

음성 인식을 위한 편집시스템의 구성

* ** *** **** *
 송도선 이천우 신천우 정중수 이행세
 (*아주 대학교 **대전실업전문대학, ***영남 대학교, ****ETRI)

Construction of Speech Editing System
 for Speech Recognition.

* ** *** **** *
 (D. S. SONG, C. W. LEE, C. W. SHIN, J. S. JEONG, H. S. LEE)
 (*AJU UNIV., **DAE JEON JUNIOR COLLEGE, ***YEUNG NAM UNIV., ****ETRI)

요 약 (Abstract)

In the study for effective speech control we designed a personal computer system with A/D converter in which the speech signal is transformed by digital data displayed graphically on the monitor and with a D/A converter in which the digital data is transformed into speech signal which people can hear.

We analyzed the character of the speech signal produced by the system.

We designed the adaptive noise cancel algorithm so that noise and interference are cancelled whenever the speech signal is recognized by the computer system.

This is a basic system for artificial intelligence.

1. 서 론

컴퓨터로 음성을 처리할 때에 음성신호를 data 화하여 컴퓨터 메모리에 저장하거나 전송을 하기 위해서는 효율적인 data 관리가 필요하며 음성의 인식 등에 관한 연구를 할 때에는 음성에 섞인 잡음이나 간섭신호등을 제거하여 순수한 음성신호만을 저장하는 기법이 필요하게 된다.

본 연구에서는 음성신호를 personal computer로 받아 들이는 A/D converter 합성한 음성 data를 다시 음성으로 확인 할 수 있는 D/A converter를 제작하여 personal computer에 연결하였으며 음성신호중 잡음이나 간섭신호를 제거 할 수 있는 Adaptive Noise Cancel Algorithm을 설계하여 음성 data가 잡음이나 불필요한 간섭신호가 있을 때에는 그 간섭신호를 제거하고 필요한 음성 data만을 저장할 수 있게 하였다. 이러한 간섭제거 Algorithm program 수행시 personal

computer에서는 시간이 많이 걸리므로 음성 data를 mincomputer(vax-11/750)로 전송시켜 mincomputer에서 음성처리를 수행시켜 시간적인 효율이 증가되도록 하였다.

2. 시스템의 구성

본 시스템을 구성하는데에는 IBM-PC/XT를 기준으로 설계, 제작하였다. A/D converter의 sampling rate는 20KHZ이며 1KHZ - 20KHZ까지 사용할 수 있게 하였다. 그리고 Resolution은 12bit이며 D/A converter의 sampling speed도 가변할 수 있게 하였다.

마이크에서 얻어진 음성 신호는 level이 낮으므로 OP Amp를 통하여 증폭 시킨후 고조파가 발생시키는 error를 줄이기 위하여 Buffer earth 4단으로 Low Pass Filter를 구성하였으며 이 filter는 음성 대역폭만 통과 시키고 4KHZ에서 20DB 5KHZ에서 40DB의 감쇠 특성을 가진다.

이 때 Analog 신호는 Sample holder 에서 hold 되어 A/D converter 에서 Analog 값이 Discrete value 로 바뀌어 controller 에 의해 1Byte 씩 computer 의 Data Bus 로 들어가게 된다.

컴퓨터에서는 각 Data 들을 받아서 file 형태로 만들어 H/D 에 save 시킨다.
 이상에서 음성 Data 를 file 로 만들어 Hard disk 에 save 시켜 여러가지 처리를 할 수 있게 하였다.

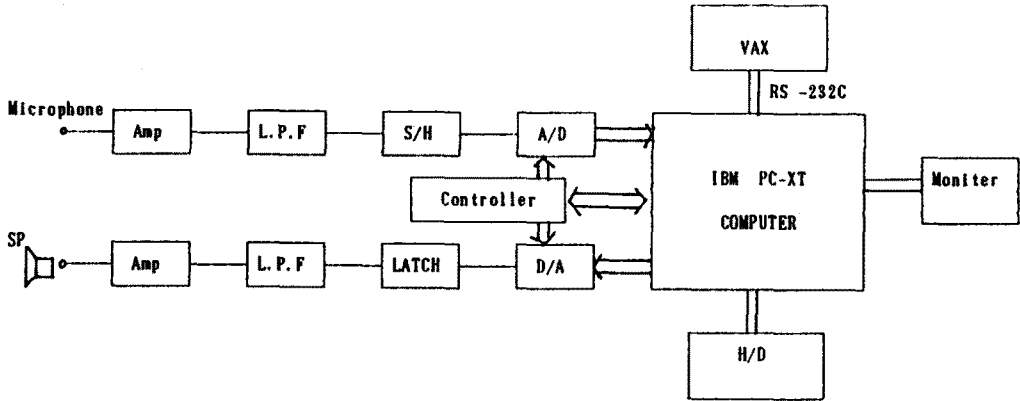


Fig.1(System Configuration)

3. 적응 필터

얻어진 음성 Data중 외부의 간섭신호나 잡음을 제거하기 위하여 Adaptive filter를 구성하였다. 원래의 신호를 변형시키지 않고 신호성분중 잡음을 제거하는 적응필터는 신호(primary input)와 Noise나 간섭신호(reference input)사이의 통계적

성질을 이용하여 필터의 전달 함수를 시간에 따라 변화시켜 최적의 output을 얻어낸다.

필터의 구성은 아래의 구성과 같으며 system의 출력 error 를 feed back시켜 전달 함수를 변형시켜 원하는 출력 (desired signal)을 얻는다.

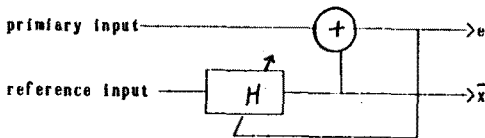


Fig.2 Adaptive Filter

얻어진 신호는 $R = S + X$

S_k -----desired signal

X_k -----noise signal

이때 원하는 신호와 얻어진 신호와의 차이인 error는

$$E_k = R_k - D_k$$

$$E_k = D_k - X_k^T W = D_k - W^T X_k$$

$$W = [W_0 \ W_1 \ \dots \ W_{M-1}]^T$$

$$X_k = [X(k) \ X(k-1) \ \dots \ X(k-M+1)]^T$$

$$e(k) = Y(k) - H^T X(k)$$

이때 평균계급 오차는

$$E \text{ 는 } E[e^2(k)] = E[Y^2(k)] - 2P_{YX}^T H + H^T R_{XX} H$$

$$\nabla_H E = -2R_{XX} H + 2P_{YX} = 0 \text{ 로 두면}$$

$$\therefore W^* = R_{XX}^{-1} P_{YX} \text{ 가 된다.}$$

4. 편집 알고리즘

이상에서 얻어진 음성 데이터 file 들을 Graphic 하여 눈으로 음성 Data를 확인할 수 있어야 하므로 Hercules Board 상에서 700x400 resolution을 가지는 Window를 만들어 그 Window 상에 음성 data 를 Graphic 할 수 있게 하였다.

한 Window 에는 각 data 를 700 Byte 씩 나타낼 수 있으나 그렇게하면 눈으로 확인할 수 있는 Data 양이 적어지므로 10개씩 평균해서 1개의 dot로 나

타내는 Program 을 만들어 7K, 14K 등으로 묶어서 Graphic 할 수 있게 하였다.

이러한 file 들은 대량의 data 를 가지므로 이 data 에서 음성의 결점을 detect 편집을 하여 D/A converter 를 통해서 귀로서 들을 수 있게 하였다. 이때에 편집된 file 의 음성 data 는 data BUS 를 통하여 D/A converter 를 거쳐 Analog value 로 바뀐다.

그 다음 Low Pass filter 를 통과하여 Sampling Noise 를 제거한 다음 Amplifier 를 통하여 Speaker 를 구동시킨다.

전화통화를 할 때의 음성이나 진찰역등에서 발생한 음성 data 를 file로 만들었을 때에는 음성 신호에 잡음이나 간섭등 음성의 신호가 많이 포함되어 있으므로 이러한 간섭 신호는 음성의 편집이나 인식 장치에 막대한 error 를 초래하게 된다. 그래서 이 때에는 Adaptive Interference Cancel Algorithm 을 적용시켜 간섭신호를 제거한 후 음성 file 을 save 시킨다.

이러한 간섭제거 Program은 microcomputer

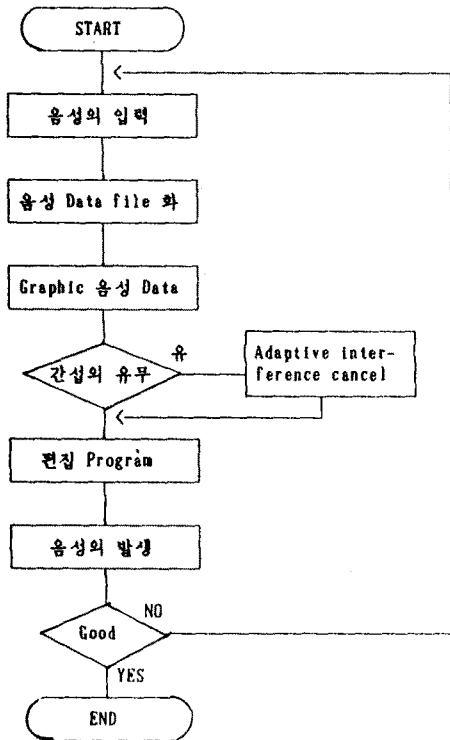


Fig.3 (System Flow Chart)

상에서 수행할 시에는 시간적인 지연이 많이 생기브 이 때에는 Datafile 을 minicomputer(vax-11/750)으로 전송하여 minicomputer(vax-11/750)에서 작업을 한 후 그 결과를 다시 microcomputer 로 전송시켜 personalcomputer 에서 결과를 분석및 측정 하였다.

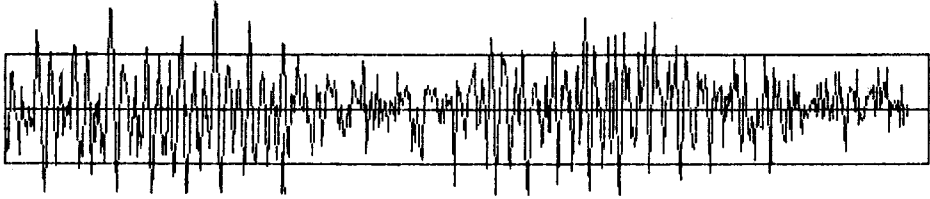
5. 결론

이상의 system 을 이용하여 음성을 분석하거나 편집시 편리하게 사용할 수 있었으며 특히 간섭이나 잡음이 많은 음성의 분석시에 아주 편리하게 사용할 수 있었다.

minicomputer(vax-11/750)에서 Adaptive Interference Cancel algorithm이 수행될때 약간 시간이 소요되었으며 이 Interference Cancel 이 H/W로 제작되어 사용할 경우는 Real time processing 이 가능하리라고 생각 된다.

Reference

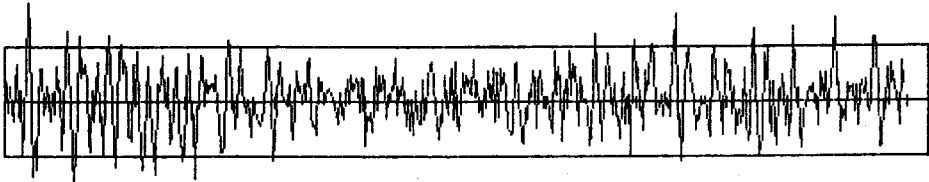
1. 신천우, 김태호, 이무영 "고속 Data Acquisition System의 설계와 제작" 한국통신학회 추계발표회 논문발표집, pp.8-11, 1986, Vol.5, NO.1.
2. Robert Lafore, "Assembly Language Primer for the IBM PC & XT", The waite Group
3. Christopher L. Morgan, "Blue Book of Assembly Routines for the IBM PC & XT", The waite Group.
4. IBM Technical Reference Manual 1984.
5. Alan V. Oppenheim & Ronald W. S. Signal Processing." Prentice-Hall Inc., 1975.



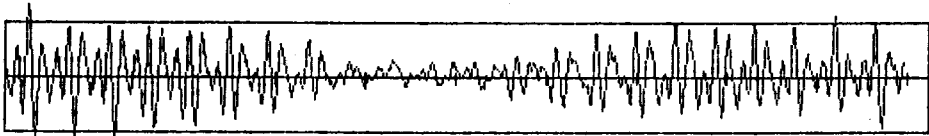
<잡음이 섞인 음성 Data>



<Adaptive filter 출력>



<잡음이 섞인 음성 Data>



<Adaptive filter 출력>
