

PCM을 이용한 16 Polyphonic One Channel Music Sound  
Synthesis에 관한 연구

구재울\*      장은영\*\*      김원후\*\*

\* 삼익악기 연구소  
\*\* 한국항공대학 항공전자공학과

A study on the 16 Polyphonic One Channel Music Sound  
Synthesis by the PCM

Jae UI KOO\*      Eun Young CHANG\*\*      Won Hoo KIM\*\*

\* Samick Musical Instrument RMD  
\*\* Dept. of Avionics Eng., Hankuk Aviation Univ.

**ABSTRACT:** In the electro-music synthesis, there has been many trial to make electro-music instrument with many acoustic sound at the least memory size. This study is presented one way that reduce memory and hardware size with DPCM, in which is used to one sound table. In this result, it can be processed 16 sound element in 1 channel at once.

I. 서론

디지털 기술의 급속한 발달은 전자 음악 분야에서 새로운 영역을 창조하는데 결정적인 기여를 하였다. 그러나 디지털 음원이 아날로그보다 더 좋은 음악을 만드는 것이 아니고, 디지털 음원 회로가 아날로그보다는 음에서의 좀더 정확한 처리와 복잡한 변화를 가능하게 해준다는 것이다. 현재까지 널리 사용되고 있는 음원합성방식에는 Subtractive Synthesis, Digital Phase Distortion Synthesis, Digital FM Synthesis, Digital Sampling (1)(2) 등이 있으며 본 연구에서는 Digital Sampling 에 기초하여 음성신호와 영상신호의 Data 감축 방식에서 많이 사용되는 DPCM (Differential Pulse Code Modulation) 기법을 이용하여 memory를 좀더 효율적으로 활용하는 방법을 제시한다.

부분, coefficients와 boundary condition 부분 그리고 control function generator와 audio waveform generator로 구성된다.

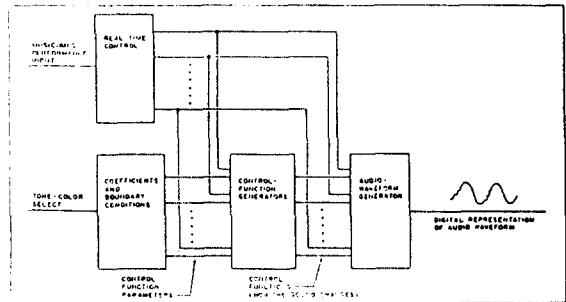


그림 1. 대표적인 음원합성 시스템

II. Musical Sound Synthesis System

본 연구의 Sound Synthesis System에 대한 control hierarchy를 그림 1에 나타낸다. 크게 Real-time control

III. Data 감축을 위한 신호처리

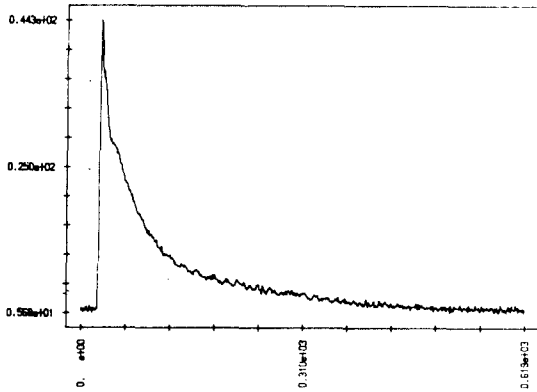
저장 memory의 용량은 key number \* sampling rate \* 지속시간 \* Data bit수로 주어지게 되어 88개 건반에 있어

한국통신학회 1990년도 추계종합학술발표회 논문집('90. 11)

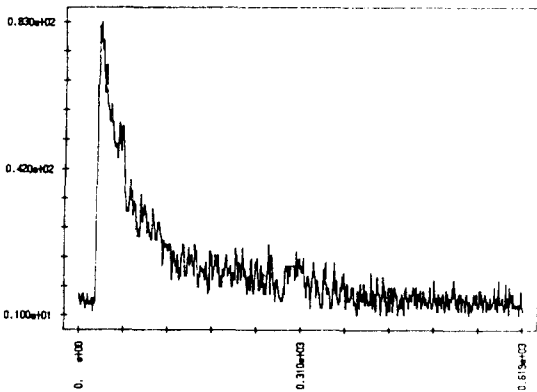
20 kHz의 표본화 속도를 가지면서 약 10초간의 지속시간과 10 bit Data를 그 채널로 저장한다면 많은 양의 memory가 필요하여 가격상승과 신뢰도 저하가 문제된다.

1. 파형의 Depetition에 의한 Data 감축

본 연구에서는 피아노 파형만을 위주로 하여 실험을 했으며 피아노 파형을 Acquisition하여 그 형태를 그림 2에 나타내었다.



(가) power envelope

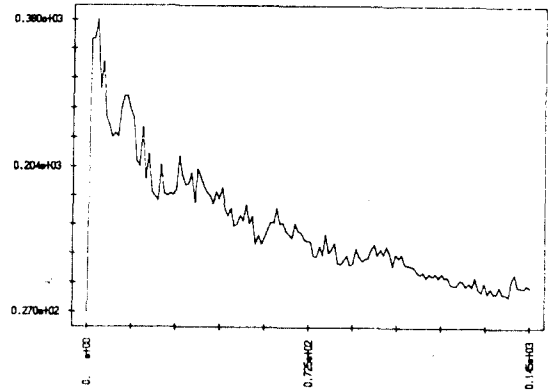


(나) peak value envelope

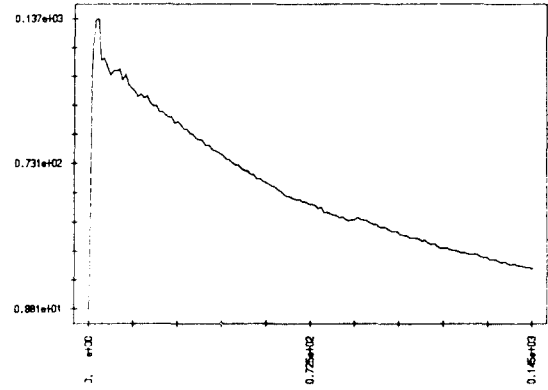
그림 2. 30번째 건반의 power envelope와 주기별 peak value envelope

그림 2와 그림 3은 피아노 파형의 변형을 시간축상에서 알아보기 위하여 각주기별로 두개의 피아노 건반에서 20kHz로 sampling하여 각주기의 power와 peak의 특성을 나타낸 것이며 여기서 전체 파형의 1/4 이후에서는 Data가 단조로운 형태로 Data가 변한다.

그림 2와 3은 피아노 건반의 누르는 세기에 따라 모양이 바뀔 수 있으나 기본적인 형태는 동일한 형식을 취한다.



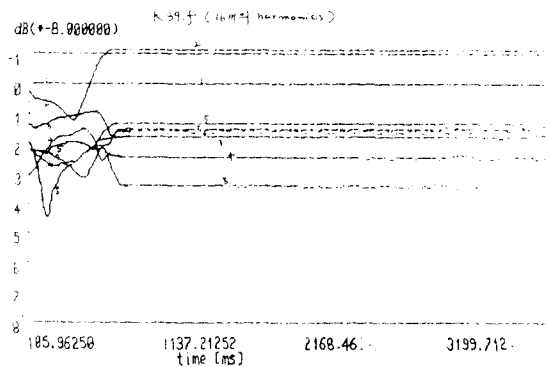
(가) power envelope



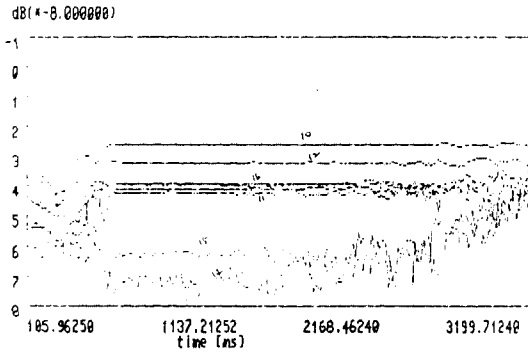
(나) peak value envelope

그림 3. 10번째 건반의 각주기 power envelope와 각주기 peak value envelope

그림 4는 주파수축상에서 피아노 파형의 16개 Harmonics들이 시간에 따라 변화된 형태로, 건반을 누르고 나서 일정 시간 지연후의 파형의 변화 형태가 단조로와짐을 알 수 있다.



(가)



(나)

그림 4. 39번째 건반의 주파수축상에서 본 Harmonics의 변화형태

위의 상태에서 볼때 attack과 delay부분의 파형은 transient한 성분이 많이 포함되어 매 주기의 파형이 불규칙하고 변화가 심하다. 반면 decay를 지나면서 sustain부분으로 들어서면 파형은 stable해져서 각 주기의 파형이 거의 같은 모양을 갖게 된다.

따라서 뒷부분의 파형을 전부 사용하는 대신 normalize 하여 파형을 단순화 시킨 후에 다음과 같이 반복주기를 구한다.

$$E_L = \sum_{n=0}^N (x(n) - (n+1))^2, \quad 1 = 1, 2, \dots, L \quad (1)$$

여기서 N : compare length  
L : interval limit

## 2. DPCM을 이용한 Music Sound 합성

DPCM을 이용한 음원합성의 양자화 식은 다음과 같다.

$$\begin{aligned} d(n) &= x(n) - \hat{x}(n) \\ u(n) &= d(n) - q(n) \\ y(n) &= \hat{x}(n) - v(n) \end{aligned} \quad (2)$$

여기서 y(n) : DPCM approximation to coder input x(n)  
d(n) : prediction error  
q(n) : quantization error  
u(n) : quantized prediction error  
v(n) : receiver version  
x(n) : input data

$$\begin{aligned} r(n) &= x(n) - y(n) \\ &= x(n) - \hat{x}(n) - (x(n) - \hat{x}(n)) + q(n) \\ &= q(n) \end{aligned} \quad (3)$$

위 식을 살펴보면 sample n의 reconstruction error r(n)은 quantization error q(n)과 동일함을 알 수 있다.

이전의 data를 가지고 다음 정보를 예측하는 predictor는 다음식 (4)와 같다.

$$x(n) = h_1 x(n-1) + h_2 x(n-2) + \dots + h_m x(n-m) \quad (4)$$

$h_1, \dots, h_m$ 의 계수들은 prediction error로

$$d(n) = x(n) - \hat{x}(n) \text{ 이}$$

최소가 되도록 선정되어야 한다.

prediction error d(n)의 pdf를 살펴보면 uniform한 분포가 아닌 Gaussian 분포가 많은 nonuniform한 pdf를 가지고 있으므로 quantizer 역시 nonuniform quantizer를 사용하게 된다.

이러한 quantizer로 simulation에 사용된 것이 Lloyd와 MAX가 제안한 Lloyd-MAX quantizer이다 (3)-(5).

quantizer는 아래와 같은 flow로 구하였다.

- 1) 각 interval limit는 이웃 level의 중간지점으로 택한다. 즉,

$$x_i = \frac{y_i + y_{i+1}}{2} \quad (5)$$

- 2) 각 level은 그 level이 속해 있는 interval의 pdf가 가지는 면적을 양분하는 점으로 한다. 즉,

$$\int_{x_{i-1}}^{x_i} (x - y_i) p(x) dx = 0 \quad (6)$$

- 3)  $y_{M/2}$ 가  $x_{M/2-1}$ 과  $\infty$  사이의 면적의 중심에 위치하고 있는지 다음식을 이용하여 확인한다.

$$\int_{x_{M/2-1}}^{\infty} (x - y_{M/2}) p(x) dx = 0 \quad (7)$$

- 4) 위의 식을 만족시키지 않으면 1)로 돌아가 y를 재설정하여 process를 반복한다.

## IV. 16 Polyphonics One Channel DPCM H/W 고찰

본 연구에서는 Hardware를 간단히 구성하기 위하여 first order predictor를 사용한다. 즉,

$$\hat{x}(n) = h_1 x(n-1) \quad (8)$$

또한 적절한 bit rate를 결정하기 위하여 critical band analysis를 이용한다.

critical band analysis를 위해 먼저 FFT를 통하여 original 신호의 spectrum  $Re(w)$ ,  $Im(w)$ 를 얻은 후 다음 식에 의하여 얻는다.

$$P(w) = Re^2(w) + Im^2(w) \quad (9)$$

spectrum은 25개의 critical band로 나눈 다음

$$B_i = \sum_{w=b_{li}}^{b_{hi}} P(w) \quad \text{를 구한다.}$$

여기서  $b_{li}$  : critical band  $i$ 의 power boundary  
 $b_{hi}$  : " upper boundary  
 $B_i$  : " energy  
 $i$  : critical band number

noise의 정량적인 측정을 위해 위의 특성을 이용한 그림 5의 noise masking threshold를 이용한다.

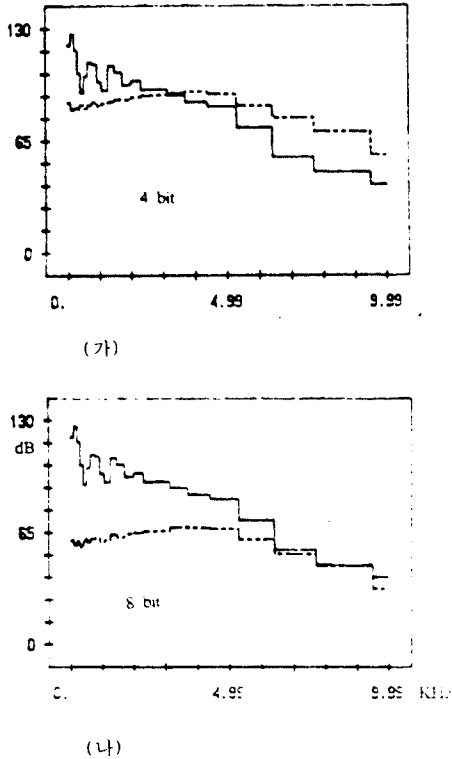


그림 5. Masking threshold

그림 5에서 알 수 있듯이 4-7 bit를 이용하여 data compression을 행한 경우는 masking threshold보다 reconstruction level이 올라가므로 8 bit로 했을 때가 가장 적절한 compression이라 판단되어 8 bit quantization을 이용한다.

### V. 결 론

본 연구는 음성합성 이론에 바탕을 둔 것으로 Digital Sampling 방법을 이용하여 Music Sound Generator를 구성 할때 큰 문제가 되는 음원합성 방식을 DPCM을 이용하여 제안하였으며 가장 중요한 문제의 하나인 Memory절약 문제에

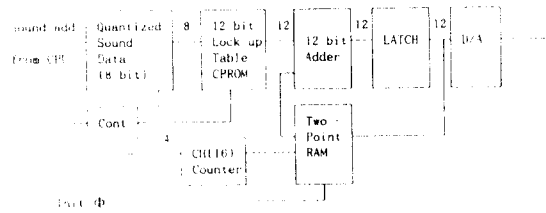


그림 6. DPCM Block Diagram

대하여 개선된 방식을 제안한다.

또한 충실한 sampling data만 있다면 거의 완벽한 원음을 재생할 수 있으며 하나의 look up table만을 사용하여 여러가지 약기음의 생성가능하다고 판단된다.

### 참 고 문 헌

- [1] Hal Chamberlin, " Musical Applications of Micro-processors , " Hayden Book Company, 1985.
- [2] Thomas J. Lynch, " Data Compression, " Lifetime learning Publications, 1985.
- [3] J.D. Johnston, " Estimation on Perceptual Entropy Using Noise Masking Criteria," in Proc. ICASSP, pp. 2524-2527, 1988.
- [4] John W. Gordon, " System Architectures for Computer Music," Computing Survey, Vol. 17, No. 2, June 1985.
- [5] N.S. Jayant, Peter Noll, " Digital Coding of Waveforms," Prentice Hall, 1984.