

악기 튜닝을 위한 피치 검출 알고리즘

이 강성

광운대학교 컴퓨터공학부

Pitch Detection Algorithm for the Tuning of Musical Instrument

Gang Seong Lee

Computer Engineering Dept., Kowangwoon Univ.

gslee@mail.gwu.ac.kr

요 약

본 연구에서는 악기 피치 검출에서 샘플링 오차로 인하여 발생하는 문제점을 살펴보고, 보간법을 적용한 피치 검출 실험 예를 소개한다. 악기의 피치 검출과 음성의 피치 검출의 차이점을 보이고 그 처리 절차를 소개한다. 제안된 알고리즘으로 검출된 피치를 이용하여 악기가 연주되고 있는 음정을 추정하여 음표로 변환한다. 단선율의 솔로 악기 연주 파일(.wav)을 midi 파일로 변환하여 구현된 알고리즘의 유용성을 평가하였다.

1. 서 론

음성의 피치 검출은 음성 분석, 음성 인식, 음성 합성, 음성 코딩등의 분야에 폭넓게 사용되어왔다. 하지만 음악을 위한 음성 혹은 악기의 피치 검출은 음성과는 차이점이 존재한다.

1. 악기의 음정은 호흡이나 연주 기법에 따라 음정은 변화할 수 있지만 비교적 안정되고 구분된 음정 영역을 갖는다.

2. 각각의 악기는 음역에 따른 음정 변화가 생긴다. 예를 들어 플루트인 경우 저음에서는 음정이 더 떨어지며, 고음에서는 음정이 상승하는 경향이 있으며, 리드 악기들은 리드의 장으로 인해 고음에서는 음정이 떨어지기 쉽다.

3. 악기의 음역은 상당히 폭 넓어서 25Hz (더블 베이

스) ~ 4200Hz(피콜로)사이의 주파수를 갖는다.

4. 저음일수록 풍부한 배음들을 만들어내나, 고음에서는 배음의 영향이 적어, 파형이 사인파에 가까워진다. 따라서 배음 효과로 인한 피치 검출의 오류는 줄어들지만, 샘플링 주파수의 한계로 인해 에러가 오히려 더 많이 발생한다.

5. 피치 검출에 있어서 음성은 자연성을 나타내는 피치 검출에 중점을 두지만, 악기는 정확한 음정(note)을 표현하는데 더 중점을 둔다. 따라서 필터를 통한 완만한 피치 트랙의 추적은 별 의미가 없다.

본 연구에서는 고음의 소리를 피치 검출하는데 있어서 발생하는 문제점[1]을 우선 살펴보고 보간법[2]을 이용한 피치 검출 알고리즘이 얼마나 효과가 있는지 알아본다. 또, 검출된 피치를 이용하여 음정(note)을 추적하는 알고리즘을 제안한다. 그리고 불안한 과도구간을 처리하여 좀 더 안정된 연주를 표현하도록 하는 음정 정정 부분의 간단한 알고리즘을 소개한다. 다음 그림 1에 전체 시스템 블록도를 보인다.

2. 피치 검출

고음의 피치 검출은 샘플링 주파수의 한계로 오차가 발생한다. 예를 들어 샘플링 주파수(Fs)를 44100Hz라고 할 때, 다음 각 주파수에 있어서 1주기당 샘플 수(sp)와 sp-1 개의 샘플에 대한 주파수 변화를 표 1에 보인다.

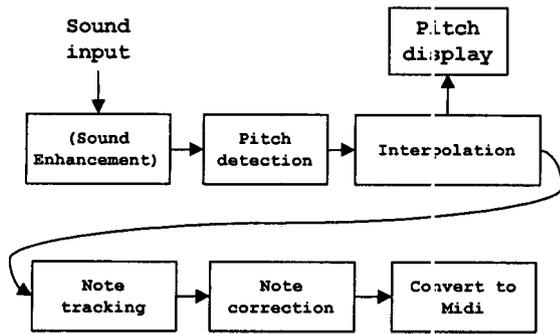


그림 1. 시스템 블록도

음계는 12평균률(옥타브를 같은 비율로 12등분 한 것)을 사용했으며, 센트(cent)는 두 음 사이의 음정을 표현하는 단위로 1 옥타브를 대수적으로 1200등분한 단위이다. 따라서 1 센트는 다음과 같이 표현된다.

$$c = (1/1200) \log 2 \quad (1)$$

두 주파수(f_1, f_2) 사이의 센트(d)는 다음 식으로 계산된다.

$$d = (\log f_1 - \log f_2) / c \quad (2)$$

주파수 f 에서 한 샘플이 줄었을 때의 주파수(f_{-1})는

음정	센트	주파수	1주기당 샘플 수	센트 변화	주파수 변화
A0	0	55	802	2	0.07
A1	1200	110	401	4	0.28
A2	2400	220	200	9	1.10
A3	3600	440	100	17	4.43
A4	4800	880	50	35	17.92
A5	6000	1760	25	71	73.16
A6	7200	3520	13	144	305.33

표 1. 주파수 변화에 따른 음정 변화 다음과 같다.

$$f_{-1} = (Fs * f) / (Fs - f) \quad (3)$$

f 와 f_{-1} 의 센트 변화는 식 (2)로 계산된다. 센트는 진동수 비를 생각하지 않고 반음 간격을 100센트, 온음을 200센트로 산술적으로 계산할 수 있어 음 사이의 간격을 계산하기에 적합하다.

표1에서도 알 수 있듯이 1760Hz에서 1샘플의 차이는 다른 음정으로의 도약을 의미한다. 이것은 음악가들이 받아들이기 어려울 만큼 상당히 큰 오차이다.

자동 상관 계수를 이용한 피치 계산법[3]을 적용할 때 계산된 피치에 2차함수(포물선) 보간법을 적용함으

로 이러한 오차를 어느 정도 줄일 수 있다. 물론 시간 축 및 주파수 축상에서 피치를 계산하는 다른 기법등에도 적용 가능하다.

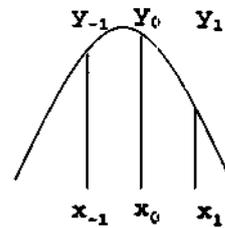


그림 2에서 피크 점(x, y)은 다음 식으로 계산된다.

$$\begin{aligned}
 c &= x_0 \\
 b &= (x_{+1} - x_{-1}) / 2 \\
 c &= (x_{-1} + x_{+1}) / 2 - x_0 \\
 x &= -b/(2a) \\
 y &= ax^2 + bx + c
 \end{aligned}$$

그림 2. 포물선 보간법

3. 음정 추적

음정 추적의 목적은 어느 정도의 연주 피치 변화에도 연주자가 의도했던 음정을 찾는 것을 말한다. 이러한 음정 변화를 일으키는 요인으로는 가장 크게는 비브라토가 있고, 악기의 특성상의 요인이 있다.

비브라토에 의한 최대 음정 변화 폭을 V_{max} 라고 하고 악기 특성상의 최대 변화 폭을 C_{max} 라 하자. 어떤 기준 점에서 변화 가능한 음정의 최대 폭은 $V_{max} + C_{max}$ 이므로 상하로 허용해야 할 음정 변화 폭은 다음과 같다.

$$-(V_{max} + C_{max}) / 2 \sim (V_{max} + C_{max}) / 2$$

하지만 악기 특성 변화에 다른 중간 기준점을 정확하게 잡는다는 가정을 할 수 없다. 기준을 C_{max} 의 최 하위점으로 했을 때 변화 폭은

$$-V_{max}/2 \sim (C_{max} + V_{max}/2)$$

이고 기준을 C_{max} 의 최 상위점으로 했을 때 변화 폭은 다음과 같다.

$$-(C_{max} + V_{max}/2) \sim V_{max}/2$$

따라서, 허용해야 할 전체 음정 변화 폭은 다음과 같다.

$$-(C_{max} + V_{max}/2) \sim (C_{max} + V_{max}/2)$$

만일 $(C_{max} + V_{max}/2)$ 가 50센트를 넘을 경우는 다른 음정과 중첩 범위가 생긴다. 이를 처리하기 위해서 비브라토 등에 의한 최대 음정 변화(V_{vm})를 계산한다. 만일 두 음정 사이가 이 음정 변화 내에 있고 또한 음정 변화 구간 $(C_{max} + V_{max}/2)$ 내에 있다면 이전 노트와 같은 음정으로 처리하고, 그렇지 않으면 가장 가까운 음정(note)를 취한다.

$V_{max}/2$
C_{max}
$V_{max}/2$

그림 3. 음정의 변화 구간

4. 음정 정정

음악은 말소리와는 달리 음과 음 경계에서 급격한 음정 변화를 일으킨다. 고정된 크기의 프레임으로 피치를 검출하다보면 이러한 과도구간에서 피치가 불안정하게 계산된다. 이러한 불안한 구간을 음성에서와 같이 메디안 필터 등으로 적당히 부드럽게 할 수는 없다. 변화 구간에서의 불안한 음정은 대부분의 경우 앞 음정과 뒤 음정의 과도 구간이므로 어느 쪽으로 취해져도 무방할 것 같아 보이나, 실제로는 음의 시작에 영향을 미치고 이것은 박자에 영향을 줄 수도 있다.

기본적인 처리 방법으로는 과도구간을 찾아내고 그 구간을 둘로 나누어 앞 뒤의 음표의 길이를 늘려주거나(1) 어느 한쪽(주로 왼쪽)의 음표만을 늘려주는 방법(2)이다. 세그먼트의 구분이 명확한 두 번째 방법을 사용하기로 한다(예비 실험결과 더 안정된 결과를 주었다).

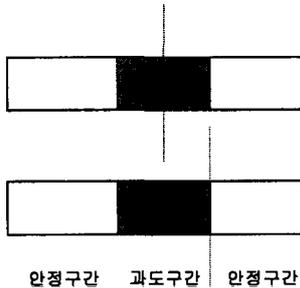


그림 4. 간단한 음정 정정 방법

5. 실험 및 고찰

5-1. 보간법 실험

피치 검출에 보간법을 사용한 경우와 사용하지 않은 경우를 비교한다. 우선 컴퓨터로 다음 8개의 주파수를 갖는 사인 웨이브 파일을 생성했다. 샘플링 주파수는 22050Hz를 사용하였다.

55Hz, 110Hz, 220Hz, 440Hz, 880Hz, 1760Hz, 3520Hz, 7040Hz

플리핑을 적용한 자동상관계수를 이용한 피치 검출 방법을 사용했을 경우 다음과 같은 결과를 얻었다.

여기서 f_n 은 보간법을 적용하지 않았을 경우의 피치 검출 결과의 차이를 센트로 표시한 것이고 f_i 는 보간법을 적용한 피치 검출 결과의 차이이다. 보간법을

기본 주파수	f_i (cent 차이)	f_n (cent 차이)
55	54.99(0)	54.99(0)
110	110.25(4)	109.99(0)
220	220.50(4)	220.01(0)
440	441.00(4)	440.00(0)
880	882.00(4)	880.03(0)
1760	1696.15(64)	1759.83(0)
3520	3675.00(75)	3529.41(5)
7040	7350.00(75)	7144.32(25)

표 2. 보간을 적용했을 경우와 적용하지 않았을 경우의 사인 파형의 주파수 분석결과 적용했을 때 상당한 보상이 이루어짐을 알 수 있다.

실제로 플롯의 높은 G음을 녹음해서 분석한 결과를 다음에 보인다.

5-2. 음정 변화 폭에 관한 실험

관악기 중에서 비브라토를 가장 폭 넓게 표현하는 플롯의 비브라토를 분석하였다. 분석에 사용된 음악은 파가니니 24개의 카프리스 작품 1 중 제5곡이고, 여기서 초반의 강한 비브라토 4구간을 추출하였다(A3, C4, F4, A5).

사용한 CD는 도이치 그라마폰, 연주/편곡 - Patrick Gallois (DG 0357)이고, 샘플링 주파수는 44100Hz, 프레임 크기는 1024를 적용하였다.

비브라토의 최대 피치 값 변화는 각각 39, 51, 19 29센트가 나왔다. 처음부터 고음의 평균 피치의 최대 변화는 30센트이다. 따라서 허용해야 할 전체 음정 변화 구간은 -55 ~ 55이다.

보간법을 적용안함	보간법 적용함
1575.00(5808 cents G4)	1600.72(5836 cents G4)
1575.00(5808 cents G4)	1596.04(5831 cents G4)
1575.00(5808 cents G4)	1599.54(5834 cents G4)
1575.00(5808 cents G4)	1592.55(5827 cents G4)

표 3. 보간법을 적용했을 때와 하지 않았을 때의 계산된 피치 또한 비브라토시에 인접 프레임에서의 최대 음정 변화는 35센트이었다. 따라서 35센트 이상의 변화는 음정 변화로 간주된다. 하지만 여기서는 안전하게 40으로 설정하였다.

5-3 음정 추적 실험

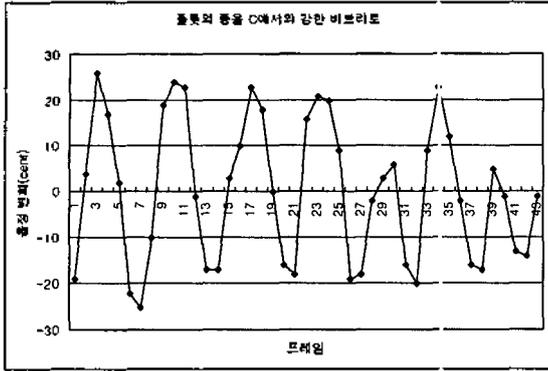


그림 5. 플롯의 중음 C에서의 강한 비브라토로 인한 음정 변화

앞서의 음정 추적 기준을 가지고 플롯으로 녹음한 비제의 '아를르의 여인' 중 미뉴엘 앞부분과 앞서 비브라토 계산에 사용되었던 같은 CD의 제2곡 초반 부분을 자동 음정 추출 실험에 사용하였다.

그 결과 비제의 '아를르의 여인'은 거의 정확하게 음정과 구간이 추출되었고, 제2곡은 빠른 패세지가 슬러없이 연주되어 음과 음 사이에 과도구간이 많이 발생한 것이 요인이 되어 일부 손실이 발생하였다.

추출 결과의 예를 그림 5에 보인다. 추출 결과에서 자동으로 생성된 Praat 레이블 파일을 이용해서 본 그림이다. 또한 미디 파일로 변환하여 미디로 연주된 음악과 원래 음악을 비교하여 평가하기도 하였다.

빠른 곡이 슬러 없이 연주되는 경우에 과도구간에 서 불안정한 점을 보완할 필요가 있다고 생각된다.

5. 결론

본 연구에서는 고음의 피치를 검출할 때 발생하는 오차를 보간법을 적용하여 정확도를 크게 개선할 수 있었고, 이 피치 검출법을 이용하여 음악의 음정 추적 알고리즘을 개발하였고, 추출된 음정과 구간 정보는 미디 파일로 변환되어 비교 평가하여 만족할 만한 결과를 얻었다.

참고 문헌

- [1] Roger J. McNab, Lloyd A. Smith and Ian H. Witten, "Signal Processing for melody transcription", Working Paper 95/22, University of Waikata, New Zealand, 1995
- [2] Panos E. Papamichalis, Practical Approaches to Speech Coding, Prentice Hall, 1987
- [3] L.R. Rabiner, R.W. Schafer, Digital Signal Processing of Speech Signal, Prentice Hall, 1978

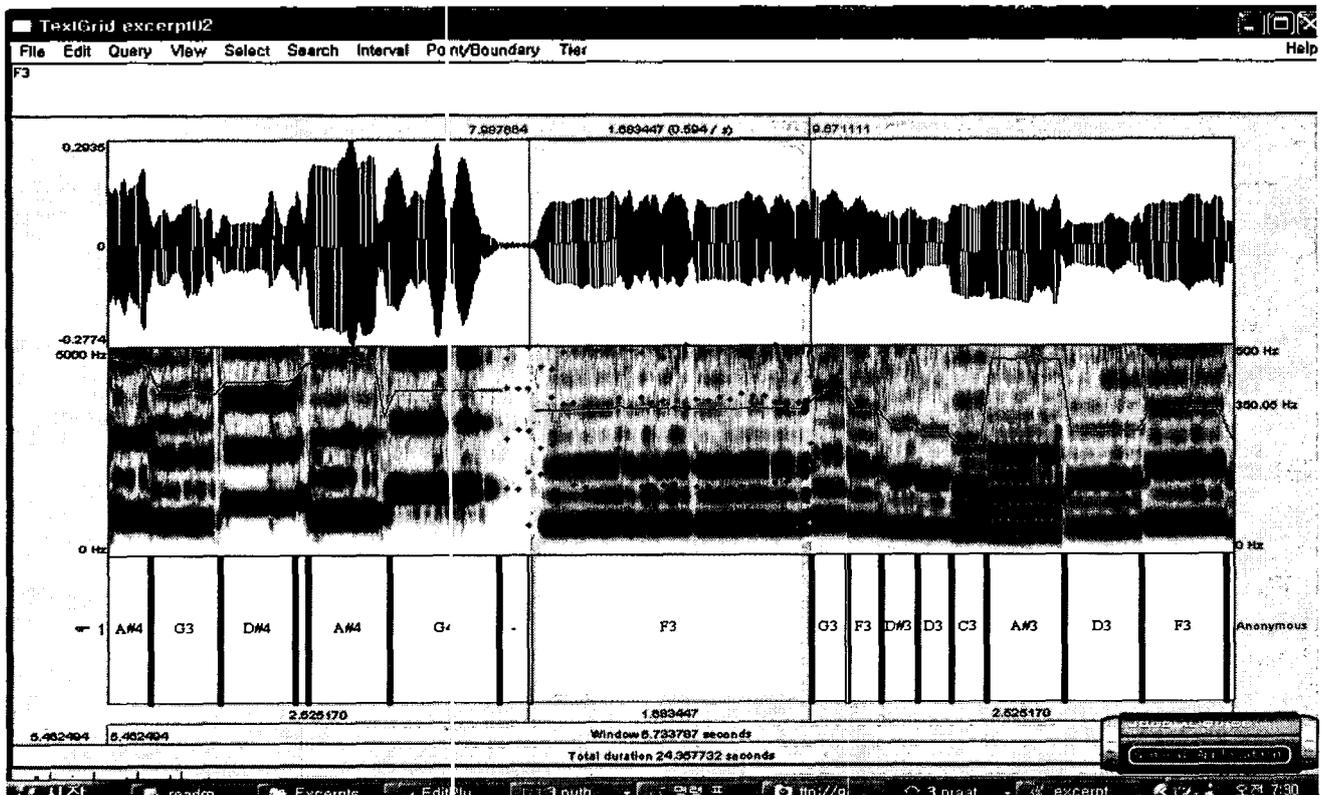


그림 6. 검출된 음정과 구간을 Praat를 이용하여 표시하였다. 레이블 파일은 자동 생성된 것이다