

범용 DSP 를 이용한 LPC 방식 실시간 음성 합성기 설계에 관한 연구

*김 홍선, **이 외택, *안 수길
*서울대학교, **한국전기통신연구소

A Study on the Design of the real-time speech synthesizer with the LPC method using Digital Signal Processor.

*Hongsun Kim, **Eetaek Lee, *Souguil ANN
* Seoul National University, **KETRI

ABSTRACT

In this paper, the implementation of the real time LPC synthesizer using NEC 77P20, the DSP (Digital Signal Processor) chip which facilitates and simplifies the digital hardware, is considered. This method shows the good quality with the low bit rate below 9.6kbps and has the advantage of the flexibility and the simplicity.

1. 서 론

LPC (Linear Predictive Coding) 음성 coding 방식은 인간의 발성구조에 근거한 source coding 방식으로 음성의 low bit rate 전송이나 음성 데이터 압축 저장등에 사용된다. 이는 최근들어 음성과 데이터의 혼용망 시스템에서 점차 그 필요성이 높아가는 9.6 kbps 이하의 Vocoder System 과 통신망의 새로운 서비스, 공공기관의 각종 Announcement, 가전제품등에의 활용이 기대되고 있다.

LPC 방식의 구성시 야기되는 문제점은 좋은 음질을 유지 하려할 경우 그 구조가 복잡하고, 일반적인 Digital Hardware 구성시 고속의 고가인 특수 Device 를 사용하여야만 Hardware 구현이 가능하다는 것이다.

일반적으로 LPC 방식의 Vocoder System 과 음성 응답 장치가 유용한 것이 되기 위하여는 다음 조건이 만족되어야 한다. 즉 소형이며, 저가격, 저전력 소비이고, 합성음질이 좋아야 한다는 것이다. 그런데 1982년부터 시판되기 시작한 DSP (Digital Signal Processor)를 사용하게 되면서 계산량이 많은 신호처리 시스템들이 Hardware 는 아주 간단하고 Software 가 주요한 부분이 되는 Hardware subroutine 형태로서 구현이 가능하게 되었다. [6]

본 연구에서는 DSP 의 하나인 NEC 77P20을 사용하여 설계한 LPC 방식 음성합성기의 구현에 관하여 다루었다. 이는 앞으로 소형 Vocoder system 과 음성 응답 장치에 활용될 것으로 기대된다.

2. LPC 방식 음성 합성 알고리즘 [1]

1) 음성 합성 알고리즘의 개요

LPC 방식 음성 합성은 LPC 분석 Program 에 의해 얻은 반사 계수와 energy, pitch 를 이용하여 frame 단위로 합성해 낸다. Driving function 은 유성음일 경우 Impulse Train 을, 무성음일 경우 Random Number Generator 를 이용한다. 본 시뮬레이션에서는 연속된 두 frame의 분석 데이터를 이용하여, Pitch Synchronous Interpolation 을 행하여, parameter 의 급격한 변화를 보상 하였다. (그림 2-1)

2) 합성필터 modelling

그림 2-2와 같이 구성되며, 각 section 의 모델은 여러 방식이 있으나 대표적인 모델로서 본 연구에서 사용한 Two-multiplier Lattice Filter 의 구조는 그림 2-3과 같다.

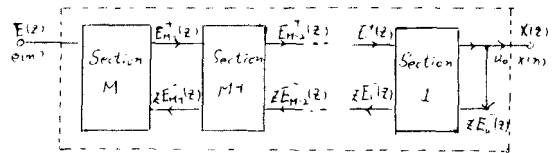
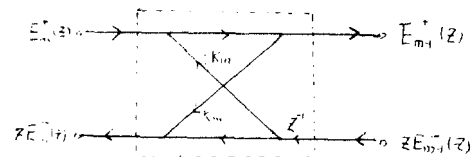


그림 2-2

그림 2-2. 음성 합성구조



$$E_{m+1}^+(z) = E_m^+(z) - K_m E_m^-(z)$$

$$zE_{m+1}^-(z) = K_m^{-1} E_m^+(z) + zE_m^-(z)$$

for m = M, M+1, ..., L

그림 2-3. Two-multiplier Lattice Filter 구조

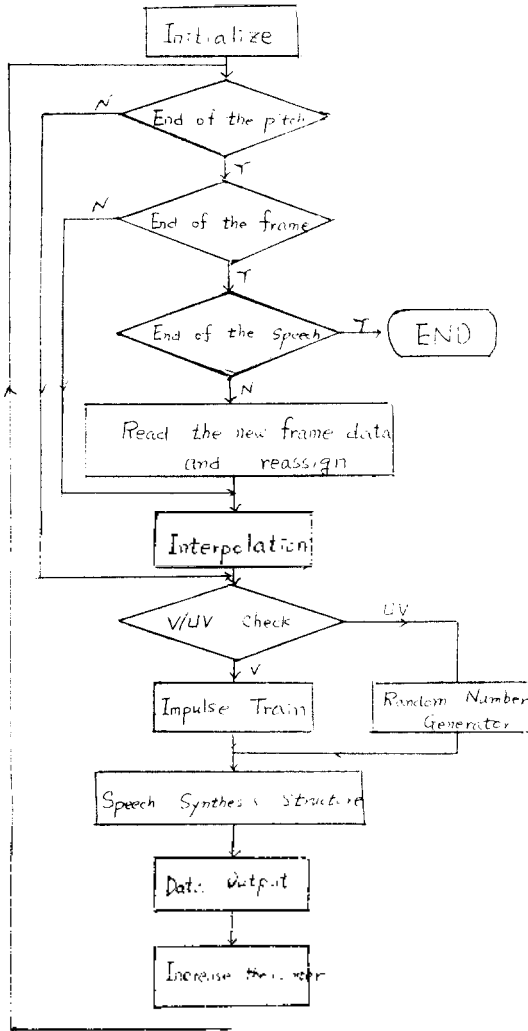


그림 2-1. 합성 알고리즘 유통도

3. 합성기의 구성

1) Hardware 구성

Hardware 의 개요는 그림 3-1 과 같다.

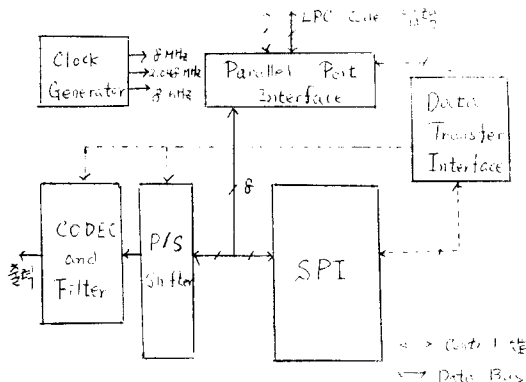


그림 3-1. 합성기 하드웨어 구성도

SPI 는 125 μ s 단위로 CODEC 에 합성된 데이터를 보내 주어야 하므로, 이 동기를 맞춰 주기 위하여 SPI 의 Interrupt 기능을 이용하였다. Frame 단위의 source parameter 전송은 micro-computer 의 parallel port 를 이용하여 SPI 의 non-DMA mode 로 구성하였다.

i) Clock Generator

CODEC 과 SPI 를 위해 2.048MHz와 8KHz 의 clock과 125 μ s 단위의 동기 신호를 발생한다.

ii) Data Transfer Interface

SPI 로의 데이터 입력과 출력을 제어하는 부분으로서, 데이터 입력시에는 SPI 의 P₀, P₁ port 로부터 입력 제어 신호를 수신하여, PIO 와 SPI 에서의 data 전송에 필요한 handshaking logic 을 만들어 내고, 데이터 출력시에는 매 125 μ s 단위의 interrupt pulse 를 SPI 에 주어 SPI 가 DMA mode 로 데이터를 출력하여 P/S shifter 에 latch 시키도록 한다.

iii) CODEC

TSx Signal 에 동기되어 SPI 로부터 출력된 8 bit pulaw data 가 2.048Mbps PCM bit-stream 에 실리도록 한다.

2) NEC 77P20 Software

그림 2-1의 Main Routine 을 고정숫자점 16bit 연산으로 구성하였으며, 연산 시간을 줄이기 위하여 data ROM 을 적절히 이용하였다.

계산 시간을 감소시키며 메모리를 효율적으로 사용하기 위해 다음과 같이 메모리를 할당해서 사용하였다.

* 기본 주파수와 frame 의 길이를 check 하기위한 data counter 를 00H-02H 에 할당한다.

* 새로 입력된 source parameter(RC,Energy, Pitch) 및 V/UV Flag 를 다음과 같이 할당한다.

New Frame Data	: 10H - 1AH
Left Frame Data	: 20H - 2AH
Right Frame Data	: 50H - 5AH
Interpolated Frame Data	: 40H - 49H
RC Buffer for the lattice filter	: 70H - 78H

4. 실험 결과 및 고찰

본 실험에서는 77P20 구조에 맞추어서 16bit 고정 숫자점 연산의 non-real time simulation 을 NOVA 4/x Mini-computer 를 사용하여 수행하였으며 (표 1), 그 결과는 그림 4-1과 같다.

항 목	명 세
Sampling Rate	8 KHz
Frame length	90 Samples(256 Samples)
Overlap window length	15 Samples(64 Samples)
필터 order	6 차 (10차)

표 1. 합성기 simulation 의 주요사항

합성기의 77P20 S/W 는 CP/M Machine 상에서 Cross-assembler 와 Simulator software 를 이용하여 개발하였으며, Hardware Evaluation 은 EVAKIT 를 이용하고 있다.

현재 실시간 음성 합성기를 구현하기 위해서 사용된 77P20 내의 memory 와 처리 시간은 다음과 같다. (표 2)

MEMORY	구 분		사용량 및 비율
		Instruction	ROM
	Data	ROM	114 (/512), 22.3%
	Data	RAM	90 (/128), 70.3%
처리시간	구 분		소요 시간
		Main Routine	30 - 110 μ s
		Interrupt Routine	6.5 μ s - 165 μ s
		RAND	10 μ s

표 2. 77P20 Memory 사용과 처리시간

합성된 음질은 상당히 'good quality' 를 갖고 있었으나, 앞으로 좀더 명료성과 intelligibility 를 높여가는 실험과 함께, 현재(약 8kbps)보다 bit rate 를 더 낮추는 실험을 진행중에 있다.

5. 결 론

본 연구에서는 기존의 DSP 인 NEC 77P20을 이용하여 LPC 방식 real-time 음성 합성기를 구성하였다. SPI 를 이용함으로써, 많은 Hardware 부분이 Software 적으로 처리 되기 때문에 flexibility 가 높고 소형 제작이 가능한 것을 보였다.

현재까지는 실시간 합성기의 구현에 연구의 초점을 맞추었기 때문에 합성 필터 구조의 변경을 시도하지 않았고 Driving Function 및 Bit-rate 의 변경에 따른 software 처리에 관하여 고려하지 않았다. 이는 앞으로의 실용적인 시스템 개발을 위하여 보완되어야 할 사항이다.

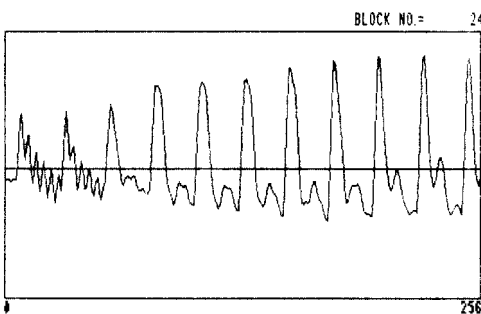
감 사 의 글

이 연구 수행중 많은 도움을 주신 KETRI 음향연구실 연구원들께 감사드립니다.

참 고 문 헌

1. J.D. Markel and A.H.Gray, Jr. "Linear Prediction of Speech"
2. Markel and Gray " A Linear Prediction Vocoder Simulation Based Upon the Autocorrelation Method ", IEEE Trans. ASSP-22, 124-134 (1974a)
3. Markel and Gray " Fixed-Point Truncation Arithmetic Implementation of a Linear Prediction Autocorrelation Vocoder ", IEEE Trans. ASSP-22, 273-281 (1974b)
4. NEC 77P20 Simulator Operating Manual
5. NEC 77P20 Assembler Operating Manual
6. Joel A.Feldman, Edward M.Hofstetter & Marilyn L. Malpass " A Compact, Flexible LPC Vocoder Based on a Commercial Signal Processing Micor-computer " ASSP-31 Feb. 1983.

SPEECH WAVEFORM



SPEECH WAVEFORM

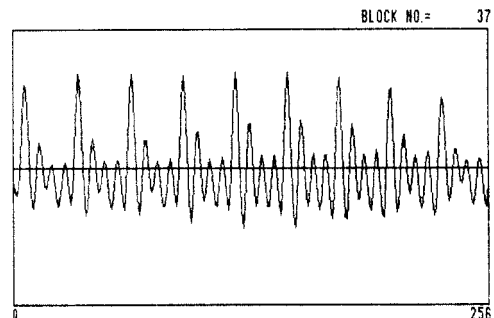


그림 4-1. 합성 음성의 파형

발성자 : 여자성우 (26세)
내 용 : 음성 메모 서비스