

피치 알고리즘의 수정 및 소음에의 적용

Modification of Pitch Algorithm and Its Application to Noise

신 성 환*, 이 정 권*,
Sung-Hwan Shin and Jeong-Guon Ih

Key Words : Pitch (피치), Pitch Strength (피치 세기), Place theory (위치 이론), Sound Quality (음질),
Sound Quality Analysis (음질 분석), Sound Quality Metrics (음질 인자)

ABSTRACT

Pitch is a perception related to the subjective frequency that is one of the psychological aspects or attributes of tones. It is also an important factor to determine the sound quality together with loudness and timber. Although the study on pitch has been active in the field of speech communication, but its application to the product sound quality is not yet enough. In this study, the empirical data by Zwicker is made use in the modification of the currently available pitch extraction model based on the place theory. By applying this modified model to various sound samples composed of tonal or banded components, the applicability of the model is suggested. As a demonstration example, the algorithm is used for the sound quality analysis of a product noise having fundamental frequency and harmonics. The result shows that the pitch should be regarded as an important subjective cue in the sound quality analysis.

1. 서 론

음질 분석 (sound quality analysis)을 위해 사용되는 음질인자는 소리의 레벨에만 집중된 기존의 분석 방법 대신 소리에 대한 주관적인 청감을 표현하기 위해 여러 가지 소리의 특성을 나타내는 요인들로 구성된다. 소리의 주관적인 크기로 정의되고 음의 크기면에서 인간의 청감과 선형적인 관계를 갖는 라우드니스 (loudness), 주관적으로 느끼는 소리의 날카로움으로 정의되는 샤프니스 (sharpness), 주관적으로 느끼는 소리의 거칠기로 정의되는 러프니스 (roughness), 진폭 및 주파수 변조에서 기인하는 소리의 주관적 짜증도와 관련된 변동강도 (fluctuation strength)는 지금까지의 연구 결과에서 음질 분석에 영향을 주는 청감으로 밝혀진 것들이다^[1]. 이런 음질인자 이외에 기본 주

파수 (fundamental frequency) 및 이와 관련된 조화음 (harmonics)로 구성된 복합음 (complex tones)의 대표적인 음의 높이와 그 세기를 표현하는 피치 패턴 (pitch pattern), 즉 피치 (pitch)와 피치 강도 (pitch strength) 역시 중요한 음질인자 중의 하나이다.

피치 및 피치 강도는 라우드니스 및 소리의 재생 시간 (subjective duration), 재생 간격 등과 함께 음질 분석을 위해 꾸준히 연구되고 있는 음색 (timber: tone quality)에 영향을 미치는 요인으로 인정되고 있지만^[2], 그 연구가 음성과 관련된 음성 인식이나 음성 분리 등의 분야에 집중된 관계로 일반적인 제품 발생 소음의 분석이나 제어에 적용하기 위해서는 매우 부족한 경향을 보이고 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

* 한국과학기술원 (KAIST)

E-mail : ihih@sorak.kaist.ac.kr, soul@cais.kaist.ac.kr

Tel : (042) 869-5035, Fax : (042) 869-8220

(temporal theory)를 기본 개념으로 사용한 방법으로 나눌 수 있다. 위치이론을 이용한 피치 분석 방법은 음향 신호의 주파수 분석을 이용하여 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 autocorrelation 을 이용하는데, 위상의 영향을 고려할 수 있고 변조음의 적용에 유리하다는 장점은 있지만 마스킹 및 배경 소음을 고려하기 어렵고, 인간이 인지할 수 있는 위상 변화 역치 (threshold) 이상으로 위상에 대해 너무 민감한 변화를 가진다는 단점을 갖는다^[3].

피치 분석을 위한 주기 이론에서의 장점은 피치 이외에 러프니스나 변동강도와 같은 음질 인자에 의해서 고려될 수 있기 때문에, 본 논문에서는 위치이론을 사용한 피치 분석 방법 중 Terhardt 에 의해서 제안된 피치 분석 알고리즘^[4,5]을 기본으로 사용하고, 소음에 적용하기 위해 Zwicker 의 임상실험 결과^[1]를 고려하여 알고리즘의 일부를 수정한다. 또한 수정된 알고리즘의 검증을 위해서 순음 (pure tone), 기본주파수 (fundamental frequency)와 고조파음 (harmonics)로 이루어진 복합음 (complex tone), 진폭 변조음 (amplitude modulated tone) 등의 소음에 적용 후, Zwicker 의 임상 실험 결과와 비교한다. 마지막으로 자동차 가속 주행 시 실내 소음에 적용하여 피치 분석이 음질 분석에 중요한 영향을 미칠 수 있다는 것을 확인한다.

2. 피치 분석 알고리즘의 수정

Terhardt 의 피치 알고리즘은 음향 신호의 주파수 분석을 통한 톤 성분 (tonal component)의 추출, 각 톤 성분의 마스킹 효과 적용, 주파수와 관련된 피치 및 피치 패턴의 분석, 이를 기본으로 한 가상 피치 및 가상 피치 패턴의 분석이 중심이 되는 총 7 단계의 과정을 갖는다. 여기서 피치 패턴은 피치 및 피치 세기의 조합을 의미하고, 가상 피치란 실제 존재하지는 않지만

인간에게 인지되는 피치를 의미하는 것으로, 완벽한 복합음 (complete complex tone)에서 기본 주파수를 포함한 일부 고조파음이 제거될 경우에도 그 세기에는 변화가 있지만 원래 기본 주파수와 관련된 피치를 인지하는 현상 등이 그 예라 할 수 있다.

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 Zwicker 에 의해서 수행된 임상 실험 결과를 고려하여 가중치에 대한 수정이 필요하다.

먼저 톤 성분의 추출을 위해 사용되는 주파수 분석을 위해서 대상 신호가 정상 신호 (steady signal)인 경우 시간에 대한 고려가 필요 없지만, 과도 신호 (transient signal)인 경우에는 시간 - 주파수 해석이 수행되어야 한다. 주파수 분석 시에 사용된 주파수 해상도를 살펴보면, 기존 알고리즘에서 사용된 주파수 해상도는 12.5 Hz 이다. 인간의 주파수 변화에 대한 역치는 주파수 변화가 일어나는 중심 주파수의 함수로 나타나는데, 500 Hz 이하의 주파수에서는 약 4 Hz 의 주파수 역치를 그 이상의 주파수 영역에서는 중심 주파수의 0.7%에 해당하는 주파수 변화를 인지할 수 있다^[1]. 이런 사실과 비교해 볼 때 12.5 Hz 이상의 주파수 해상도는 약 2 kHz 이상의 주파수 영역에 적당한 것으로 2kHz 이하의 주파수 영역에 대한 적용에는 적당하지 않다. 따라서 본 연구에서 적용한 주파수 분석에서는 5.4 Hz 의 주파수 해상도를 갖도록 하였고, 과도음에 대한 분석을 위해서 STFT (short time Fourier transform)을 이용하였다.

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 슬롭 가진 (L_{EV} : slope excitation)을 위한 마스킹 패턴은 해당 주파수

이상 ($f_\mu > f_v$) 영역에서 Aures 에 의해서 제안된 (식 1) 과 같이 수정한다:

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

피치 인지에 대한 각 토널 성분의 주파수 피치 세기 (spectral pitch strength)는 토널 성분의 음압 초과 정도 (LX_μ : SPL excess)와 주파수의 함수로 결정되는데, LX_μ 의 값이 0 보다 클 경우 실제 피치 인지에 영향을 주는 토널 성분을 의미한다:

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

여기서 우변의 첫 항은 음압 초과 정도에 대한 영향을 두번째 항은 주파수에 대한 영향을 표현한다. (식 2)의 0.7 kHz 는 앞에서도 언급했듯이 기존의 알고리즘은 음성 신호에 대한 분석에 그 초점이 맞춰 있기 때문이다. 그러나 Zwicker 의 임상 실험 결과를 보면, 레벨에 따른 피치 세기는 기존 알고리즘 보다 선형적인 관계를 가지며, 주파수에 따른 순음의 피치 세기는 1 ~ 3 kHz 에서 가장 큰 값을 갖는다^[1]. 본 연구에서는 이러한 차이를 보정하기 위해서 (식 3)과 같이 수정된 식을 제안한다. 그림 1 과 그림 2 는 주파수 피치 세기를 나타내는 각 항에 대해서 기존의 알고리즘, 수정된 알고리즘, Zwicker 의 임상 실험 결과를 비교한 것이다. 그림에서 수정된 알고리즘에 의한 결과가 기존 알고리즘의 결과와 비교할 때 임상 실험 결과와 매우 근사함을 확인할 수 있다:

3. 수정된 피치 알고리즘 모델의 검증

음질을 고려한 소음 분석을 하기 위해서 가장 먼저 검증되어야 하고 기준을 제시할 수 있는 음향 신호는 순음 성분이다. 피치 분석에서도 순음의 피치 패턴, 즉 피치 및 피치 세기는 다른 소음에 대한 기준을 제시할 수 있는 예가 된다. 본 연구에서는 1 kHz 순음의 레벨을 변화시키면서 기존의 알고리즘, 수정된 알고리즘에 의한 계산 결과와 Zwicker 의 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 3 의 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.

다음으로 수정된 피치 분석 알고리즘을 사용하여 125 Hz 와 관련된 5 개 소음에 대해서 주파수 피치 패턴 및 가상 피치 패턴을 분석하였다. No. 1 은 125 Hz 순음, No. 2 는 기본주파수가 포함된 complete complex tone, No. 3 은 중심주파수 500 Hz 이고, 변조주파수 125 Hz 로 100% 변조된 소음, No. 4는 500 Hz 와 1.5 kHz 를 차단주파수로 갖는 대역통과 필터를 적용한 주파수 간격 125 Hz 의 residue complex tone, No. 5는 백색잡음 (white noise)에 125 Hz 를 중심으로 100 Hz 의 밴드폭을 갖는 대역통과 필터를 적용한 대역 소음이다.

$$WS_\mu = 1.12 \cdot \left[1 - \exp\left(\frac{-LX_\mu}{30\text{dB}}\right) \right] \cdot \left[1 + 0.05 \left(\frac{f_\mu - 1.5\text{kHz}}{f_\mu} \right)^2 \right]^{\frac{1}{2}} \quad f_u \leq 1500\text{Hz} \quad (\text{식 3})$$

$$= 1.12 \cdot \left[1 - \exp\left(\frac{-LX_\mu}{30\text{dB}}\right) \right] \quad 1500\text{Hz} \leq f_u \leq 2500\text{Hz}$$

$$= 1.12 \cdot \left[1 - \exp\left(\frac{-LX_\mu}{30\text{dB}}\right) \right] \cdot \left[1 + 0.65 \left(\frac{f_\mu - 2.5\text{kHz}}{f_\mu} \right)^2 \right]^{\frac{1}{2}} \quad f_u \geq 2500\text{Hz}$$

표 1 은 각 소음에 대한 피치 분석 결과이다. 여기서, 인지된 피치가 정확히 125 Hz 와 일치하지 않는 이유는 인간의 청각이 피치를 인지할 때 마스킹 효과 및 소리의 크기에 의존하여 피치이동 (pitch shift)이 나타나고 이를 고려하는 관계식이 알고리즘 상에 포함되어 있기 때문이다. 순음에 대한 피치이동을 확인하기 위해 $\frac{\Delta f}{f}$ 를 계산하면 0.032 로 피치이동과 관련된 임상실험 결과^[6]와 거의 일치한다. 위에서도 언급했듯이 No. 1 순음의 가상피치 패턴을 제외하면 모두 주파수 및 가상피치 패턴을 갖고, No. 5 소음을 제외한 나머지 소음에 대해서는 125 Hz 와 관련된 피치가 인지되었다. No. 5 소음에 대해서 정확한 피치 분석이 이루어 지지 않는 이유는 본 연구에서 피치 분석을 위해 사용된 모델이 주파수 관점에서 접근하고, 토널 성분을 기본으로 사용 하기 때문에, 대역 소음에 대한 피치 추출에 적당하지 못하기 때문이다. 따라서 본 연구에서 제안된 알고리즘의 적용 범위는 토널 성분이 존재하는 소음에 대한 피치 추출로 한정한다.

4. 소음에 대한 피치 분석의 적용

자동차 실내 소음에 영향을 주는 주요한 소음원 중 엔진의 부밍 현상은 일반적으로 엔진의 점화주파수 (firing frequency)와 관련된 고조파음 (harmonics)으로 구성된다. 이러한 엔진의 부밍 현상을 분석하기 위한 기존의 방법으로는 시간 (또는 회전수) - 주파수 분석을 통한 오더 해석 (order analysis)이 일반적이지만, 이러한 해석 결과만으로는 엔진의 부밍 현상에 영향을 주는 주파수나 시간 영역을 정확히 파악할 수 없고, 실제 인간이 느끼는 부밍의 정도를 예측하기에는 부족한 점이 있다. 그 이유는 기존의 방법들이 인간의 청각 특성을 전혀 고려하지 않은 물리적인 측정 방법에 의존하기 때문이다.

따라서 본 연구에서는 수정된 피치 분석 알고리즘을 자동차의 가속 주행 소음 중 엔진의 부밍 (booming) 현상을 분석하기 위해 적용하였다.

이는 위에서도 언급했듯이 부밍 현상은 기본 주파수와 그와 관련된 고조파음으로 구성되기 때문에 기존의 다른 음질 인자보다 피치 분석이 가장 이상적인 음질 분석 도구로 사용될 수 있기 때문이다. 대상 소음은 4기통 엔진 및 자동 변속기를 가진 자동차를 대상으로 성능 시험장 (proving ground)에서 2 단 가속 주행 중 차량 실내 소음을 측정 한 것이다.

차량의 회전수 (rpm) - 주파수 해석 결과를 나타내는 그림 4 와는 달리, 본 연구에서 수정된 피치 분석을 적용한 그림 5 에서는 실제 인간의 청각에 영향을 주는 주요 성분만이 나타나고 있다. 그림 5 를 보면 4000 rpm 까지는 회전수에 대한 1 차 점화 주파수 성분이, 4000 rpm 에서 5500 rpm 까지는 회전수에 대한 2 차 점화 주파수 성분이 전체 음질에 주요한 영향을 주고 있음을 알 수 있다. 이는 피치 분석 단계에서 마스킹 효과 (masking effect)를 고려하고 임상 실험을 기본으로 각 토널 성분에 가중치를 적용했기 때문이다. 이러한 피치 분석 결과를 보면 기본 주파수 및 그와 관련된 조화음을 갖는 소음에서 기본 주파수 성분을 찾는 것이 가능하고, 특히 전체 소음의 음질에 영향을 주는 주요 주파수 성분을 분석하는 것이 가능하다. 결과적으로 피치 분석이 점화 주파수 (firing frequency)와 관련된 주파수 성분들에 의해서 나타나는 엔진 소음의 부밍 현상을 분석하고 주관적인 음감을 표현하기에 적당한 방법이 될 수 있음을 의미한다.

5. 결론

본 연구는 Terhardt 의 피치 분석 알고리즘을 소음에 적용하기 위해서 Zwicker 의 임상 실험 결과를 적용하여 그 일부를 수정하였다. 그 결과 기존의 방법에 비해 수정된 알고리즘을 적용하면 순음의 레벨에 따른 피치 분석과 주파수에 따른 피치 분석 결과가 임상 실험 결과와 매우 유사한 경향을 보인다는 것을 확인했다. 그리고 125 Hz 와 관련된 5 가지의 소음에 적용하여 현 알고리즘이 토널 성분에 의해 나타나는 피치 패턴 분석에는 용이하지만, 대역 소음에 의해서

나타나는 피치 분석에는 적용에 문제가 있음을 지적하고, 그 적용 범위를 전자의 경우로 한정하였다. 이를 바탕으로 자동차 실내 소음에서 엔진의 부밍 현상의 분석 및 부밍의 주관적 평가를 위하여 피치 분석이 적절한 음질 분석 도구로 사용될 수 있음을 확인하였다.

그러나 가상피치 패턴과 주파수 피치 패턴 사이의 관계를 얻어 피치 세기에 대한 대표값을 구하거나, 레벨에 따라 달라지는 마스킹 패턴의 적용 등에 대해서는 지속적인 연구가 이루어져야 할 것이다. 또한 현 알고리즘에서 밴드 스펙트럼을 갖는 소음의 경우 정확한 피치 인지가 어려운 것이 문제로 남아있다.

참고문헌

[1] E. Zwicker and H. Fastl, *Psychoacoustics, Facts and Models*, 2nd updated Ed., Springer, New York, 1999.

[2] Y. Ando, "A Theory of Primary Sensations and Spatial Sensations Measuring Environmental Noise", *J. Sound and Vibration*, vol. 241(1), pp. 3~18, 2001.
 [3] 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.
 [4] 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.
 [5] 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.
 [6] S. S. Stevens and H. Davis, *Hearing: Its Psychology and Physiology*, 2nd printing, American Institute of Physics, Inc., New York, 1983.

No.	Noise	Spectral		Virtual		
		Pitch	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

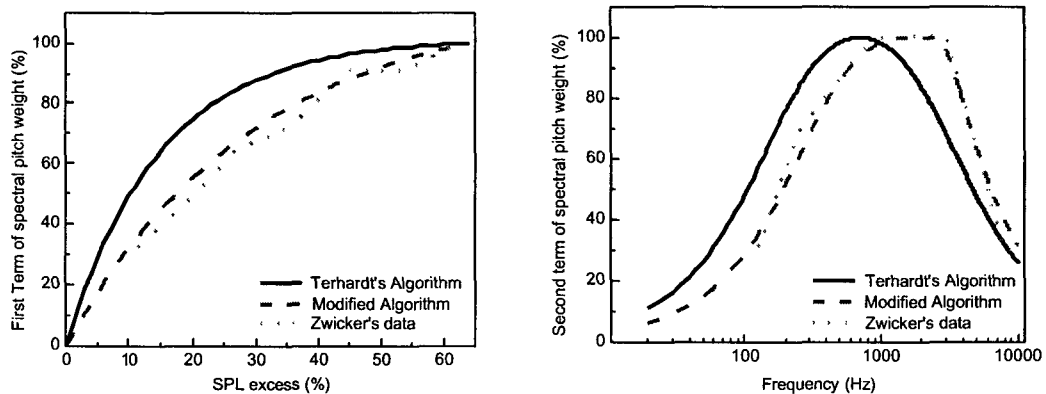


그림 1. 주파수 피치 세기에 대한 레벨의 기여도. 그림 2. 주파수 피치 세기에 대한 주파수의 기여도.

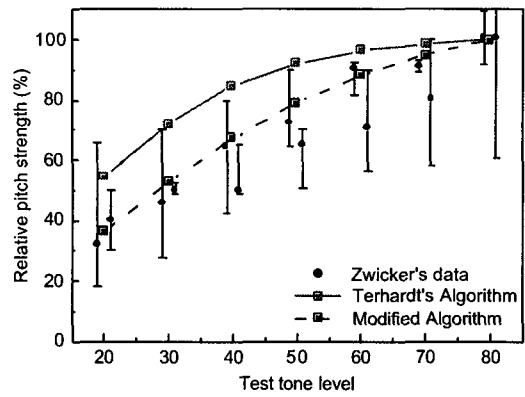


그림 3. 1kHz 순음의 레벨에 따른 피치 세기

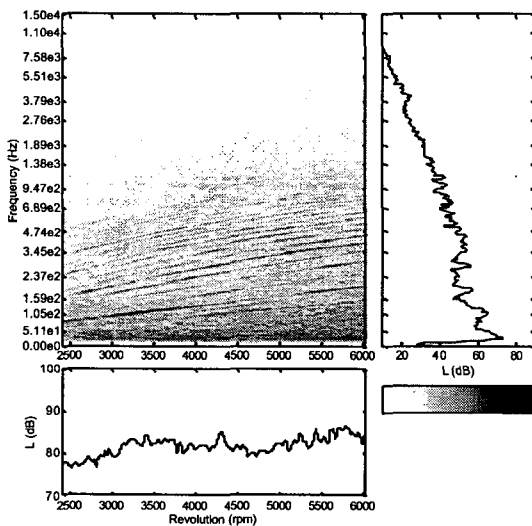


그림 4. 자동차 실내 소음의 시간 주파수 분석.

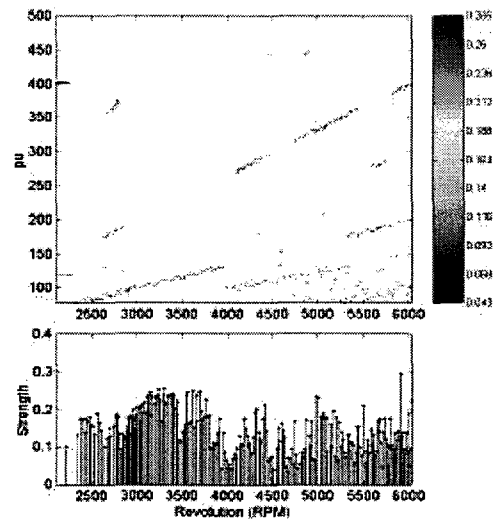


그림 5. 자동차 실내 소음의 피치 분석