

객관적으로 정량화된 복합 신호음의 조화도

Objectively Quantified Consonance of Complex Sounds

전 상 배*, 최 인 용*, 이 민 구*, 성 평 모*

(Sang Bae Chon*, In Yong Choi*, Mingu Lee*, Koeng-Mo Sung*)

*서울대학교 전기·컴퓨터공학부 뉴미디어통신공동연구소

(접수일자: 2007년 7월 31일, 채택일자: 2007년 8월 17일)

본 논문에서는 객관적으로 정량화된 복합 신호음의 조화도를 새로운 심리 음향 파라미터로 제안하였다. 사람이 복합 신호음을 인지하는 과정에서 발생하는 masking 효과, 등청감 곡선, criticalband 등의 심리음향학적인 모델들을 적용시킨 후, 그 소리가 어느정도의 조화도를 갖는지를 정량화시키는 모델을 제안하였고, 이를 검증하기 위하여 복합 음으로 구성된 자동차 Horn 신호로 실험을 수행하였다. 실험 결과, 제안한 알고리즘에 의하여 정량화된 객관적인 조화도와 청취 평가를 통한 주관적인 조화도가 0.95의 cross correlation을 보였다. 현재 객관적인 심리음향학적인 파라미터로는 Zwicker 파라미터 외에는 거의 전무한 상태에서 제안하는 알고리즘을 통하여 소리의 심리음향학적인 효과를 보다 객관적으로 파악하는 데에 도움이 될 것으로 보인다.

핵심용어: 조화도, 심리 음향 파라미터

투고분야: 음악 음향 및 심리분야 (8)

In this paper, objectively quantified consonance of complex sound is proposed as a new psychoacoustical parameter. Proposing algorithm quantifies consonance of complex sound after applying psychoacoustical models which are parts of human perception such as masking effect, equal loudness contour, and critical band. To verify proposing algorithm, experiments with 10 car horn signals which have different complex sound were performed. The experiments show cross correlation of 0.95 between objectively quantified consonance by proposing algorithm and subjectively assessed consonance by listening tests. Considering the fact that there are few psychoacoustical parameter except Zwicker parameter, proposing algorithm will help to quantify psychoacoustical effect of complex sounds objectively.

Key words: Consonance, Psychoacoustical parameter

ASK subject classification: Musical Acoustics and Psychoacoustics (8)

I. 서론

신호음은 그 제품과 사용자를 연결하는 수단으로 많은 전자제품 및 기계제품에서 사용되어왔다. 제품의 종류나 그 기능에 따라 신호음은, 현재 상태에 대한 표현, 일련의 과정의 종료, 외부로부터의 호출, 위험한 상황에서의 경고, 발생 가능한 재난에 대한 경보 등을 사용자에게 신속하게 전달하는 역할을 하였다. 사용자의 입장에서 그 음에 대한 학습에 의하여 그 음이 의미하는 바를 결정하기도 하지만, 그 신호음이 적절히 의미와 조화로운 때 더욱 그 의미의 직접적인 전달 효과는 더욱 커진다.

그러나, 현재 라우드니스 (Loudness), 샤프니스 (Sharpness), 러프니스 (Roughness), 플럭추에이션 스트레NGTH (Fluctuation Strength) 등의 쯔비커 파라미터 (Zwicker Parameter) [10]를 제외하면, 신호음에 대한 심리음향학적인 파라미터들은 거의 전무한 실정이다. 따라서 어떠한 신호음 설계나 검증의 과정에서 일정한 지표 없이 임의로 만들어지는 상황이다.

음의 조화도 [2,3,7]에 있어서는 두 단일음 (pure tone)에 대하여서만 그 정량화가 이루어져 있고, 실질적으로 많이 사용되는 둘 이상의 단일음으로 이루어진 복합음 (complex tone)에 대하여서는 정량화에 대한 연구가 진행되어있지 않았다. 이에, 본 논문에서는 신호음에 대한 평가 파라미터로 쓰일 수 있을 정량화된 조화도 (Quantified Consonance)에 대하여 제안하고자 한다.

책임저자: 전 상 배 (strlen@acoustics.snu.ac.kr)
151-744 서울특별시 관악구 신림동 산 56-1번지 301동 915호
서울대학교 전기·컴퓨터공학부 음향공학연구소
(전화: 02-880-7263)

II. 조화도 이론 (Consonance Theories)

조화도 (consonance)는 음악의 분석 과정 중, 화음에 대한 연구에 의하여 정의되어왔다. Pythagoras는 두 단 음에 있어서, 근사화된 두 주파수의 비가 1:1, 1:2, 2:3, 3:4 등과 같이 낮은 정수의 비로 나타낼 수 있을 때, 그 음이 더 조화롭다고 [2,7] 하였다. Helmholtz는 두 음의 배음 성분들간의 간섭 (interference) 현상으로부터 발생하는 비트 (Beat) 현상에 의하여 부조화 (Dissonance)가 발생한다는 논리에 근거하여, 배음성분들이 일치하는 경향이 많은 두 음의 조합의 경우에 간섭현상 (Interference)가 적어 더욱 조화롭게 인자된다 [2,7]고 해석하였다.

이후, 1960년대에 Plomp와 Levelt에 의하여 단일 음들의 조화도에 대하여 연구 [7]가 이루어졌는데, 조화도는 주파수의 비가 아닌, 두 음의 주파수 간격 (혹은 차이)에 의하여 결정된다고 해석되었고 그 간격에 따른 조화도가 구체적으로 수치화 되었다.

그림 1은 Plomp와 Levelt의 실험 결과인 주파수 간격 (Frequency Interval)과 조화도와 관계를 보여준다. 한 크리티컬 밴드 (critical band)안에 두 성분의 음들이 존재할 경우, 그 간격 (interval)과 크리티컬 밴드 폭의

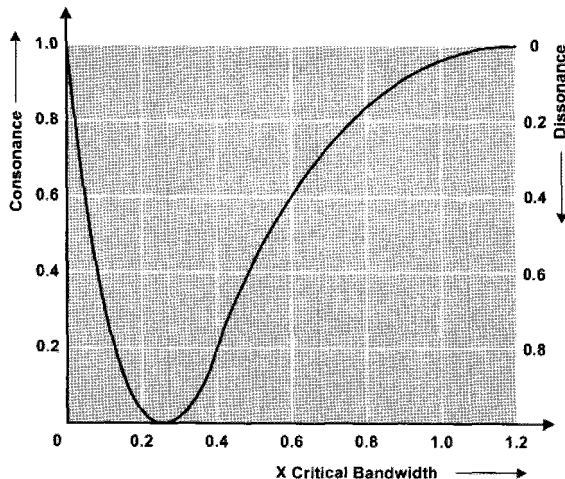


그림 1. 주파수 차이와 조화도와 관계
Fig. 1. Frequency Interval and Consonance.

- 두 음이 같은 주파수일 경우 (unison), 두 음은 완전히 조화롭다고 한다.
- 두 음이 하나의 Critical Band안에 위치하지 않는 경우, 두 음은 조화롭다고 판단한다.
- 두 음의 주파수 차이가 Critical Band폭의 5~50%인 경우, 두 음은 부조화스럽다고 판단한다.
- 가장 부조화스러운 경우의 두 음의 주파수 차이는 Critical Band폭의 약 4분의 1인 경우이다.

상대적인 비로 그 조화도 또는 부조화도가 정의된다. 이러한 Plomp와 Levelt의 크리티컬 밴드에 근거한 조화도 이론은 다음과 같이 결론지어진다.

III. 복합 신호음의 조화도 (Consonance of Complex Sound Signal)

3.1. 개요

2장에서 살펴본 바와 같이 Plomp와 Levelt에 의하여 두 음의 조화도 (consonance)가 정의되었다. 본 논문에서는 이러한 조화도 이론을 현재 다양한 기계, 전기, 전자 제품에서 사용되는 신호음에 적용시켜 각각의 신호음의 심리음향학적인 파라미터로서의 사용을 제안하고자 한다.

3.2. 복합 신호음의 조화도 정량화 알고리즘

그림 2는 제안하는 복합 신호음의 조화도 정량화의 블록 다이어그램이다. 본 알고리즘에서는 등청감 곡선 [8], 마스킹 (Masking) 효과 [1,5,8], 크리티컬 밴드 [3,8,9] 등의 심리 음향 모델이 사용되었다.

3.2.1. T/F 변환

음원으로부터 조화도를 산출하기 위해서 시간 도메인 (Time Domain)보다 주파수 도메인 (Frequency Domain) 상에서 분석이 용이하므로, 주파수 영역으로 변환한다.

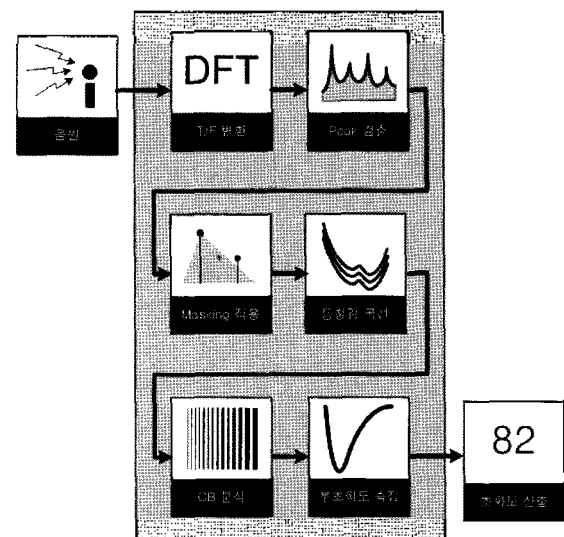


그림 2. 복합 신호음의 조화도 정량화 흐름도
Fig. 2. Consonance Quantification Procedure.

실험에서는 가장 일반적인 변환인 이산 푸리에 변환 (Discrete Time Fourier Transform)이 사용되었다.

3.2.2. peak 검출

주파수 도메인상에서 주로 들리는 음들을 분리해내기 위하여, 주파수 도메인에서의 peak들을 검출한다.

3.2.3. Masking 적용

주파수 도메인 상에서 에너지가 충분히 많은 피크 (peak)의 경우에도, 마스킹 현상에 의하여 인접한 주파수의 성분에 따라 마스크 (maskee)가 되어 들리지 않는 경우도 있으므로 이러한 유효하지 않은 성분을 제거하기 위하여 마스킹 현상을 적용시켜 마스크가 되는 피크를 제거한다.

3.2.4. 등청감곡선 적용

물리적으로 같은 에너지를 갖는 신호에 대해서, 사람의 청각기관 특성상 그 크기를 다르게 인지하는데, 이러한 특성을 보정하기 위하여 등청감곡선 (Equal Loudness Contour) [8]을 적용시켜 사람이 인지하는 크기로 변환시킨다.

3.2.5. 크리티컬 밴드 분석

위의 과정들을 통하여 사람의 청각기관 특성이 반영된 신호들을, 모든 피크 성분들끼리 Plomp와 Levelt에 의하여 정의된 조화도 이론을 통하여 부조화도를 측정하기 전에, 각각의 피크 쌍들이 동일한 크리티컬 밴드안에 위치하는지의 여부를 판별한다. 그림 1과 같이 Plomp와 Levelt의 조화도 이론이 크리티컬 밴드의 1.2배가 되는 주파수 간격을 갖는 신호에 대하여 정의되어 있으므로, 크리티컬 밴드의 1.2배 이내에 존재하는 피크들을 쌍 (pair)으로 구분지어 놓는다. 크리티컬 밴드의 폭은 한쌍의 peak들로부터 그 중심 주파수 (f_c)를 구하고 이를 다음과 같이 정의되는 ERB (Equivalent Rectangular Bandwidth) 모델을 적용시켜 구한다. [6]

$$CBW = [6.23 \times 10^{-6} \times f_c^2] + [93.39 \times 10^{-3} \times f_c] + 28.52 [\text{Hz}]$$

3.2.6. 부조화도 (Dissonance) 측정

크리티컬 밴드 분석 이후, 각각의 피크의 쌍의 주파수 간격과 크리티컬 밴드 폭을 그림 1의 Plomp와 Levelt의 조화도 이론을 통하여 부조화도를 구한다.

3.2.7. 조화도 (Consonance) 산출

심리음향 모델을 거쳐 유효한 피크 쌍들의 부조화도를 모두 측정한 후, 이로부터 최종적인 조화도를 산출해낸다. 이러한 조화도 산출에서는 두가지의 가중치가 적용된다.

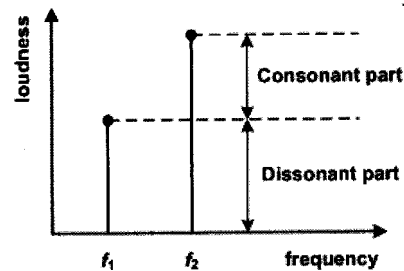


그림 3. 주파수 차이와 조화도의 관계

Fig. 3. Frequency interval and Consonance.

첫째, 부조화도가 계산된 피크의 쌍들의 경우, 그 라우드니스가 동일하지 않기 때문에, 계산된 부조화도에 이를 고려한 보정이 필요하다. 그림 3은 주파수 f_1 과 f_2 성분의 두 피크에 있어서 그 라우드니스가 다른 경우에 대한 예시이다. 이러한 경우의 부조화도는 라우드니스가 상대적으로 적은 f_1 성분의 피크에 의하여 결정되므로, 이에 f_1 성분의 피크의 라우드니스를 곱한다.

둘째, 부조화도는 음원 전체의 라우드니스의 함수가 아니므로, 음원의 전체 에너지로 측정된 부조화도를 normalize해야 한다.

이러한 두가지의 가중치를 적용시키면 총부조화도 (D)는 다음과 같이 표현할 수 있다.

$$D = \frac{1}{LOUD} \sum_{all\ i, j} \min(loud(i), loud(j)) \times d(i, j)$$

where $loud(i) = \text{loudness of } i^{th} \text{ peak}$

$d(i, j) = \text{dissonance between } i^{th} \text{ and } j^{th} \text{ peaks}$

$$LOUD = \sum loud(i)$$

최종적으로 산출된 부조화도 (D)로부터 총조화도 (C)를 $C = 1 - D$ 의 관계로 산출한다.

IV. 조화도 확인 실험 (Verification Test)

4.1. 실험 개요

제안하는 알고리즘을 확인하고자 제안하는 알고리즘을 통하여 객관적으로 조화도를 계산하였고, 청취평가를 통하여 조화도를 평가하여 각각의 결과를 비교해 보았다. 사용된 음원으로는 10가지의 각기 다른 자동차 경적 신호음이 사용되었다.

4.2. 제안한 알고리즘에 의한 조화도

제안한 알고리즘은 표 1과 같이 구현되었고, 10개의 자동차 경적 신호음에 대한 결과는 표 2와 같다.

표 1. 구현 세부 사항
Table 1. Simulation Details.

Block	세부 사항
T/F 변환	4096 Point FFT Hamming Window
Masking 적용	MPEG Psychoacoustic Model 2 (1,5)
등청감 곡선	Equal Loudness Contour (4) (ISO 226)
Critical Band Width	ERB Model (6)
부조화도 측정	Plomp and Levelt Model (7)

표 2. 신호음의 조화도 객관적 평가 결과
Table 2. Objective Assessment Result.

신호음	총조화도
horn 1	0.642
horn 2	0.618
horn 3	0.642
horn 4	0.607
horn 5	0.614
horn 6	0.573
horn 7	0.604
horn 8	0.578
horn 9	0.648
horn 10	0.641

표 3. 신호음의 조화도 주관적 평가 결과
Table 3. Subjective Assessment Result.

신호음	총조화도
horn 1	0.88
horn 2	0.42
horn 3	0.58
horn 4	0.43
horn 5	0.52
horn 6	0.08
horn 7	0.45
horn 8	0.10
horn 9	0.95
horn 10	0.84

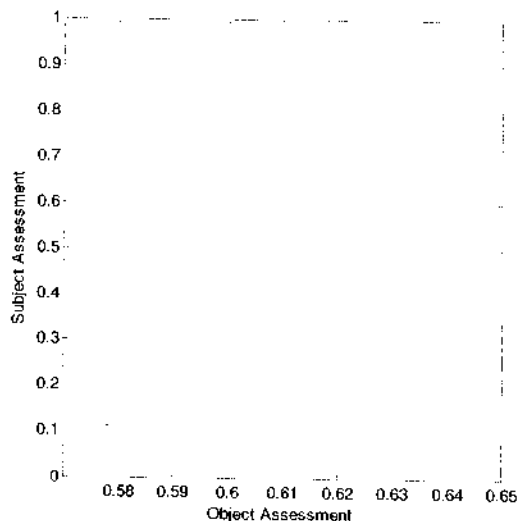


그림 4. 객관평가와 주관평가 관계
Fig. 4. Object Assessment vs. Subject Assessment.

4.3. 청취평가에 의한 조화도

주어진 10개의 음원들에 대하여 주관적으로 어느정도 조화도를 보이고 있는지 확인하기 위하여 기존에 다른 청취평가에 경험이 있는 10명을 대상으로 청취평가를 실시하였다.

청취평가에서는 보다 나은 결과를 얻기 위하여 Plomp와 Levelt 이론에 근거하여 임의로 만든, 조화도가 0에서 0.1 간격으로 1까지 11개의 음원을 통하여 트레이닝과정을 포함시켰다. 트레이닝 데이터는 청취 평가 도중에도 들을 수 있도록 하였고, 각각의 청취 평가 음원들을 0~1 사이 0.1 간격의 10개의 등급 중 하나를 주도록 하였다. 청취 평가 결과는 표 3과 같이 나타났다.

4.4. 실험 결과

그림 4는 주어진 10개의 음원들에 대하여 제안하는 알고리즘을 통한 객관 평가 결과와 주관 평가 결과의 관계를 보여준다. 그림에서 알 수 있듯이 주관평가와 객관평가는 0.95의 상당히 높은 상관도를 보였다. 따라서, 제안하는 알고리즘인 복합 신호음의 조화도는 심리음향학적인 지표로 사용될 수 있을 것으로 보인다.

V. 결론 (Conclusion)

본 논문에서는 복합 신호음들의 조화도를 객관적으로 측정할 수 있는 알고리즘을 제안하였다. 제안한 알고리즘은 masking 현상, 등청감곡선 등의 심리음향 모델들을 사용하여 각각의 부조화를 야기하는 성분들에 가중치를 부가하였고, 단일음에서의 Plomp와 Levelt의 조화도 모델을 사용하였다.

제안하는 알고리즘의 검증을 위하여 구현한 알고리즘을 통한 객관 평가와 청취 평가를 통한 주관 평가와의 관계를 살펴보았다. 보다 정확한 평가를 위하여 객관 평가를 위한 알고리즘 구현시에는 MPEG의 마스킹 모델, ISO의 등청감곡선 등 표준 데이터를 사용하였고, 주관평가를 위해서 조화도의 등급에 따른 참조 음원들을 사용하였다. 그 결과 제안한 알고리즘에 의한 객관 평가 결과와 청취 평가에 의한 주관 평가 결과가 0.95의 상관도를 갖는 것으로 확인 되었고, 이로서 제안하는 복합 신호음의 객관적으로 정량화된 조화도는 심리 음향학적인 지표로 사용 가능할 것으로 보인다.

참 고 문 헌

1. Bosi, Marina and Goldberg, Richard E., *Introduction to Digital Audio Coding and Standards*, (Kluwer Academic Publishers, 2003)
2. Burns, Edward M., *The Psychology of Music*, (San Diego: Academic Press, 1999)
3. Howard, D.M. and Angus, J., *Acoustics and Psychoacoustics*, (Forcal Press, 1996)
4. ISO 226:2003, Acoustics "Normal Equal-Loudness-Level Contours," ISO, 2003
5. ISO/IEC 11172, Information Technology, "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s," 1993
6. Moore, B.C.J. and Glasberg, B.P., "Suggested Formulae for Calculating Auditory-Filter Bandwidths and Excitation Patterns," *Journal of the Acoustical Society of America*, 74 (3), 750-753, 1983.
7. Plomp, R. and Levelt, W.J.M., "Tonal consonance and critical bandwidth," *Journal of the Acoustical Society of America*, 38, 548, 1965.
8. Rossing, T.D., Moore, F.R. and Wheeler, P.A., *The Science of Sound*, 3rd Ed., (Addison Wesley, 2002)
9. Scharf, B., *Foundation of Modern Auditory Theory*, (London: Academic Press, 1970)
10. Zwicker, E. and Fastl, H., *Psychoacoustics : Facts and Model*, (Springer-Verlag Berlin Heidelberg, 2000)

저자 약력

• 전 상 배 (Sang Bae Chon)

2000년: 전북대학교 전자공학과 공학사
 2005년: 서울대학교 전기컴퓨터공학부 공학석사
 2005년~현재: 서울대학교 전기·컴퓨터공학부 박사과정

• 최 인 용 (In Yong Choi)

2001년: 서울대학교 전기공학부 공학사
 2003년: 서울대학교 전기컴퓨터공학부 공학석사
 2003년~현재: 서울대학교 전기·컴퓨터공학부 박사과정

• 이 민 구 (Mingu Lee)

2004년: 서울대학교 물리학과 이학사
 2004년~현재: 서울대학교 전기·컴퓨터공학부 석박통합과정

• 성 평 모 (Koeng-Mo Sung)

1965~71: 서울대학교 전자공학과
 1971~73: 독일 아헨공대 Vordiplom
 1973~77: 독일 아헨공대 전자통신공학 Dipl.-Ing.
 1977~82: 독일 아헨공대 음향공학 Dr.-Ing. (공학박사)
 1977~83: 독일 아헨공대 음향공학연구소 연구원
 1983~현재: 서울대학교 전기·컴퓨터공학부 교수