

소음의 피치 분석을 위한 알고리즘

신성환, 이정권

한국과학기술원 기계공학과 소음 및 진동제어 연구센터

The Algorithm for pitch analysis of noise

Shin Sung-Hwan, Ih Jeong-Guon

NOVIC, Department of Mechanical Engineering, KAIST

soul@cais.kaist.ac.kr, ihih@sorak.kaist.ac.kr

Abstract

피치 (pitch)는 주파수와 관련된 인간이 실제 느끼는 음의 높이로 라우드니스 (loudness), 음색 (timbre)과 함께 소리의 음질을 결정하는 중요한 요소로 알려져있다. 이러한 피치는 음성 해석 및 분리를 위해 많은 연구가 이루어진 반면 소음 분석 및 음질 향상을 위한 방향으로의 연구는 부족한 상황이다. 본 연구에서는 기저막 (basilar membrane)의 위치에 따른 주파수 분리 이론인 위치이론 (place theory)을 기본으로 한 기존의 가상 피치 (virtual pitch) 분석 알고리즘을 소음에 적용하기 위해서 수정하고, 여러가지 소음에 적용하였다.

본 연구에서의 알고리즘은 소음의 주파수 특성에 의존한 방법이기 때문에, 톤 (tonal) 성분이 존재하는 소음의 적용에는 적합한 결과를 나타냈지만, 그 이외의 소음에 대해서는 정확한 분석이 어렵다. 따라서 기본 주파수 (fundamental frequency)와 이와 관련된 고조파음 (harmonics)이 음질에 중요한 영향을 미치는 소음의 음질 해석 및 음질 향상을 위해 본 연구의 알고리즘에 의한 피치 분석과 기존의 음질 인자를 적용하면 보다 효율적인 결과를 얻을 수 있을 것이다. 이런 소음의 예로는 엔진의 부밍 소음이나 기어 whine 소음 등이다.

I. 서론

인간은 소리를 인지할 때 소리의 크기, 소리의 지속시간 및 재생 간격, 주파수와 관련된 소리의 높이, 즉 피치에 의해서 소리의 음색을 결정한다^[1]. 이러한

피치는 지금까지 음성 인식 및 분리 등의 음성 해석 분야에서 많은 연구가 이루어진 반면, 제품에서 발생하는 소음 분석 및 그 소음의 음질 향상을 위한 적용에 대해서는 매우 부족한 경향을 나타내고 있다. 그러나 피치 분석은 라우드니스, 샤프니스 (sharpness), 러프니스 (roughness), 변동강도 (fluctuation strength) 등 기존의 인간 청각을 고려한 음질 해석을 위해 사용되었던 여러 음질 요소들과 더불어 소음의 특성을 파악하고, 제어하는데 매우 중요한 영향을 줄 수 있다.

피치 분석은 크게 기저막의 위치에 따라 주파수를 분리하여 인식한다는 위치이론을 기본 개념으로 사용한 방법과 청각 신경 (auditory nerve: AN)을 통해 전달되는 임펄스 자극의 시간적 분포에 따라 주파수를 분리, 인식한다는 주기이론 (temporal theory)을 기본 개념으로 사용한 방법으로 나눌 수 있다. 위치이론을 이용한 피치 분석 방법은 음향신호의 주파수 분석을 통하여 피치 인식에 영향을 주는 모든 주파수 성분을 찾고 각 레벨에 따른 마스킹 현상을 고려할 수 있는 장점이 있는 반면 위상에 의한 영향을 고려하기가 어렵다는 단점이 있다. 일반적으로 음성의 차이를 인지하는데 사용되는 주기이론을 이용한 방법은 피치 분석을 위해서 autocorrelation을 이용하는데, 위상의 영향을 고려할 수 있고 변조음의 적용에 유리하다는 장점은 있지만 마스킹 및 배경 소음을 고려하기 어렵고, 인간이 인지할 수 있는 위상변화 역치 (threshold) 이상으로 위상에 대해 너무 민감한 변화를 가진다는 단점을 갖는다^[2].

본 연구에서는 위치이론을 사용한 피치 분석 방법

중 E. Terhardt에 의해서 제안된 피치 분석 알고리즘^{[3][4]}을 기본으로 사용하고, 소음에 적용하기 위해 E. Zwicker의 임상실험 결과^[5]를 고려하여 알고리즘의 일부를 수정한다. 그리고 수정된 방법을 이용하여 순음 (pure tone), 기본주파수와 고조파음으로 이루어진 complex tone, 진폭변조음 (Amplitude modulated tone) 등의 소음에 적용 후, Zwicker의 임상 실험 결과와 비교한다.

II. 피치 분석 알고리즘의 수정

E. Terhardt의 피치 알고리즘은 음향신호의 주파수 분석을 통한 톤 성분 (tonal component)의 추출, 주파수와 관련된 피치 및 피치 패턴의 분석, 이를 기본으로 한 가상 피치 및 가상 피치 패턴의 분석이 중심이 되는 총 7단계의 과정을 갖는다. 여기서 피치 패턴은 피치 및 피치 세기의 조합을 의미하고, 가상피치란 실제 존재하지는 않지만 인간에게 인지되는 피치를 의미하는 것으로, 완벽한 complex tone에서 기본주파수를 포함한 일부 고조파음이 제거될 경우 그 세기에는 변화가 있지만 원래 기준주파수와 관련된 피치를 인지하는 현상 등이 그 예라 할 수 있다. 그림 1은 가상피치를 시각적인 착시 현상으로 표현한 것이다.

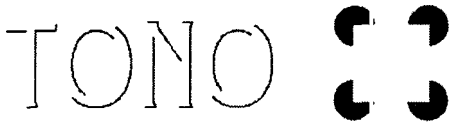


그림 1. 가상피치의 시각적 착시^[6]

그러나 기존의 알고리즘에 사용된 주파수 분석의 주파수 해상도, 주파수나 음향 신호의 음압 레벨과 관련된 가중치 등은 음성 신호를 분석하기 위한 것으로 소음에 대한 적용에 적당하지 못하다. 따라서 기존에 발표된 인간의 청각 특성과 E. Zwicker에 의해서 수행된 임상 실험 결과를 고려하여 가중치에 대한 수정이 필요하다.

먼저 주파수 분석 시에 사용된 주파수 해상도를 살펴보면, 기존 알고리즘에서 사용된 주파수 해상도는 12.5Hz이다. 인간의 주파수 변화에 대한 역치는 주파수 변화가 일어나는 중심주파수의 함수이다. 500Hz 이하의 주파수에서는 약 4Hz의 주파수 역치를 그 이상의 주파수 영역에서는 중심주파수의 0.7%에 해당하는 주파수 변화를 인지할 수 있다^[5]. 이런 사실과 비교해 볼 때 12.5Hz 이상의 주파수 해상도는 약 2kHz 이상의 주파수 영역에 적당한 것으로 2kHz 이하의 주파수

영역에 대한 적용에는 적당하지 않다. 따라서 수정된 알고리즘에서는 5.4Hz의 주파수 해상도를 갖도록 하였다. 그러나 일반적으로 알려진 것처럼 주파수 해상도를 높일 경우 시간에 대한 해상도가 낮아져서 과도음의 분석에 문제가 발생할 수 있는데, 이런 문제점을 극복하기 위해서 인간의 청각 특성을 고려한 VFR-STFT^[7]등의 시간-주파수 스펙트럼 분석 방법을 사용하도록 한다.

주파수 분석을 통한 톤 성분의 추출이 이루어지면 각 톤 성분 사이의 마스킹 (masking) 영향을 고려해야 한다. 이는 인접한 톤 성분들의 영향으로 실제 인지되지 않는 톤 성분을 제거하고, 각 톤 성분의 음압 레벨이 실제 피치 인지에 영향을 주는 정도를 계산한다. 이 과정에서 적용되는 슬로프 가진 (L_{Ev} : slope excitation)에 사용된 마스킹 패턴은 해당 주파수 이상 ($f_{\mu} > f_{\nu}$) 영역에서 (식 1) 과 같이 수정한다.

$$s = \left[\begin{array}{c} -24 - \frac{0.23kHz}{f_{\nu}} + 0.2 \frac{L_{\nu}}{dB} \\ -1.95 \times 10^{-6} \left(\frac{L_{\nu}}{dB} \right)^3 \end{array} \right] \frac{dB}{Bark} \quad (\text{식 1})$$

전체 피치 인지에 대한 각 톤 성분의 주파수 피치 세기는 톤 성분의 음압 초과 정도 (LX_{μ} : SPL excess)와 주파수에 의해서 결정된다.

$$WS_{\mu} = \left[1 - \exp\left(\frac{-LX_{\mu}}{15dB}\right) \right] \left[1 + 0.07 \left(\frac{f_{\mu}}{0.7kHz} - \frac{0.7kHz}{f_{\mu}} \right)^2 \right]^{\frac{1}{2}} \quad (\text{식 2})$$

(식 2)는 음압 초과 정도가 '0'보다 큰 값을 가질 때 ($LX_{\mu} > 0$), 기존 알고리즘에서 주어진 각 톤 성분의 주파수 피치 세기를 나타낸다. 여기서 우변의 첫항은 음압 초과 정도에 대한 영향을 두번째 항은 주파수에 대한 영향을 표현한다. (식 2)의 0.7kHz는 앞에서도 언급했듯이 기존의 알고리즘은 음성 신호에 대한 분석에 그 초점이 맞춰 있기 때문이다. 그러나 E. Zwicker의 임상 실험 결과를 보면, 레벨에 따른 피치 세기는 기존 알고리즘 보다 직선에 가까운 모양을

보이고, 주파수에 따른 순음의 피치 세기는 1kHz ~ 3kHz에서 가장 큰 값을 갖는다^[5]. 본 연구의 알고리즘에서는 이와 같은 임상 실험과의 차이를 보정하기 위해서 (식 3)과 같은 수정된 식을

적용하였다. 그림 2 와 그림 3은 주파수 피치 세기의 각 항에 대해서 기존의 알고리즘, 수정된 알고리즘, Zwicker의 임상 실험 결과를 비교한 것이다.

$$\begin{aligned}
 WS_{\mu} &= 1.12 \cdot \left[1 - \exp\left(\frac{-LX_{\mu}}{30dB}\right) \right] \left[1 + 0.05 \left(\frac{f_{\mu}}{1.5kHz} - \frac{1.5kHz}{f_{\mu}} \right)^2 \right]^{\frac{1}{2}} & f_{\mu} \leq 1500Hz & \text{(식 3)} \\
 &= 1.12 \cdot \left[1 - \exp\left(\frac{-LX_{\mu}}{30dB}\right) \right] & 1500Hz \leq f_{\mu} \leq 2500Hz & \\
 &= 1.12 \cdot \left[1 - \exp\left(\frac{-LX_{\mu}}{30dB}\right) \right] \left[1 + 0.65 \left(\frac{f_{\mu}}{2.5kHz} - \frac{2.5kHz}{f_{\mu}} \right)^2 \right]^{\frac{1}{2}} & f_{\mu} \geq 2500Hz &
 \end{aligned}$$

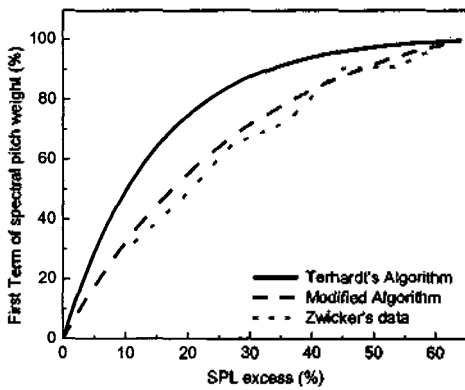


그림 2. 주파수 피치 세기에 대한 레벨의 기여도

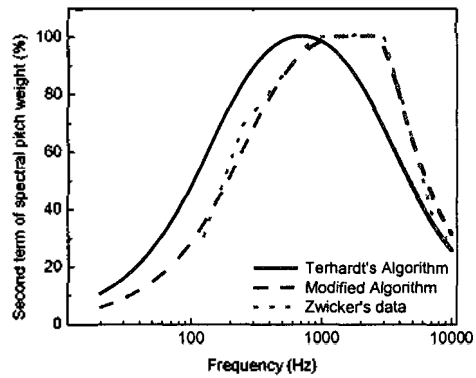


그림 3. 주파수 피치 세기에 대한 주파수의 기여도

III. 수정된 피치 분석 알고리즘의 적용

음질을 고려한 소음 분석을 하기 위해서 가장 먼저 검증되어야 하고 기준을 제시할 수 있는 음향신호는 순음 성분이다. 피치 분석에서도 순음의 피치 패턴, 즉 피치 및 피치 세기는 다른 소음에 대한 기준을 제시할 수 있는 예가 된다. 본 연구에서는 1kHz 순음의 레벨을 변화시키면서 기존의 알고리즘, 수정된 알고리즘에 의한 계산 결과와 Zwicker의 임상 실험 결과를 비교하였다. 가상피치 패턴을 얻기 위한 과정 중 일치성 시험 (coincidence test) 과정에 의해서 주파수 피치 (spectral pitch)와 일치하는 가상피치는 존재하지 않는다. 따라서 순음의 경우는 단지 주파수 피치 패턴만이 존재한다. 그림 4의 결과에서 볼 수 있듯이 각 음압 레벨에서 기존 알고리즘의 계산 결과는 임상 실험의 편차 밖에 위치한 반면 수정된 알고리즘에 의한

계산 결과는 임상 실험의 중앙값 (median value)에 근접함을 확인할 수 있다.

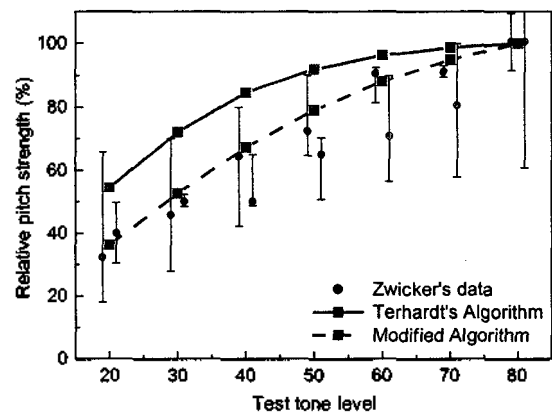


그림 4. 1kHz 순음의 레벨에 따른 피치 세기

다음으로 수정된 피치 분석 알고리즘을 사용하여 125Hz와 관련된 5개 소음에 대해서 주파수 피치 패턴 및 가상 피치 패턴을 분석하였다. No. 1은 125Hz 순음, No. 2는 기본주파수가 포함된 complete complex tone, No. 3은 중심주파수 500Hz이고 변조주파수 125Hz로 100% 변조된 소음, No. 4는 500Hz와 1.5kHz를 차단주파수로 갖는 대역통과 필터를 적용한 주파수 간격 125Hz의 residue complex tone, No. 5는 백색소음 (white noise)에 125Hz를 중심으로 100Hz의 밴드폭을 갖는 대역통과 필터를 적용한 대역소음이다. 표 1은 피치 분석 결과이고, 각 소음에 대해서 피치 세기가 가장 큰 피치를 명시하였다. No. 5 소음을 제외한 나머지 소음에 대해서는 125Hz의 주파수에 대한 피치이동 (pitch shift)이 적용된 피치를 나타내고 있고, No. 1 순음의 가상피치 패턴을 제외하면 모두 주파수 및 가상피치 패턴을 갖는다. 부분적인 피치 세기를 비교하면 Zwicker의 임상 실험 결과와 비슷하지만 정확한 피치 분석을 위해 피치 세기의 대표값을 얻기 위한 연구가 아직은 부족한 상황이다.

IV. 결 론

본 연구는 E. Terhardt의 피치 분석 알고리즘을 소음에 적용하기 위해서 E. Zwicker의 임상 실험 결과를 적용하여 수정하였다. 그 결과 기존의 방법에 비해 순음의 레벨에 따른 피치 분석은 임상 실험 결과와 매우 유사한 경향을 보인다는 것을 확인했다. 그러나 가상피치 패턴과 주파수 피치 패턴 사이의 관계를 얻어 피치 세기에 대한 대표값을 찾는 연구는 지속적으로 이루어져야 한다. 특히 순음의 경우

알고리즘 상 가상피치가 나타나지 않는 다는 것과 밴드 스펙트럼을 갖는 소음의 경우 정확한 피치 인지가 어려운 것이 문제로 남아있다. 그러나 지속적인 연구를 통해서 본 연구의 결과를 엔진의 부밍 소음이나 기어의 whine 소음 등, 토널 성분 특성을 갖는 소음에 적용하면 소음 제어에 필요한 보다 효율적인 정보를 얻을 수 있을 것이다.

참고문헌

- [1] Y. Ando, "A Theory of Primary Sensations and Spatial Sensations Measuring Environmental Noise", *J. Sound and Vibration*, 241(1), pp. 3~18, 2001.
- [2] Ray Meddis and Michael J. Hewitt, "Virtual Pitch and Phase Sensitivity of a Computer Model of the Auditory Periphery. I: Pitch identification", *J. Acoust. Soc. Am.*, 89(6), pp. 2866~2882, 1990.
- [3] E. Terhardt, G. Stoll and M. Seewann, "Algorithm for extraction of pitch and pitch salience from complex tonal signals", *J. Acoust. Soc. Am.*, 71(3), pp. 679~688, 1982.
- [4] E. Terhardt, G. Stoll and M. Seewann, "Pitch of complex signals according to virtual pitch theory: Tests, examples, and predictions", *J. Acoust. Soc. Am.*, 71(3), pp. 671~678, 1982.
- [5] E. Zwicker, H. Fastl, *Psychoacoustics: Fact and Models*, 2nd updated Ed., Springer, New York, 1999.
- [6] E. Terhardt, "Pitch, consonance and harmony", *J. Acoust. Soc. Am.*, 55(5), pp. 1061~1069, 1974.
- [7] H. Jeong and J.-G. Ih, "Implementation of the New Algorithm Using the STFT with Variable Frequency Resolution for the Time-Frequency Auditory Model", *J. Audio Engineering Soc.*, 47(4), 1999.

No.	Noise	Spectral		Virtual		
		Pich	Pitch strength	Pitch	Pitch strength	Pitch strength sum
1	125Hz pure tone	129	0.324	129		
2	Complex tone (-3dB/oct low pass)	129	0.264	124	0.484	2.147
3	AM tone (m=1)	377	0.455	122	0.177	0.342
4	Complex tone (band pass)	498	0.561	120	0.378	0.378
5	Band pass noise (96dB/oct)	123	0.060	38	0.02	0.03

표 1. 피치 분석에 사용된 소음