

피아노 연주곡의 wave sound에서의 피치추출 및 악보생성¹⁾

최제현^o, 김규년
울산대학교 대학원 전자계산학과

Pitch extraction and generating a piano score from a wave sound of the piano music

Jae-Heon Choi, Kyoo-Nyun Kim
Dept. of Computer Science, University of Ulsan

e-mail: heony@cic.ulsan.ac.kr
knkim@uou.ulsan.ac.kr

요 약

본 논문에서는 피아노 음으로 구성된 간단한 웨이브(wave) 형식의 연주곡으로부터 각 음들의 음높이인 피치(pitch)를 추출하는 방법과 연주곡의 악보를 생성하는 방법에 대해 연구한다. 본 연구의 최종 목적은 위의 방법을 사용하여 클래식 악기로 연주된 CD-DA방식의 음악을 같은 음질이 보장되는 MIDI[1,3,7,8] 데이터로 변환하는 데 있다.

이번 연구에서는 하나의 악기(피아노)로 연주된 음악으로 고정시켰으며, 잡음이 없는 깨끗한 웨이브 사운드로 제한하였다. 웨이브 사운드 형식으로 된 간단한 피아노 연주곡을 입력 데이터로 사용하였다. 이 웨이브 사운드에서 Autocorrelation[4,5,6]방법을 사용하여 각 음들의 피치를 추출하였다. 또한, 웨이브 사운드로부터 추출한 음의 피치, 음의 지속시간 등의 정보를 사용하여 연주곡에 대한 악보를 생성하는 방법에 대해 연구하였다.

본 논문에서는 웨이브 사운드 데이터에서 악보로의 변환까지를 다루고 있지만, 향후 실제 클래식 악기와 흡사한 MIDI 음원모듈이 제공될 경우 CD-DA방식의 원음과 같은 음질의 MIDI 음악을 생성하는 것도 가능할 것이다.

1. 서론

일반적으로 MIDI 데이터를 웨이브 형식의 데이터로 변환하는 방법은 이미 제시되었으며, 여러 가지 도구가 개발되어 있다. 하지만, 그 반대의 웨이브 데이터를 MIDI 데이터로 변환하는 방법은 음악인식(Music Recognition)의 문제이며[9], 아직 만족할만한 방법이 없다. 또한, 현재도 연구중인 부분이다. 현재 인쇄문서를 스캐닝하여 인식하는 것은 거의 99%정도 해결되어 있으며, 음성인식 역시 80 ~ 90% 정도의 정확성을 보이고 있다[9]. 하지만, 음악인식은 단지 하나의 악기로 연주된 음악에서만 약간의 정확도를 보일 뿐이다.

따라서, 본 논문에서는 음악인식의 시작단계로서 웨이브 데이터로부터 피치를 추출하는 방법과 그 결과를 이용하여 악보를 재구성하는 방법을 연구하였다. 물론 하나의 악기로 연주된 음악으로 제한하였으며, 사용한 음악은 피아노 음으로 구성된 간단한 음악이다.

2. 연구내용

2.1 웨이브 데이터로부터의 피치추출

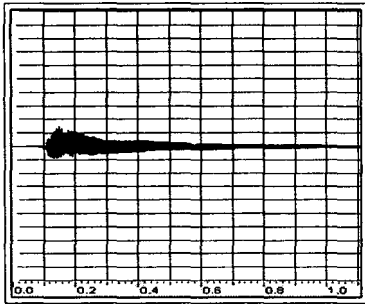
웨이브 데이터로부터 피치를 추출하기 위해 반복주기 추출법[10] 중 하나인 Autocorrelation 방법을 적용하였다. 입력으로 주어진 웨이브 데이터를 일정구간으로 나누어 각 구간에서 Autocorrelation을 적용하여 피치를 추출한다. 구간은 (Sampling rate / 10) 으로 설

1) 정보통신부에서 지원하는 태학기초연구지원사업으로 수행

정하였으며, 이 구간에서 (식1)을 사용하여 반복주기를 이루는 결과를 얻고, 이 결과로부터 피치를 추출한다.

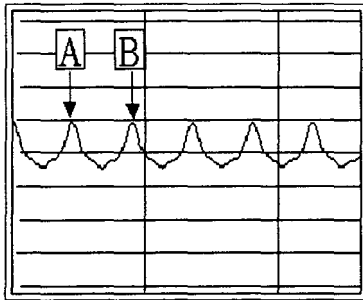
$$Autocorrelation(k) = \frac{1}{m} \sum_{n=0}^{m-1} x(n)x(n+k) \quad (\text{식1})$$

$x(n)$ 은 구간내의 원래 샘플값이며 $x(n+k)$ 는 해당 신호를 k 만큼 이동(shift)시킨 신호이다. k 는 0에서 $m-1$ 까지 변한다. [그림1]은 피아노의 4옥타브 라 음인 A4음의 파형을 보여준다.



[그림 1] 피아노 A4음의 파형

[그림1]을 입력으로 했을 경우, 어떤 구간에서의 Autocorrelation을 사용하여 결과로 얻은 파형의 일부를 확대한 것이 [그림2]이다. [그림2]에서는 반복되는 주기를 가지는 형태가 나타내며, 이 한 주기가 해당 음의 주기가 된다. 따라서, 이 한 주기의 샘플 개수를 구한 후 (식2)를 사용하여 음의 피치를 계산할 수 있다.



[그림 2] A4음의 주기추출

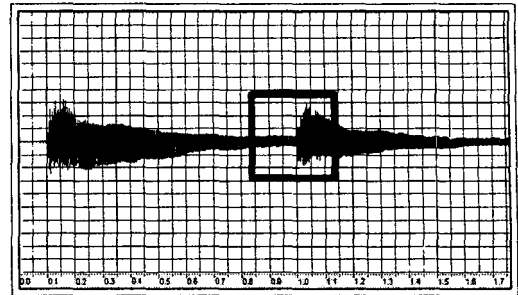
$$\text{피치(Hz)} = \frac{\text{Sampling rate}}{\# \text{ of Samples between A and B}} \quad (\text{식2})$$

본 연구에서 사용한 A4음은 sampling rate가 44100 Hz이며, A4음의 주파수는 440Hz이다[2]. 실험결과 약 436Hz ~ 445Hz 사이의 값으로 추출되었으며 이는 A4음에 근사한 값이며, A4음으로 간주할 수 있다.

2.2 두 개 이상의 음의 인식

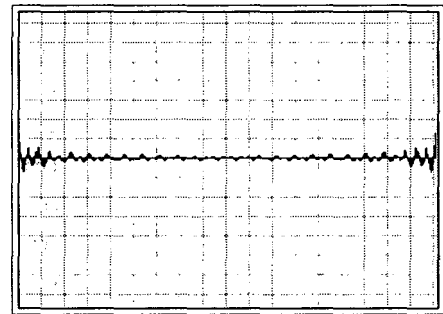
하나의 피아노 음으로부터 피치를 추출하는 것은 비교적 근사한 값을 얻을 수 있었다. 이번에는 두 개 이상의 음이 이어져 있을 경우의 피치추출에 대해 살펴본다. [그림3]은 피아노의 C4음과 A4음이 이어져 있는

웨이브 모습이다.



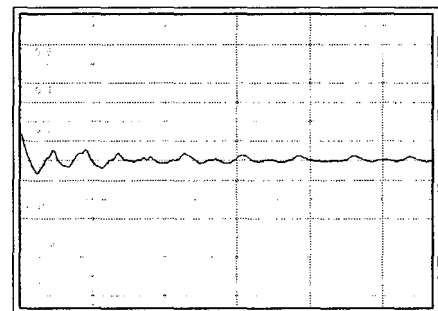
[그림 3] C4음과 A4음의 파형

[그림3]에서 중앙의 사각형 영역은 C4음과 A4음이 이어져 있는 부분이다. 이 부분에서 Autocorrelation을 하여 얻은 결과는 [그림4]와 같다. [그림4]에서 알 수 있듯이 음과 음이 이어진 부분에서는 Autocorrelation 결과 파형이 0에 가까이 간다는 것을 알 수 있다.



[그림 4] 그림3의 사각형 영역의 Autocorrelation 결과 파형

[그림4]의 앞부분을 확대하면 [그림5]와 같다. [그림5]를 보면 Autocorrelation결과 파형이 주기적인 패턴이 아닌 약간의 불규칙적인 형태를 이루고 있음을 알 수 있다.



[그림 5] 그림4의 앞부분의 확대

[그림5]에서는 [그림2]와 같이 비슷한 주파수의 주기가 반복되는 것이 아니라 서로 다른 주파수 값들이 이어져 나타난다. 이 부분을 처리하기 위하여 가장 작은 음표의 연주시간을 고려하였다. 즉, 비슷한 주파수가 반복되어 나타나는 총 시간이 가장 작은 음표의 연주시

간보다 작다면 그 부분의 주파수들은 무시하는 것이다.

본 연구에서는 $J=120$ 으로 설정하였다. 이것은 4분음표를 1분에 120번 연주하는 속도이다[11]. 즉, 4분음표 하나의 연주시간은 0.5초이며, 가장 작은 음표인 32분음표의 연주시간은 $0.5/8 = 0.0625$ 초이다. 따라서, 비슷하게 반복되는 주파수의 연주시간이 0.0625초보다 작을 경우는 피치인식에서 제외하였다. 여기에서 약간의 오차가 발생했으며 앞으로 줄여 나가야할 부분이다.

2.3 간단한 음악을 이용한 피치분석

이번에는 간단한 노래 하나를 입력 데이터로 사용하여 피치추출을 해 보았다. 초등학교에서 배웠던 "학교중"이라는 노래의 피아노 음으로 구성된 웨이브 데이터를 사용하였다. "학교중"은 도, 레, 미, 솔, 라 의 5개 음으로만 구성된다. 실험에 사용한 음은 C3, D3, E3, G3, A4 이며, 각 음의 주파수는 [표1]과 같다.

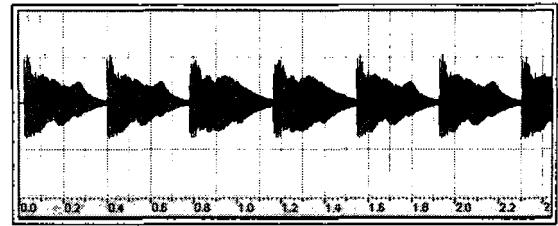
음 (피치)	주 파 수
B3 (시)	247.21 Hz
C3 (도)	261.86 Hz
C3#	277.47 Hz
D3 (레)	293.91 Hz
E3b	311.43 Hz
E3 (미)	329.99 Hz
F3 (파)	349.55 Hz
F3#	370.39 Hz
G3 (솔)	392.34 Hz
G3#	415.72 Hz
A4 (라)	440.00 Hz
B4b	466.60 Hz

[표 1] 각 음의 주파수(평균음)

[표1]에서 나타난 주파수들은 여러 가지 조율체계[2] 중의 하나인 평균음(Temperament)[2]에 기준을 둔 것이다. 본 실험에서 나온 결과의 주파수와 [표1]의 주파수를 비교하여 가장 가까운 피치를 선택한다. 실험결과 해당 피치의 주파수에 가까운 값들이 반복적으로 측정되었다.

특정 피치의 주파수대가 반복된 후 새로운 주파수가 나오는 부분에서 다시 Autocorrelation 구간을 지정하여 새로이 Autocorrelation을 적용하여 결과를 구하는 방식으로 피치추출을 하였다.

[그림6]은 입력 데이터인 "학교중" 음악의 웨이브의 앞부분을 보여준다. 실험결과 측정된 5개음의 주파수는 [표2]와 같이 나왔다.

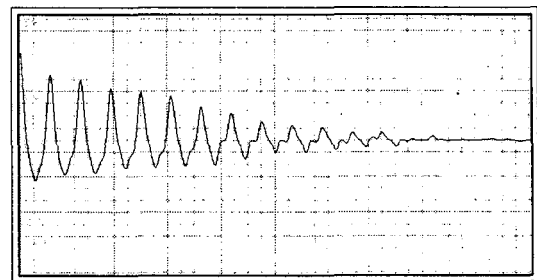


[그림 6] "학교중"노래의 피아노 음의 웨이브

음 (피치)	측정된 주파수대
C3 (도)	260.95 ~ 264.07 Hz
D3 (레)	292.05 ~ 295.97 Hz
E3 (미)	326.67 ~ 331.58 Hz
G3 (솔)	383.48 ~ 390.27 Hz
A4 (라)	436.63 ~ 445.45 Hz

[표 2] 측정된 5개음의 주파수

[표2]의 결과를 보면 비교적 [표1]의 평균음 주파수에 근접한 값들이 나왔다. 하지만, 음과 음이 이어지는 부분에서는 전혀 다른 주파수 값들이 계산되어졌다. 이것은 2.2절에서 설명한 바와 같은 오차이다. 즉, "학교중" 노래의 음과 음이 이어지는 부분에서는 [그림7]과 같이 Autocorrelation 결과 파형이 다양한 패턴을 나타낸다. 이런 패턴을 육안으로는 구분이 가지만 알고리즘적으로 구분하는데 에서 오차가 발생한 것이다. 현재 몇가지 규칙을 두고 패턴을 구분하여 주기를 찾고 있지만, 각각의 음들마다 다양하게 패턴이 발생하여 약간의 오차가 발생하고 있다. 이 부분은 앞으로도 계속 연구해야할 부분이다.



[그림 7] Autocorrelation결과의 다양한 패턴

2.4 악보의 생성방법

웨이브 데이터로부터 악보를 생성하려면 앞서 구한 피치정보와 해당 피치의 지속시간 정보가 필요하다. 어떤 피치의 지속시간을 구하려면 (식3)과 같이 구하면 된다.

$$\text{피치의 지속시간} = \frac{\text{연속된 동일 주파수 구간의 샘플개수}}{\text{Sampling rate}} \quad (\text{식}3)$$

(식3)에서 분자는 Autocorrelation결과 파형에서 비슷한 주파수대가 반복되는 구간내의 샘플개수를 나타낸다. 예를 들어, Sampling rate가 44100 Hz이고 결과 파형에서 441 Hz의 주기가 10번 나타났다면, 한 주기당 샘플의 개수는 $(44100 / 441) = 100$ 개, 총 샘플 개수는 $(100 \times 10) = 1000$ 개이다. 따라서, 441 Hz 피치의 지속시간은 $(1000 / 44100) = 0.023$ 초이다.

또한, 4분음표 하나의 길이가 얼마인지를 알아야 하는데, 이것은 보통 음악에서 $J=120$ 처럼 수치로 표현한다. $J=120$ 는 4분음표 하나를 1분에 120번 연주하는 속도라는 말이다. 즉, 4분음표 하나가 0.5초의 연주 시간이다. 하지만, 웨이브 데이터에는 이런 정보가 없다. 현재 연구단계에서는 미리 이 값을 설정한 후 악보 정보를 구하고 있다. 이것을 해결할 방법으로는 추출된 모든 피치들의 연주시간의 상대적인 비교로 그 값을 계산하는 방법을 고려해볼 수 있다. 현재 이 방법으로 연구중이며 이것만 해결된다면 웨이브 데이터로부터 악보를 어느 정도 구성할 수 있을 것이다. 물론 원래 악보와 동일하게 생성하기 위해서는 그 외 부가적인 정보가 더 필요하지만, 위의 정보만으로도 어느 정도의 악보는 생성할 수 있을 것이다.

3. 결론 및 향후 연구계획

본 논문에서는 단일 악기음으로 구성된 웨이브 데이터로부터 피치정보와 각 피치의 연주시간 정보를 추출하고, 이를 바탕으로 악보를 생성하는 방법에 대해 연구하였다. 다소의 오차는 발생하였지만 계속적으로 연구하여 이런 오차를 감소시키고 보완한다면 단일 악기음으로 연주된 웨이브 사운드로부터 완전한 악보를 만들어내는 것이 가능할 것이다.

앞으로 계속 연구해 나갈 방향은 본 연구에서 나타난 결과를 더 보완하여 웨이브 데이터로부터 악보를 생성하고, 또 나아가 웨이브 데이터를 MIDI 형식의 데이터로 변환하는 방법을 연구할 것이다.

4. 참고문헌

[1] Curtis Roads, *The Computer Music Tutorial*, MIT Press, 1998
 [2] Donald E. Hall, *Musical Acoustics*, Brooks/Cole Publishing Company, 1990
 [3] Max V. Mathews and John R. Pierce, *Current Directions in Computer Music Research*, MIT Press, 1989
 [4] Alan V. Oppenheim and Alan S. Willsky, *Signals & Systems*, Prentice Hall, 1997

[5] Ken Steiglitz, *A DSP Primer with Applications to Digital Audio & Computer Music*, Addison-Wesley Publishing Company, 1996
 [6] Emmanuel C. Ifeachor and Barrie W. Jervis, *Digital Signal Processing A Practical Approach*, Addison-Wesley, 1996
 [7] Moog R., "MIDI: Musical Instrument Digital Interface.", *Journal of the Audio Engineering Society* 34(5): 394-404, 1986
 [8] Rothstein J., "MIDI: A Comprehensive Introduction", Madison: A-R Edition, 1992
 [9] AKoff Sound Labs, "What is Music Recognition", <http://akoff.hypermart.net>, 1999
 [10] 박경범, *음성의 분석 및 합성과 그 응용*, 도서출판 그린, 1996
 [11] 이성천, *음악통론과 그 실습*, 음악예술사, 1997