

# 오디오/음성 컬러 처리 기술 연구

김 광 기, 김 상 진, 손 백 권, 한 민 수  
한국정보통신대학원대학교

## A Study on Audio/Voice Color Processing Technique

Kwangki Kim, Sang-Jin Kim, BeakKwon Son, Minsoo Hahn  
Information and Communications University  
E-mail : {kkkim, sangjin, neverland, mshahn}@icu.ac.kr

### Abstract

In this paper, we studied advanced audio/ voice information processing techniques, and trying to introduce more human friendly audio/voice. It is just in the beginning stage. Firstly, we approached in well-known time-domain methods such as moving average, differentiation, interpolation, and decimation. Moreover, some variation of them and envelope contour modification are utilized. We also suggested the MOS test to evaluate subjective listening factors. In the long term viewpoint, user's preference, mood, and environmental conditions will be considered and according to them, we hope our future technique can adapt speech and audio signals automatically.

### I. 서 론

미래 음성 및 음향 정보처리 기술은, 이미 저장되어 있던 내용을 단순히 녹음, 압축 및 복원, 전송, 재생하여 수동적으로 듣는 것이 아니라, 사용자의 기호, 감정 상태 등과 사용 환경에 능동적으로 대처하는 적응형 기술로 진화할 것이다. 그러나 이에 대한 연구는 심리 음향 학자들에 의해 정의된 청감인자들을 모델링하는 것에 그치고 있다.

본 논문에서는 적응형 음성 및 음향 정보처리 기술의 원천 기술인 음향정보의 음색변환 기술로, 음성 및 음향신호에서 사물의 질감을 느낄 수 있도록 하는 음색의 변환을 연구하였다. 이를 위하여 먼저 시간영역

에서 다양한 신호처리 방법들을 적용하여 보았으며, 이렇게 변환된 신호에 대해 변화도를 측정할 수 있는, 음의 심리적 질감의 정량화에 대하여 간략히 소개하였다.

2절에서 다양한 음의 질감을 부여할 수 있는 시간 영역에서의 신호처리 방법들을 설명하고, 3절에서는 청감인자들과 질감의 정량화에 대해서 설명한다. 4절에서 피아노 연주곡 및 음성을 이용한 실험 및 결과를 설명하며, 5절에서 결론을 맺었다.

### II. 시간 영역의 신호처리 방법들

시간 영역에서의 대표적인 신호처리 방법에는 이동평균처리와 differentiation이 있으며, 이를 응용하여 이동편차강조, 이동편차교차를 실험하여 보았다. 또한, envelope contour 조절, 재생속도조절 등을 적용하여 보았다. 이러한 방법들을 오디오 및 음성 신호에 적용하여 몇 가지 느낌 및 질감을 부여할 수 있었다.

#### 2.1 이동평균처리(Moving Average Processing)

이동평균처리는 주어진 오디오/음성 신호에서 현재 샘플과 이전 샘플의 차이를 감소시키는 일반적인 신호 처리 방법 중 하나이다. 이 처리는 신호에 저주파 통과 필터의 효과를 주게 되어, 고주파 성분이 상대적으로 감소되며 보다 부드러운 한편 장중한 느낌을 가지게 한다. 본 논문에서는 20개 샘플의 창단위로 평균처리를 하였다.

#### 2.2 이동편차강조(Moving Double Differentiation)

이동편차강조는 이동평균처리와는 달리 현재 샘플과 이전 샘플의 차이를 강조하는 방법으로 신호의 저

주파 성분보다 고주파 성분을 강조하며, 신호에 조금 가벼운 느낌을 부여한다. 그러나, 고주파 성분으로 인한 약간의 잡음이 함께 들려 약간 거친 느낌을 주게 된다. 이동평균처리와 마찬가지로 20개 샘플의 창단위로 처리를 하였다.

### 2.3 Differentiation Processing

Differentiation Processing은 현재 샘플과 이전 샘플의 단순 비교/차감을 하는 방법으로 과거와 현재 값의 차이만으로 이루어졌다. 이로 인해 전체적인 에너지가 줄어들어 작게 들리므로, 원래 신호의 에너지와 비슷한 레벨로 정규화 하였다. 이 방법 역시 이동편차 강조와 같이 고주파 성분을 강조하기 때문에 이로 인한 잡음이 발생하게 된다. 이 처리를 거친 신호는 가볍고 거친 느낌이 든다.

### 2.4 이동교차편차(Moving Cross Differentiation)

이동교차편차는 현재 샘플과 이전 샘플의 차이를 교차해서 강조해주는 방법으로 이동평균의 값과 원래 값의 차이를 번갈아 더하거나 차감한 값으로 이루어져 있다. 즉 고주파 성분을 번갈아가며 가감하여 음질의 저하 없이 날카로운 질감을 부여할 수 있다. 이 처리를 거친 신호는 카랑카랑하고 날카로운 금속성의 느낌을 가지게 된다.

### 2.5 Envelope Contour 조절

Envelope Contour 조절은 비교적 원 음색에 손상을 주지 않으며 느낌을 조절할 수 있는 방법이다. 원 신호의 포락선에 대해 다양한 창 함수를 적용함으로써 구현할 수 있으며, 여기에서 사용된 창 함수들은 그림 1과 같다.

### 2.6 Replaying Speed 조절

주어진 오디오 데이터는 곡의 빠르기를 조절함으로써 유쾌한 느낌이나 우울한 느낌이 들도록 조절이 가능하다. 빠르기가 감소하면 우울한 느낌이 들며, 증가하면 보다 유쾌한 느낌이 든다. 이를 위해서 시간 영역에서의 decimation과 interpolation을 사용하였다. Decimation은 곡의 빠르기의 증가를, interpolation은 감소를 가져온다.

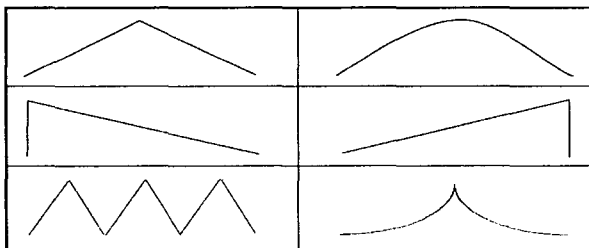


그림 1. Envelope Contour 조절을 위한 창 함수

## III. 질감의 정량화

기본적인 음질의 평가는 물리적 평가와 주관적 평가로 나눌 수 있다. 물리적 평가는 계측기에 의한 측정의 한계가 있기 때문에, 보편적인 음질 측정법으로 주관적 평가가 사용된다. Zwicker 등의 심리 음향 학자들에 의해 다양한 청감인자들이 정의되고, 이를 정량화하기 위한 연구가 진행되어 왔으나, 이는 기본적으로 음질 자체에 대한 정량화였으며, 신호의 느낌이나 질감의 정량화와는 다소 거리가 있다. 그러나, 본 연구에서는 다음과 같은 인자들에 대해 주관적인 MOS(Mean Opinion Score) 평가를 실시해, 신호처리에 의한 느낌이나 질감의 변화의 측정을 시도하고자 한다.

### 3.1 Loudness

Loudness는 주관적으로 느끼는 소리의 크기를 나타내며, Zwicker의 loudness 모델에 기초해서 계산된다. Zwicker의 loudness계산 방법은 DIN 45631과 ISO recommendation 532B로 표준화 되어있다[1],[2].

### 3.2 Roughness

Roughness는 주관적으로 느끼는 소리의 거칠기에 대한 크기를 나타낸다. 20 Hz이상의 진폭 및 주파수 변조되는 소음의 경우에는 음의 크기 변화를 느끼지 못하는 대신 소리의 거칠기를 인지하는데, 이를 roughness로 정의한다[1],[3].

### 3.3 Sharpness

Sharpness는 고주파 성분의 소리로 인해 느끼는 날카롭거나 높은 소리의 느낌을 나타낸다. 스펙트럼 구조나 level 차는 sharpness에 큰 영향을 미치지 못하며, 스펙트럼 envelope 안의 중심 주파수가 결정적인 영향을 미친다. 즉 중심 주파수가 높은 곳에 위치해 있을수록 소리는 더 날카롭게 들린다[1].

### 3.4 Fluctuation Strength

Fluctuation Strength는 20 Hz이하의 modulation 주파수를 가진 sound 신호의 주기적인 파동(fluctuation)은 귀에서 시간에 대한 volume의 변화로 느껴지는데, 이에 대한 지각을 나타낸다[1].

### 3.5 Pleasantness

Pleasantness는 소리가 유쾌하게 들리는 정도를 나타내며, loudness와 roughness, sharpness가 작을수록 pleasantness가 크다[1].

### 3.6 Unbiased Annoyance

Unbiased Annoyance는 듣는 사람의 감정, 소리에

표 1. 주관적 질감 평가표

청감인자	질감 변화 전	질감 변화 후	변화도
Roughness	1/2/3/4/5	1/2/3/4/5	
Sharpness	1/2/3/4/5	1/2/3/4/5	
Fluctuation Strength	1/2/3/4/5	1/2/3/4/5	
Pleasantness	1/2/3/4/5	1/2/3/4/5	
Unbiased Annoyance	1/2/3/4/5	1/2/3/4/5	

대한 주관적 느낌 등 감정적인 요소를 배제한 불쾌함의 정도를 나타낸다[1].

### 3.7 주관적 질감 평가

음질의 물리적 평가는 계측기에 의한 측정의 한계를 가지므로, 일반적으로 음질의 평가는 주관적인 방법에 의해 수행된다. 주관적 평가를 수행하기 위해서는 각각의 청감인자들에 대한 기준음을 설정해 주는 것이 바람직하다. 본 연구에서는 주관적 평가를 통한 질감의 정량화를 시도하기 위해, 각각의 청감인자들에 대해 표1과 같은 평가표를 이용하여 5단계의 MOS 평가를 실시하였다. 아직 이에 대한 구체적이고 보편화된 방법은 제시된 바가 없으므로, 지속적인 연구가 필요하다.

## IV. 실험 및 결과

### 4.1 사용된 오디오 및 음성 신호

실험에 사용된 오디오 신호는 피아노 연주곡(모차르트 피아노 소나타 K.467)의 도입부로 16 bit, 44.1 kHz로 샘플링 된 신호를 16 bit, 16 kHz로 다운 샘플링하여 사용하였다. 음성의 경우 5초 분량의 대화체 음성이며 역시 16 bit, 16 kHz로 다운 샘플링하여 사용하였다.

### 4.2 이동평균처리 결과

그림 2에서 보는 것과 같이 이동평균처리된 신호는 고주파 성분이 상대적으로 감소되었음을 알 수 있다. 이로 인해 신호는 부드러우면서 장중한 느낌을 갖는다.

### 4.3 Differentiation 결과

Differentiation된 신호는 그림 3에서와 같이 고주파 성분이 강조된다. 이로 인해 비교적 잡음이 많이 들리며, 가볍고 거친 느낌이 든다.

### 4.4 이동편차강조 결과

그림 4에서 보는 것과 같이 이동편차강조된 신호는

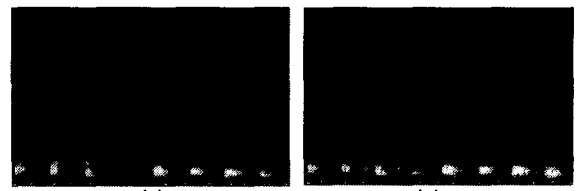


그림 2. 원 신호 및 이동평균처리된 신호의 스펙트럼 : (a) 원신호, (b) 이동평균처리된 신호

원 신호에 비해서 고주파 성분이 강조되었음을 알 수

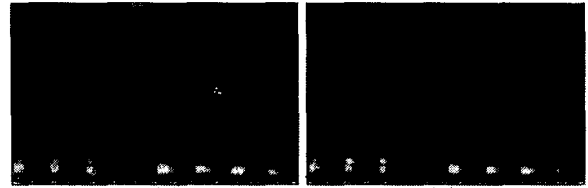


그림 3. 원 신호 및 Differentiation된 신호의 스펙트럼 : (a) 원 신호, (b) differentiation된 신호

있다. 이로 인해 신호는 고주파 성분이 강조된 보다

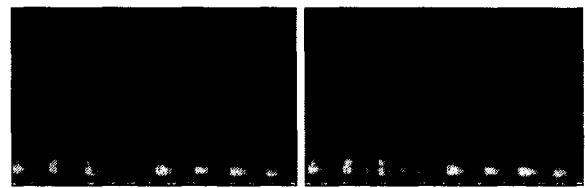


그림 4. 원 신호 및 이동편차강조된 신호의 스펙트럼 : (a) 원 신호, (b) 이동편차강조된 신호

가벼운 느낌을 갖는다.

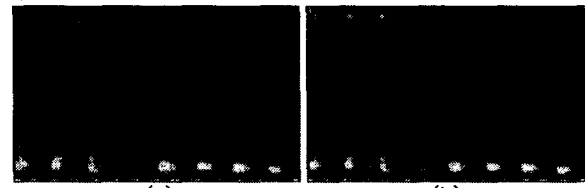


그림 5. 원 신호 및 이동편차교차된 신호의 스펙트럼 : (a) 원 신호, (b) 이동편차교차된 신호

### 4.5 이동편차강조 결과

그림 4에서 보는 것과 같이 이동편차강조된 신호는 원 신호에 비해서 고주파 성분이 강조되었음을 알 수 있다. 이로 인해 신호는 고주파 성분이 강조된 보다 가벼운 느낌을 갖는다.

### 4.6 이동교차편차 결과

이동교차편차된 신호는 그림 5에서 보는 것과 같이 고주파 성분이 변갈아가면서 가감되었음을 알 수 있

다. 이동교차편차된 신호는 들어보았을 때 음질의 저하 없이 카랑카랑하고 날카로운 금속성의 느낌이 든다.

#### 4.7 Decimation과 Interpolation 결과

그림 6은 시간영역에서의 decimation과 interpolation결과를 보여준다. 그림 6에서 가운데 신호가 원 신호이고, 위에 있는 신호가 interpolation, 아래 신호가 decimation된 신호이다. 즉, interpolation된 신호는 곡의 속도가 감소되고, decimation된 신호는 증가되었음을 알 수 있다. 이로 인해 interpolation된 신호는 다소 무겁고 우울한 느낌이 들며, decimation된 신호는 가볍고 상쾌한 느낌이 든다. 본 실험에서는 7% 정도 증가시키거나 감소시켜 평가하였다.

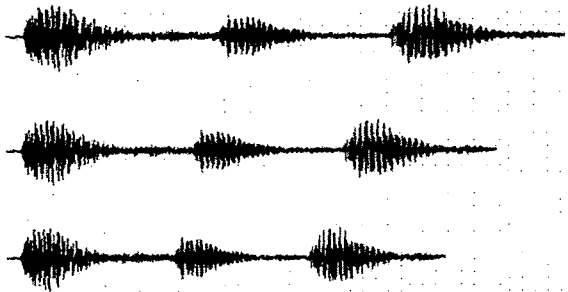


그림 6. 원 신호 및 Decimation/Interpolation된 신호

#### 4.8 Envelope Contour 조절 결과

그림 7은 원 신호에 envelope contour 조절을 적용한 결과이다. 그림 7에서 가운데 신호는 원 신호이고, 위에 있는 신호는 톱니 모양의 창 함수에 의한 신호이며, 아래 신호는 삼각 창 함수에 의한 신호이다. 창 함수에 의해서 원 신호는 비교적 잡음이 많으며, 가볍고 거친 느낌을 갖게 되었다.

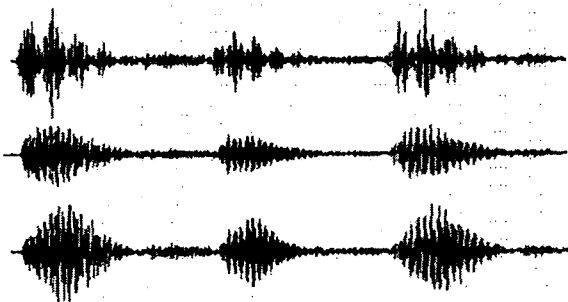


그림 7. 원 신호 및 Envelope Contour가 조절된 신호

#### 4.9 주관적 질감 평가 결과

5 명의 평가자들에게 앞에 설명한 신호처리의 결과를 원 신호와 함께 비교하여 반복 청취를 한 후, 표 1의 평가표를 이용하여 5 단계의 MOS 평가를 실시하였다. 그 결과는 표 2와 같으며, 각 항목마다 5에 가까울수록 그 느낌이 가장 강함을 나타낸다. 사용된 주관

표 2. 주관적 평가 결과

	Roughness	Sharpness	Fluc.St.	Pleasantness	Annoyance
원 신호	2.2	2.6	2.6	3.2	2.6
이동평균	1.4	1.8	2.8	3.0	2.2
이동편차	1.6	3.2	2.2	2.8	2.8
Differenciation	2.0	3.8	1.8	2.4	3.4
이동편차교차	2.4	3.6	3.0	2.2	3.6
Decimation	2.4	2.8	2.4	2.6	2.2
Interpolation	2.6	1.6	3.8	2.0	2.6
ECM-tri	2.2	2.2	2.8	2.2	2.2
ECM-saw	4.0	2.6	4.2	1.0	4.6

적 질감 평가는 아직 객관성을 보여주지 못하고 있으며 앞으로 계속 보완되어야 할 것이다.

## V. 결 론

지금까지 시도된 바 없는 오디오 및 음성신호에 대하여 느낌 및 질감 부여를 시도하였다. 그 첫 번째 접근 방법으로 시간 영역에서의 다양한 신호처리 방법들을 시도하였고, 결과에서 볼 수 있듯이 몇몇 청감인자들에 대해 그 변화를 확인할 수 있었다. 향후 보완을 통해서 '부드럽다', '울퉁불퉁하다', '가볍다', '무겁다', '날카롭다', '거칠다' 등의 느낌이나 질감 부여를 실험할 계획이다. 특히 음성신호의 경우, 감정 정보를 부여하는 연구를 계획 중이다. 또한, 주파수 영역에서의 처리 방법들에 대한 연구도 진행 중이며, 피아노 외의 다른 악기들에 대해서는 확장 실험 중이다. 더불어 음의 질감의 정량화에 대한 보다 구체적인 연구도 계속되어야 할 것이다.

## 참고문헌

- [1] 임준희, 이신렬, 김순협, 정완섭, "바이노럴 음질 평가 시스템 구현에 관한 연구", 한국 음향학회 학술 발표대회 논문집 제17권 제1호, pp.283-286, 1998.
- [2] 정혁, 이정권, "임계 대역 필터를 이용한 과도 음의 라우드니스 모델", 한국음향학회 학술발표대회 논문집 제17권 제1호, pp.279-282, 1998.
- [3] 정혁, 김현빈, 이정권, "소리의 주관적 평가인자 러프니스 계산 알고리즘 구현", 한국음향학회 학술발표대회 논문집 제20권 제1호, pp.787-790, 2000.