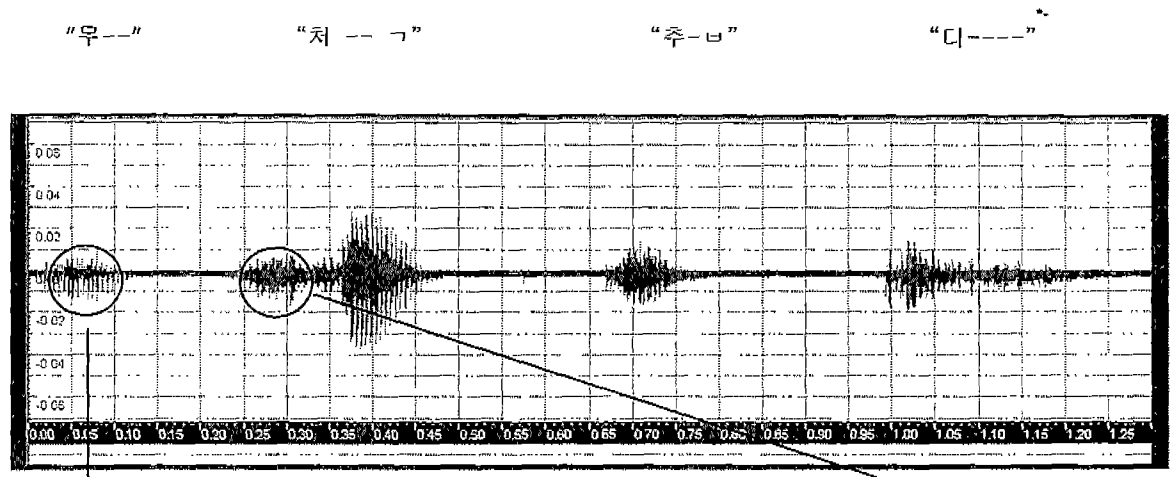
는 A/D 변환기에서 약 10kHz로 표본화되고 DSP 프로세서에 입력되어 S/W적인 신호처리를 거치게 된다. DSP 프로세서 내에서는 pre-emphasis, 피치 추정, 다중 필터 처리, 최대값을 갖는 대역의 검출등의 일련의 신호처리 과정을 수행한 뒤, 검출된 피치 및 다중 스펙트럼을 토대로 피치와 동기화된 다중 채널 자극신호를 합성하여 출력한다.

#### 1. Analog 신호처리부

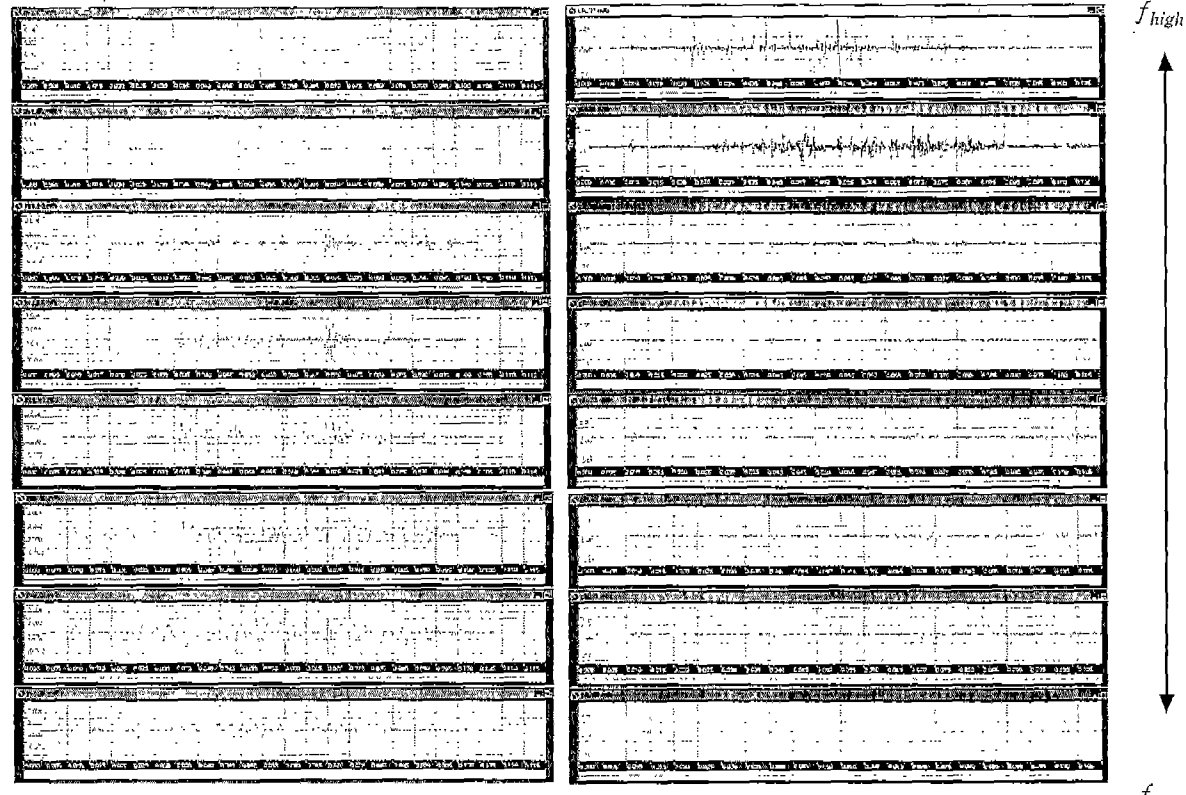
귀의 외이와 내이는 증폭, 저역통과 필터, 자동이득조절등의 기능을 수행하는데, 이에 대응하는 처리를 아날로그 신호처리부에서 행한다. 이 회로는 전치증폭기, LPF, ALC, clamp 회로 등으로 구성한다. 사람이 일상적으로 음성신호를 발생할 때, 신호의 대부분의 정보는 4kHz 이하에 존재하기 때문에 LPF의 차단주파수를  $f_c = 4.5\text{kHz}$ 로 제한한다. LPF는 2차 Butterworth active LPF( $f_c = 4.5\text{kHz}$ )이며, 진폭 특성, 위상지연 특성( $\tau_p$ ), 군지연 특성( $\tau_g$ )은 그림 4와 같다.  $f_c = 4.5\text{kHz}$  부근에서  $\tau_p = -90^\circ$ ,  $\tau_g = 50\mu\text{sec}$  이므로, 대체로 선형특성을보임을 알 수 있다. ALC 회로의 특성은  $G_{vo} = 80\text{dB}$ (개회로전압이득),  $R_{ALC} = 45\text{dB}$ (ALC 범위)로서 과도한 크기를 갖는 음성신호의 진폭을 일정수준이하로 제한한다. 또한 아날로그 입력이 일정한 DC offset을 갖도록 low-level clamp를 사용하여 최저 전압 수준을 고정한다.

#### 2. Analog / digital interface

아날로그 입력음성신호를 고속 디지털 표본화하기 위해서 Analog Devices(A.D.)사의 A/D 변환기인 AD678(12bits, 200ks/s,  $t_s = 5\mu\text{s}$ )를 사용한다. 음성신호의 대역을 4.5kHz로 제한하므로 D/A 할 때 발생할 수 있는 aliasing 현상을 방지하기 위해  $f_s = 10\text{kHz}$ 로 표본화 한다. Digital 신호처리부에서 신호처리가 끝난 최종 출력은 디지털 신호(12 bits) 및 아날로그 신호 두 가지로 출력 가능하며, 아날로그 출력을 위한 D/A 변환은 역시 A.D.사의 D/A 변환기인 AD667(1 channel, 12bits,  $t_s = 3\mu\text{s}$ ) 및 AD7247(2channel, 12bits,  $t_s = 10\mu\text{s}$ )를 사용했다.



(a)



유성음(voiced sound)

무성음(consonant sound)

(b)

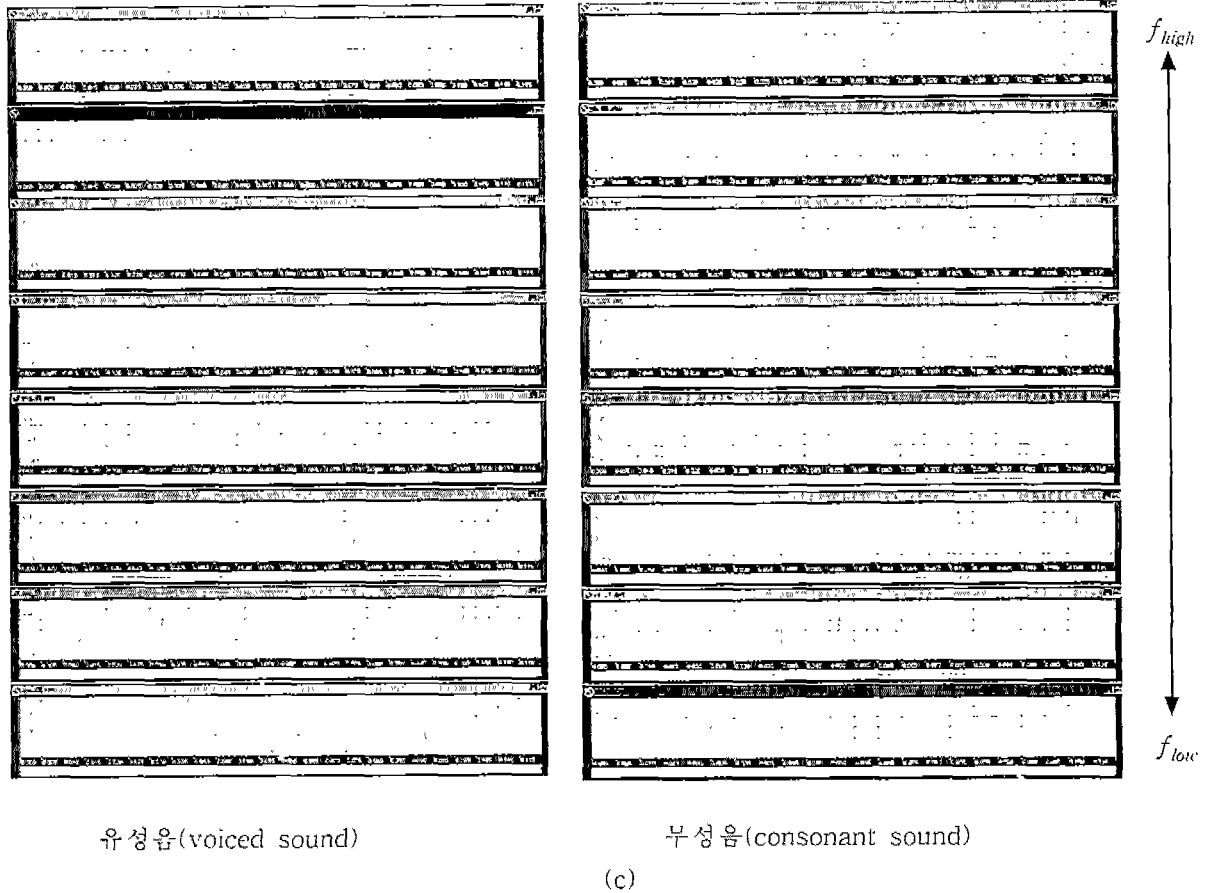


그림 5. CA방식과 pulsatile방식의 출력 비교  
 (a) 원신호 (13000샘플, 1.3초) (b) CA 방식 출력 (c) Pulsatile 방식 출력  
 Fig. 5. Comparison of output between CA and pulsatile method  
 (a) Speech signal (13000 sample, 1.3sec) (b) Output of CA method (c) Output of pulsatile method

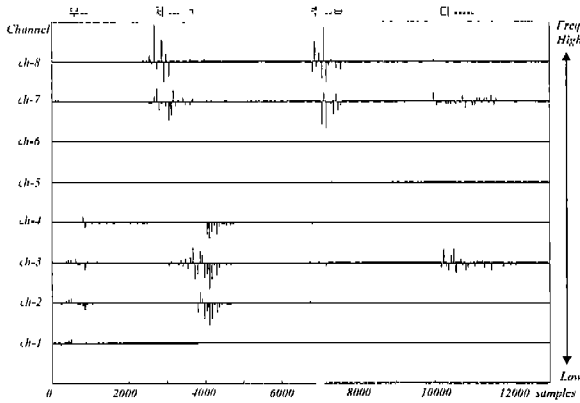
3. Digital 신호처리부

본 시스템은 입력된 음성신호에 대해 실시간 출력을 할 수 있는 고속연산이 요구되므로, 이를 위해 고속 DSP 프로세서를 사용한다. DSP 프로세서로는 T.I사의 32 bit 마이크로 프로세서인 부동소수점 전용 TMS320C31-40(40MHz, 40MFLOPs, 20MIPs)을, 프로그램 저장용으로 DRAM(256 kbits, 32kWords), 데이터 저장을 위해서는 SRAM(32kWords)을 사용하였다. 또한, 단독으로 동작할 수 있도록 reset/watchdog timer용으로 MAX1232를 내장했고, PC용 외부기기와의 접속은 RS-232C 직렬 포트를 이용할 수 있도록 설계하였다. Digital 신호처리의 대부분은 DSP 프로세서내에서 S/W로 수행된다[13-14].

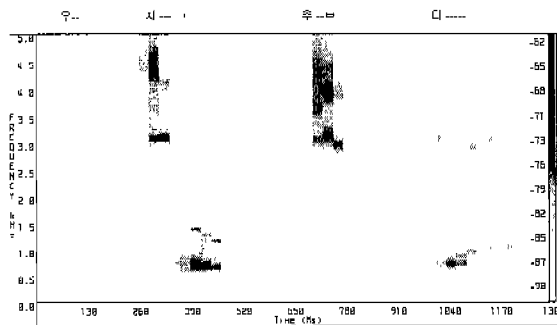
실험 및 결과 고찰

본 연구에서는 25세의 남성이 발음한 한국어 5개의 단모음

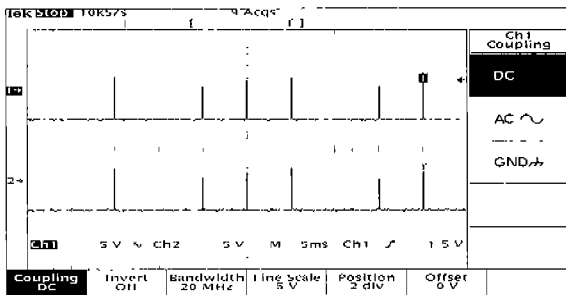
(vowel, /ㅏ/, /ㅓ/, /ㅣ/, /ㅜ/, /ㅡ/, 각 5000샘플씩), 7개의 자음(CV, “바”, “다”, “가”, “카”, “파”, “사”, “타” 각 7000 샘플) 그리고 31세의 여성이 발음한 한국어 11개의 연속음(“내가”, “사는 곳은”, “개울가라”, “지대가 높고”, “겨울이면”, “무척 춥다”, “영하 20도를”, “오르내린다”, “나가는걸 보면”, “일기예보란”, “대관령이”, 각 5000~13000샘플)을 음성 샘플( $f_s = 10kHz$ )로서 사용하였다. 실험 데이터를 중에서 연속음 “무척 춥다”(13,000 샘플, 1.3초) 샘플의 CA/pulsatile 방식의 8채널 출력은 그림 5, 그리고 최종 출력 및 스펙트로그램은 그림 6 (a),(b)와 같다. 그림 5에서 알 수 있듯이, “쿠--” 및 “치--기” 부분은 각각 유성음 및 무성음에 가까운 부분으로서 그 필터 출력 분포의 대부분이 각각 시주파대역 및 고주파대역에 분포하고 있다. 이는 청각계의 place principle이 이 다중 필터 출력신호들에 의하여 만족될 수 있음을 의미한다. 또한, 그림 5(c)에서는 일정 간격의 출력을 보이고 있는데, 이는 피치와 동기화된 출력으로서 periodicity를 만족할 만한 패턴을 출력함을 의미한다. 그림 5(b)의 CA 방식 출력은 다중필터 출력이 음성신호의



(a)



(b)



(c)

그림 6. 연속음 “우척 춤다”의 출력 특성

- (a) 다중 채널 출력 특성 (b) 스펙트로그램
- (c) 채널 1, 2의 일부 구간만을 오실로스코프로 관측한 파형

Fig. 6. Output characteristics of continuous speech sound

/mu:tʃʌk tʃub tʰa:/

- (a) Output characteristic of multi-channel (b) Spectrogram
- (c) Observed waveforms of some part of channel 1 and 2 using oscilloscope

중요정보인 피치 정보를 소실할 우려가 있기 때문에 periodicity를 상대적으로 덜 만족시킬 우려도 있다. 그림 6(b)에서 스펙트로그램의 진한부분은 입력 음성신호의 에너지가 높은 주파수 부분을 나타낸다. 스펙트로그램과 다중 채널 출력(그림 6(a))의 전체적인 스펙트럼 형상이 유사함을 알 수 있다. 이는 시스템의 다중 채널 출력이 입력된 음성신호의 주파수특성, 특히 중요 형성음 성분을 잘 반영하고 있다는 것을 의미한다. 실시간 출력(pulsatile 출력인 경우)을 관측한 결과, 두 채널(첫번째, 두번째 채널)의 최종 출력중 일부 구간만을 오실로스코프로 관측한 파형은 그림 6(c)와 같다. 펄스의 간격이 약 5ms인데, 이 부분에 대응되는 음성신호의 피치는 약 50 sample이기 때문에 피치와 동기화된 출력이 나오고 있다는 것을 알 수 있다. 중간에 펄스가 없는 부분은 신호가 존재하지 않는 부분임을 나타낸다. 일반적으로 어셈블리(assembly) 언어를 사용하면 CPU를 최적으로 제어할 수 있지만, 프로그램 전체의 실행이 실시간이 될 수만 있다면 프로그램 수정이 용이한 C 언어로 작성하는 것이 유리하기 때문에 본 연구에서는 프로그램 언어로서 C 언어를 사용했다. 40MHz clock을 기준으로 할 때, 전체 알고리즘 연산결과는 약 46만 instruction으로서, 1 프레임 전체 용량(25ms, 50만 inst.)의 약 92%를 사용하고 있으므로 저가형의 C31 프로세서를 사용해도 실시간 처리가 가능함을 알 수 있다. 또한, 일반적인 ANSI-C 규격에 근거한 분법으로 대부분의 프로그램을 작성했기 때문에 다른 CPU를 사용하는 시스템에도 쉽게 이식할 수 있다. 이는 곧, 상품화 시에는 프로그램 수정이 대단히 용이하며, 저가격화가 가능함을 의미한다.

결론

본 연구에서는 내이의 손상에 의한 감각성 난청환자들의 청력 회복을 위한 청각보철장치내의 가장 중요한 부분인 어음발체기의 음성신호처리 알고리즘 및 하드웨어를 개발하였다. 실험 데이터인 단모음/자음/연속음 등의 다중 필터 출력, 스펙트로그램과 H/W를 통해 출력된 다중 채널 출력의 패턴들을 정량적으로 분석한 결과 청각계의 tonotopy와 periodicity 원리들을 만족시킬 수 있는 음성신호의 주파수특성을 관측할 수 있었다. 이는, 내이의 기저막 특성을 비선형 다중 필터 뱅크 방식으로 모델링한 후, PSMS등의 전략을 이용한 결과라고 생각된다. 또한 음성신호처리의 대부분은 프로그램으로 동작되므로 환자의 개인적인 특성에 따른 적응(필터뱅크의 대역폭 변경, 채널수의 변경 등)이 아주 쉽고 다양한 실험을 위한 시스템 수정 또한 용이하며 더욱이, 다른 프로세서를 사용하는 H/W에도 쉽게 이식될 수 있다는 장점을 갖는다.

특히, 본 연구에서는 시작품 형태의 H/W 샘플을 제작하여 상품화의 가능성을 확인하였다.

본 연구는 1995년 보건복지부에서 시행한 G7 의료공학 기술개발사업(HMP-95-G-2-31)의 3차년도 연구결과임을 밝힙니다.

### 참 고 문 헌

1. J. G. Webster, *Electronic Device for Rehabilitation*, A Wiley Medical Pub., 1985.
2. B. S. Wilson, C. C. Finley, D. T. Lawson and R. D. Wolford, "Speech Processors for Cochlear Prostheses", *Proceedings of IEEE*, Vol.76, NO.9, pp.1143-1154, Sep. 1988
3. H. J. McDermott, A. E. Vandali, R. J. M. Hoesel, C. M. McKay, J. M. Harrison and L. T. Cohen, "A portable programmable digital sound processor for Cochlea implant research", *IEEE Trans. on Rehabilitation Engineering*, Vol.1 NO.2, pp.94-100, June 1993
4. J. L. Flanagan, *Speech Analysis, Synthesis and Perception*, Springer-Verlag, 1972
5. D. O'Shaughnessy, *Speech communication, Human and Machine*, Addison-Wesley Publishing Company, 1990.
6. G. V. Bekesy, *Experiments in Hearing*, Robert E. Krieger Pub., Huntington, New York, 1980.
7. E. Zwicker, "Subdivision of the audible frequency range into critical bands", *J. Acoust. Soc. Am.*, Vol.33, pp.248-249, 1961
8. E. Zwicker and E. Terhardt, "Analytical Expressions for critical band rate and critical bandwidth as a function of frequency", *J. Acoust. Soc. Am.*, Vol.68, pp.1523-1525, 1980
9. D. K. Eddington, "Speech discrimination in deaf subjects with cochlear implant", *J. Acoust. Soc. Am.*, Vol.68, NO. 3, pp.885-891, 1980
10. N. Tonder, R. Hartmann and R. Klinke, "A Versatile System for the Generation and the Development of Speech Coding Strategies in Cochlear Implants", *IEEE Trans. on Biomedical Engineering*, Vol.45 NO.6, pp.773-782, 1998, 6,
11. G. M. Clark et al., *The university of Melbourne-Neucleus multi-electrode cochlear implant*, Karger, 1987
12. M. M. Merzenich et al., *The UCSF cochlear implant project : State of development*, *Advances in Audiology*, Karger, 1984
13. Teax Instruments, *TMS320C3x User's guide*, 1994
14. Teax Instruments, *TMS320 Floating-Point DSP Optimizing C Compiler*, 1994