論文98-35T-12-2

長母音 認識裝置 設計 製作

(Design and Manufacture of a Device for the Recognition of Long Vowels)

具 龍 會 *

(Yong Hea Ku)

요 약

장 모음 음성인식을 전자회로로 수행 하였다. 레벨 압축은 음성 파형을 직렬 펄스로 변화 시킬수 있었다 이 펄스들로 모음을 구별하는 정보가 된다. 펄스의 샘플링은 한다위로 모음의 피치 직렬신호를 얻어지는 레지스터에 의해서 이루어진다. 샘플링 펄스에 의한 시간제어 펄스 는 음성파형의 첨두치 펄스에 의해 발진 되어 진다. 이 레지스터에 있는 병렬 데이터는 만약 ○○이면 ○○이다 는 규칙으로 이루어지는 의지결정 회로의 뜻에 따라 음성 심볼이 인식되어진다.

Abstract

The speech recognition on long vowels are carried out by electric circuits. A level compressor is able to transform the wave of voice to serial pulses. The obtained pulses have informations to distinguish the vowels. The sampling of the pulses is carried out by the register which picks up a series of serial signals in a pitch of a vowel as an unit. The timing control pulses such as sampling pulses are generated by using peak pulses in the speech wave. The parallel data in the register assign the phonetic symbol by means of the decision making circuit which carries out the IF-THEN rule.

Ⅰ. 서 론

컴퓨터를 이용하여 인간의 知的기능을 기계에 대체하기 위하여 노력 하여 온 결과 현재와 같이 인간이 행동하여 작업하는 많은부분을 컴퓨터가 개입하여 대행하고 있다. 기계 언어를 이용하여 인식을 시행해온 결과 한정된 음성 인식이 실용화되었다. 그러나, 人間과 컴퓨터간에 거리는 존재한다. 즉, 기계와 人間간의 커뮤니케이션은 인간과 같은 수준으로 도달하기는 어렵다.

人間의 知性이라고 하는 기능은 단순히 현재의 컴퓨터의 기억용량이나 연산속도 를 고성능으로 하여도 달성하기에는 어렵다.

인간은 체험에 따라 습득된 많은 양을 조건 반사 유지하기 위한 시스템을 가지고 있다. 이것은 인간의 大腦에 구성을 전자 회로와 같이 만들어졌다고 생각하여 본다면 知性의 구성을 이해 할 수 있다고 본다.

인간의 大腦와 같은 전자회로 시스템을 만들 수 있다면 기계가 같은 지능을 인공지능으로 생각해 볼 수 있다. 본 연구는 이 전자회로 시스템을 다음과 같은 전자회로에 기초를 둔다.

「만약 ○○이면 ○○이다.」 같은 조건 반사에 대한 동작에서 의지 결정 회로를 병렬 내지 직렬 접속한다. 이 의지 결정 회로는 항상 동작 상태에 있다. 실시간으로 차차 변화하는 사항에 대응하여 의지 결정 정보처리를 할 수 있다.

여기에서 의지 결정 회로는 음성을 인식하는 여러가지 규칙을 조합하는 장치를 제작한다. 이 전자 장치

* 正會員, 仁荷工專 電子科

(Dept. of Elec. Inha Junior Colleage)

接受日字: 1998年10月13日, 수정완료일: 1998年12月11日

는 이용자가 필요로 하는 지능만을 익혀 사용하게 된다.

본 논문은 음성 신호의 파형으로부터 피치주기에 따라 음성 인식의 특징을 전자회로로 검출하여 모음을 분리하는 전자회로를 제작한다.

Ⅱ. System 구성

장모음은 성인 남성은 6~11 [msec] 여성은 3~4 [msec] 정도의 주기가 같은 파형이 반복된다. 여기서 같은 모음의 경우에서 고음과 저음은 피치 주기가 다르고 모음의 종류에 따라 피치내의 파형이 다르다. 이 피치 내의 파형을 조합하여 모음을 분리한다. 피치 주기에 대하여 음성 신호를 조합한다. 이 시스템은 회로 동작을 제어하는 타이밍 펄스회로와 음성 데이터의 처리회로로 대별한다. 음성 신호의 레벨 변동을 제거하기 위해서 진폭의 대수 압축을 한다. 이 진폭 압축은 음성의 인식에서는 정보가 손실되지만 회화의 전반적인 이해성은 유지된다.

인식 회로의 설계는 「만약 ○○ 이면 ○○이다.」 즉 IF THEN Rule에서 입력 신호와 기준 데이터와 일치한다고 판단시 회로가 음성을 인식하는 회로가 구성된다. 구체적으로 음성 신호를 압축하여 얻은 펄스열을 직렬/병렬 변환하여 레지스터에 저장시키고 해석기에 병렬 입력시킨다. 해석기는 조합논리의 해석기를 확장한 논리 회로이다. 우선 Digital 신호를 H(1)와 L(-1)의 2진수로 변환하여 비반전 신호와 반전 신호가 출력되면 양자간에 하나가 기준 상태의 신호 H(1)상태인선을 가산회로에 접속한다. 이 가산회로는 기준 데이터와 입력 데이터가 일치하면 가산하고 불일치하면 감산한다. 이렇게 해서 가산 합산된 값으로 분리한다.

그림 1은 제작된 회로의 Block Diagram이다.

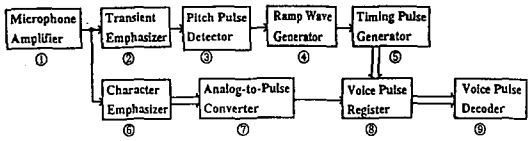


그림 1. 장모음 인식 장치의 Block Diagram
Fig. 1. Block Diagram for the recognition of long vowels.

Ⅲ. 제어용 Timing Pulse 회로

1. Peak 후속 신호의 압축회로

음성 신호는 진폭의 변화가 크다 이 진폭 변동이 급격히 변화하는 신호이기 때문에 음성의 피치를 확실하게 검출하는데 어려움이 있다. 먼저 음성 진폭을 일정하게 하는 목적으로 그림 2에 나타낸 것처럼 진폭 Level 압축기를 제작하였다.

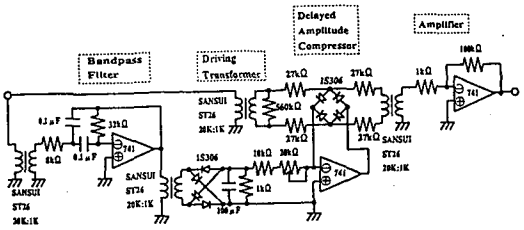


그림 2. Peak 후속 신호 압축회로
Fig. 2. Peak follow signal compressor circuit.

하단의 진폭 정보로 상단의 신호 회로의 이득을 제어하는 신호를 검출한다. 이 회로는 제어 신호가 상승될 때까지 지연되는 것을 이용하여 후속 신호의 Level을 압축하여 선두치 Peak 신호가 강조된다.

다음은 그림 2의 회로의 동작을 설명한다.

入力의 음성파형은 교류이나, 양과 음이 대칭이 아니므로 능동소자가 포함된 전자회로를 통과하면 Level이 변동된다. 이 Level이 변동되는 것을 억제시키기 위하여 트랜스를 사용하여 신호를 전달한다.

그림 2의 하단은 진폭 정보 검출을 나타내고 있다. 중심 주파수 (f_0=100Hz), 첨예도(Q=1), 중심의 이득(H=2)의 BPF에 남성의 pitch 주파수대는 강조하여 증폭하였고 전파정류에서 진폭 정보를 얻는다. 여기서 τ =100μs 시정수를 유지하여 평활회로를 통과한 진폭 정보에서 제어 전류를 얻는다.

음성 신호의 이득 제어 회로는 다이오드 브릿지에 음성진폭에 비례하는 제어 전류가 흐른다. 제어 전류에 반비례하는 동저항에 의해서 진폭을 압축한다.

2. Pitch pulse 검출회로

Peak Hold 회로에서 방전하는 시정수 회로를 접속하고 모음 Pitch를 반복하여 단시간에 충전하는 파형을 얻어 Pitch를 검출한다. 이 Peak Hold 회로를 이용하여 pitch pulse 검출하는 회로를 그림3에 나타냈다.

첫단의 IC OPAmp는 100μF콘덴서 단자 전압이 입력전압과 같을 때까지 충전된다. 입력전압이 Hold 전압보다 낮을 때 다이오드가 도통 되어 폐 Loop 동작

이 된다. 여기서 1MΩ저항은 리-그 전류를 보충하는 역할을 유지하게 된다.

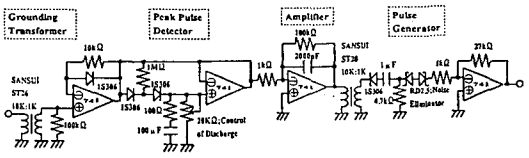


그림 3. Pitch pulse 검출회로
Fig. 3. Pitch pulse detection circuit.

다음은 Peak치 주변 Pulse 1개를 적분회로에 통과시키면 적분회로에서 약간의 Time이 변동되는 문제가 생긴다. 정극성 impulse를 만들기 위하여 미분회로의 앞에 다이오드를 삽입한다. 또한 낮은 Level을 제거하기 위하여 제너 다이오드를 직렬 삽입한다.

3. Ramp 파형 발생 회로

음성 파형을 합성하는 경우에는 직.병렬 변환 Register를 사용한다. 이 레지스터는 Flip-Flop 회로를 병렬 사용하며 각각의 Flip-Flop의 Clock 입력은 Pitch기간을 시간적으로 분할하여 병렬 Pulse가 필요하게 된다. Clock pulse가 Pitch 주기에 반복으로 분배하게 되면 이때의 data는 레지스터에 유지한다. 이 레지스터로부터 병렬음성 Data가 조합회로에 전송되어진다. Pitch 기간을 정확하게 분할하여 Pitch 기간 동안 Ramp 파형이 발생한다.

이 Ramp 파형의 최대치를 유지시, 이 전압을 분할하고, 비교회로의 기준 전압을 원래의 Ramp파형과 비교하면 시간 분할의 시간 Pulse를 얻게된다. 이 시간 Pulse 발생등 Ramp 파형의 발생회로를 그림 4에 나타내고 있다.

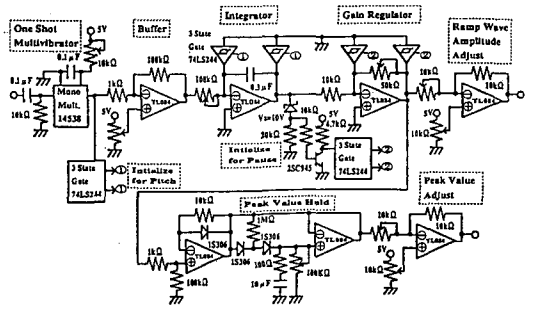


그림 4. Ramp 파형 발생 회로
Fig. 4. Ramp waveform generation circuit.

그림 4의 회로는 Pitch pulse를 트리거로 이용하여 단안정 멀티 바이브레이터를 구동시 진폭Level이 일정한 동기신호를 얻는다. 이 단안정 멀티 바이브레이터는 Digital IC로써 전원전압은 5V를 사용한다. 다음 단계의 적분을 전원 전압 ±15V의 Analog IC를 사용하여 동작시키면 중간 Level을 합하여 완충 증폭기를 설계한다.

Ramp파형은 동기신호를 IC OP Amp를 사용하여 적분기에 적분하여 얻는다. 적분 초기치의 설정은 3 state gate ①을 이용하여 동기신호가 High때 콘덴서 양단에 접속된다. 3 state gate는 개방시 전압은 0 V~7V 영역에서 정상 동작하게 신호 Level을 조정할 필요가 있다. 즉 입력신호가 없을 경우에도 적분을 계속하여 출력치가 포화되어 최대치가 된다. 이것은 적분치가 초월하는 경우 출력은 "0"으로 되고 3 state gate ②회로에 가해진다. 최후 파형의 출력단 Ramp 파형의 Level이 최종조정 되어진다.

비교 회로의 기준전압으로 제거하여 필요한 Ramp 파형의 최대치를 유지하는 회로가 그림 4의 하단에 나타나 있다. 유지되는 Peak치의 출력단의 Level을 최종 조정한다.

4. Sampling pulse 발생회로

직/병렬 변환 레지스터에 데이터를 입력 하면 입력 Timing을 지정하기 위해서 Sampling pulse가 필요하게 된다. 그림 5는 Sampling pulse 발생회로를 나타내고 있다.

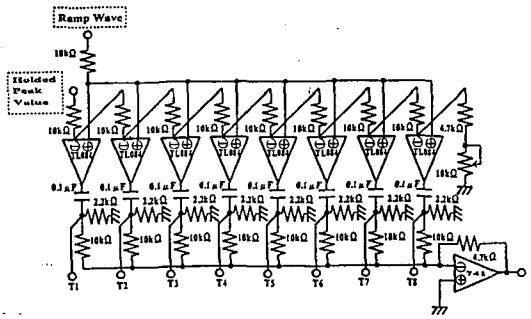


그림 5. Sampling pulse 發生回路
Fig. 5. Sampling pulse generation circuit.

전절 3.4 Ramp 발생회로에서 얻은 Ramp 파형을 각각을 비교회로에 입력하면 Ramp 파형의 최대치를 저항 분압으로 구분하여 각각의 기준전압으로 비교하고 Pitch 반복하여 Sampling pulse를 발생하게 된다. 각 Sampling pulse는 OP IC 741을 이용하여 가

산된다. 이 Pulse 열을 Monitor하게 되는 것이다.

Ⅳ. 음성 회로 처리 회로

1. 마이크로폰 증폭기

그림 6에 나타낸 마이크로폰 증폭기에서 콘덴서 마이크로폰으로 나온 신호를 Impedance가 높은 회로에 가한다. 이 신호는 직류 성분이 함유되어 있는데 전압 Shift 회로를 통과하여 증폭한다. 여기서 Trans의 직류 Level을 접지 하면 적은 Level의 잡음을 Diode에서 제거 증폭한다. 최종의 출력을 조정하는 회로가 설계되어 있다.

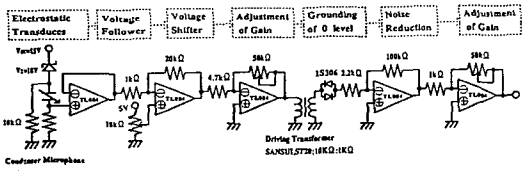


그림 6. 마이크로폰 증폭기
Fig. 6. Microphone amplifire.

2. 주파수 특성 조정회로

음성신호의 진폭을 단순히 대수 압축한다. Pulse 신호에서 낮은 주파수 성분은 큰 영향이 있고「이」와「우」의 구별이 안된다. 여기서 음성의 파형을 Pulse로 변환하게 되면 주파수 성분이 증가된다.

모음의 주파수 특성의 특징이 불분명하게 된다. 그러므로 모음별에다가 성도(聲道)의 형태에 따라서 주파수 특징을 강조하여 모음별 Pulse 신호를 얻어 음성인식 Date로 Sampling 한다.

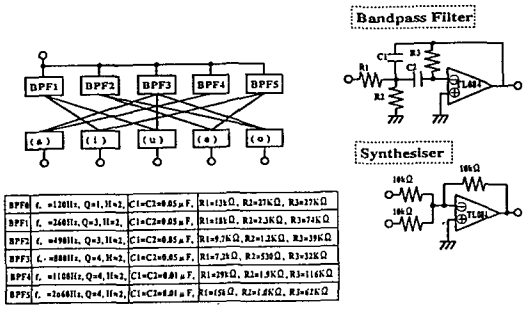


그림 7. 주파수 특성 조정 회로
Fig. 7. Frequency characteristics tuning circuit.

그림 7은 주파수 특성 조정 회로를 나타내고 있다. 이 회로망은 「아」「이」 우」「에」「오」등 각 모음에 대해 제 1 Format 내지 제 2 Format의 주파수 성분[6]을 BPF에 강조하는 회로이다. IC OPAmp를 사용하여 BPF와 가산기로 구성되어 있다.

3. 음성 신호를 대수 압축한 Pulse 신호화 회로

Diode의 각단자에 흐르는 전류량은 대수적인 관계를 이용하여 진폭을 압축한다. 이 대수 압축한 Pulse 신호 변환회로를 그림 8에 나타 냈다. 여기서, 대수 압축하면 미소 잡음이 크게 증폭되어지는데 이것은 Diode로 제거한다. 초단의 출력을 근사적으로 약 0.55V 정측에 Shift하고 그후에 정류하여 5V로 Digital IC 레지스터에 가한다. 이 Pulse 신호화 회로는 5종류의 각각 모음에 대하여 개별적으로 제작된다.

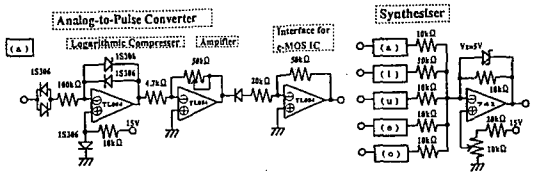


그림 8. 음성 신호를 대수 압축한 Pulse 신호화회로
Fig. 8. Logarithmic compressed pulse signalize circuit for the voice signal.

초단의 레지스터는 각 모음에 대하여 Sampling 수(8개) 만큼 Flip-Flop 접속한다. 여기서 F-F회로의 수가 40개가 8개로 감소하게 된다. 각 모음의 Pulse 신호를 가산하여 합성 Pulse 신호화를 Sampling 한다.

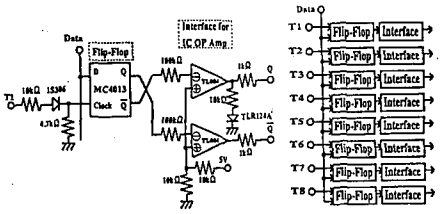


그림 9. 음성 Pulse 레지스터
Fig. 9. Voice pulse register.

4. 음성 Pulse 레지스터

Sampling Pulse 발생 회로에서 얻은 8종류의 Timing pulse를 Clock 신호로 사용한다. 직렬 음성 Pulse 신호를 병렬 레지스터의 Data에 가한다. 이 Digital 레지스터의 출력은 비반전 신호선과 반전 신

호선에 분배하여 IC OPAmp인 Interface에 접속하여 다음의 해석기에 보내진다.

출력 상태는 발광 다이오드로 확인 할 수 있다. 병렬 배열된 F-F에 Sampling pulse에 대한 음성 Pulse를 분배하는 회로를 그림 9에 나타내고 있다.

5. 음성 Pulse 해석기

음성 Pulse 해석기는 그림 10에 나타냈다. 이 해석기는 꾸며진 논리의 해석기를 확장하여 비 반전과 반전 신호의 양자를 참조하여 신호가 H(1) 상태가 되면 이 선은 가산 회로에 접속시킨다. 본 시작 장치에서 예를 들면「i」를 발성할 때 발광 다이오드가 불이 들어오는 선은 DIP SW를 사용하여 가산기에 접속한다. 이때에 참조 Data와 같은 부분에 같은 입력이 있으면 이 둘은 가산한다. 불일치 부분은 감산한다. 이렇게 해서 얻은 합계의 가산치를 비교회로에서 판정한다.

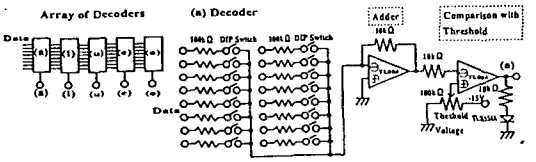


그림 10. 음성 Pulse 해석기
Fig. 10. Voice pulse analyzer.

V. 결 론

본 논문은 장 모음 인식 장치를 9개의 Block 회로로 나누어 각 Block 회로에 대하여 회로 특성 측정과 개선을 반복하여 제작하였다 특히 Pitch 신호의 확실한 검출과 Sampling pulse의 發生 회로 방법을 신중히 검토하였다.

인식의 종합 특성에서 「i」와「u」를 명확히 식별하기 위해서 조절하므로 「e」나「o」의 인식이 불안정 하게되므로 개선의 여지가 남는다.

「a」는 주파수 성분이 높으므로 Sampling pulse의 수를 증가할 필요가 있다. 그러나, 본 장치를 조정하면 「i」「u」「e」「o」의 장모음의 인식이 가능하다는 것을 알게 되었다. 또한 Ramp 파형 최대치를 유지하는 회로는 응답이 지연되므로 3 State gate에서 Pitch를 취하는 방식을 개량할 필요가 있다.

앞으로 더욱 연구하여 전후 관계를 인식판단하여 정보를 취할 필요가 있다. 여기서 출력은 2종류 값을 설계하여 Clearly Yes [1], Clearly No [-1] 내지, Not Clearly [0] 같이 3 값을 나오게 하면 System에서 정보처리가 적절하고 명확하게 처리 될 수 있다고 생각 된다.

참 고 문 헌

[1] 中川聖一, 鹿野清宏 ,東倉洋一, (コュ-ロサイエンス&セクノロジ-シリ-ズ) [音聲・聽覺 と 神經回路網モテル] オ-ム 社 (1990)

[2] 國際電氣通信基礎技術研究所篇, ATR 先端テクノロジ-シリ-ズ, 自動飜譯電話 オ-ム 사, (1994)

[3] S. Karasawa, Y. Ku, "Methods and Circuits for a Real-Time Automatic Translation from Voice into Phonograms", Proc. of Inter. Tech. Conf. on Circuits/ Systems, Computers and Communicatons, pp. 1130-1133, Seoul Korea, July. 15-17, (1994).

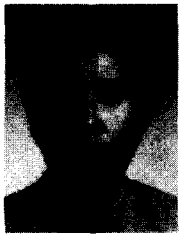
[4] 唐澤信司, 具龍會, 関泰三弘, [振幅壓縮による音聲 の デジタル 信號化] 宮城工業高等專門學校研究紀要 vol. 32, pp. 1-3, (1996)

[5] トランジスタ 技術編集部 , [實用電子回路 ハンドラック] pp. 236, xza CQ 出版社, (1978)

[6] P. Lieberman, "Speech Physiology and Acoustic Phoneteics" pp. 175, Macmillan Publishing Co., Inc., (1975).

[7] 丹野賴元, [演習オペアンプ回路] 森出版社, pp. 148, (1982)

저 자 소 개



具 龍 會(正會員)

1941년 1월 4일생. 1968년 2월 광운대학교 응용전자공학과 졸업(공학사). 1973년 2월 연세대학교(석사). 1989년 2월 광운대학교(공박사). 현재 인하공업전문대학 전자과 교수