

# Spectral Band Replication 대역폭에 따른 MPEG USAC 부호화 성능 평가

안 경 준<sup>a)</sup>, 정 유 선<sup>a)</sup>, 백 승 권<sup>b)</sup>, 강 경 옥<sup>b)</sup>, 김 인 철<sup>a)‡</sup>

## Performance Evaluation of the MPEG USAC According to the Spectral Band Replication Bandwidth

Kyung Jun An<sup>a)</sup>, Yoosun Jung<sup>a)</sup>, Seungkwon Beack<sup>b)</sup>, Kyeongok Kang<sup>b)</sup>, and Rin Chul Kim<sup>a)‡</sup>

### 요 약

본 논문에서는 SBR의 대역폭이 MPEG USAC의 전체 성능에 끼치는 영향을 다룬다. 여기서 SBR 대역폭은 SBR 코덱이 적용되는 주파수 영역을 말하고 SBR 비트열 성분 중 하나인 `bs_stop_freq`에 의해 규정된다. 다섯 개의 서로 다른 SBR 대역폭을 가질 때 USAC의 성능을 MUSHRA 테스트라는 주관적인 방법으로 비교하였다. 이 비교에서 전송률은 14~24Kbps로 한정하였고, 주부호화기로는 LPD만을 사용하였다. 실험결과 SBR 대역폭이 18KHz 이상일 때가 다른 대역폭에 비해 좋은 성능을 보이는 것으로 관측되었다.

### Abstract

This paper deals with the effect of SBR bandwidth on the overall performance of the MPEG USAC. Here, the SBR bandwidth is termed the frequency region covered by the SBR codec, and is specified by the `bs_stop_freq`, which is one of the SBR bitstream components. The performance of the USACs with 5 different SBR bandwidths are compared in a subjective manner using the MUSHRA test. In the comparison, the bit rate is confined to 14~24kbps and only the LPD unit is selected for the core codec. From the comparison, it is observed that the SBR bandwidth that stretches up to 18KHz or above gives the better performance than the others.

Keyword : MPEG USAC, SBR, MUSHRA test, SBR bandwidth

## 1. 서 론

최근 수년간 음성과 오디오 신호를 통합적으로 서비스하

려는 노력이 진행되고 있다. 그 대표적인 예로 2007년부터 표준화 작업이 시작되어 현재 표준화 작업이 마무리 단계인 MPEG USAC(unified speech and audio codec)이 있다. USAC<sup>[1][2]</sup>이 요구하는 성능 조건은 12~64Kbps 전송률에서 음성과 오디오에 대하여 각각 최고 성능을 가지는 AMR-WB<sup>[3]</sup>와 HE-AAC v2<sup>[4]</sup>보다 동등하거나 우수한 품질을 제공하는 것이다.

MPEG USAC은 저주파 대역 부호화기로 AAC(advanced audio coding), ACELP(algebraic code excited line-

a) 서울시립대학교 전자전기컴퓨터공학부  
Dept. of Electrical and Computer Eng., Univ. of Seoul  
b) 한국전자통신연구원 실감음향연구팀  
Realistic Acoustics Research Team, ETRI  
‡ 교신저자 : 김인철 (rin@uos.ac.kr)  
※ 이 논문은 한국전자통신연구원의 무안경 다시점 3D 지원 UHD TV 방송 기술 개발 과제 일환으로 수행된 연구결과물입니다.  
· 접수일(2011년6월26일), 수정일(2011년8월29일), 게재확정일(2011년9월2일)

ar prediction), 혹은 TCX(transform coded excitation)를 선택적으로 채택한다. 그리고 고주파 대역 부호화를 위해 SBR(spectral band replication), 다채널 확장을 위해 MPEG surround<sup>[5]</sup> 기법을 채택하였다. 또한 각 부호화 기술들을 확장 및 수정하여 향상된 성능을 얻을 수 있도록 표준을 제정하고 있다.

ACELP<sup>[3]</sup>, TCX<sup>[3]</sup> 혹은 AAC<sup>[6]</sup>로 구성되는 저주파 대역 부호화기는 입력된 신호를 하향 표본화(down sampling) 후 변환 혹은 선형 예측(linear prediction) 등의 기법을 사용하여 비교적 많은 비트량을 소비하면서 신호를 충실히 부호화한다. 따라서 일반적으로 저주파 대역 부호화기를 주부호화기(core codec)로 부른다. 반면에 SBR<sup>[7]</sup> 및 MPEG surround 기법은 이미 부호화된 저주파 대역 신호를 참조하여 적은 비트량으로 표현되는 파라미터만으로 고주파 대역 혹은 다채널 신호를 합성한다. 따라서 SBR과 MPEG surround 기법을 파라미터 부호화(parametric coding)라고 부른다.

MPEG surround 기법에서는 단일 채널 신호로부터 다채널 파라미터를 이용하여 2 혹은 3채널 신호로 합성한다. 이 기법에서는 다채널 파라미터의 선택 및 추정 방법 등이 부호화 성능에 영향을 준다. SBR 기법에서는 저주파 대역 신호로부터 SBR 파라미터를 이용하여 고주파 대역 신호를 합성하고 이로부터 전 대역(full band) 신호를 복원한다. SBR 기법에서는 SBR 파라미터 선택 및 추정 방법도 중요하지만, SBR 대역폭 또한 부호화 성능에 큰 영향을 준다.

SBR 대역폭은 주부호화기에서 부호화되는 저주파 대역 이후에 SBR 기법에 의해 확장되는 대역폭을 말한다. 전체 부호화 대역폭은 주부호화기에서 부호화되는 저주파 대역폭에 SBR 대역폭을 합하여 이루어진다. 일반적으로 주부호화기에서 처리하는 저주파 대역폭은 3~12KHz 범위에서 설정되고, SBR 대역폭은 저주파 대역폭과 같은 넓이로 설정된다. 즉, 전체 부호화 대역폭은 일반적으로 저주파 대역폭의 2배로 설정된다. 전체 부호화 대역폭이 넓으면 공간감이 향상되고 풍부한 음색의 소리를 재현할 수 있다. 그러나 주부호화기에서 처리하는 저주파 대역폭을 넓히면 많은 비트량이 추가적으로 요구되어 낮은 전송률에서는 높은 품질의 오디오를 재현하기 어렵다. 반면에, SBR 파라미터는 매

우 적은 비트량을 요구하므로 SBR 대역폭은 매우 적은 비트량으로 넓힐 수 있다.

본 논문에서는 주부호화기의 대역폭을 고정된 상태에서 SBR 대역폭을 가변시킴으로써 전체 부호화 대역폭을 조정하였을 때 부호화 성능을 평가하고자 한다. 성능 평가를 통해 전체 부호화 성능을 최적화할 수 있는 적절한 SBR 대역폭 및 전체 부호화 대역폭을 조사한다. 본 논문에서 주부호화기로는 ACELP와 TCX를 선택적으로 사용하는 LPD(linear prediction domain) 유닛(unit) 및 양자화 모듈을 사용한다. SBR 부호화에서는 전체 부호화 대역폭이 10.875KHz, 13.125KHz, 15.375KHz, 18.375KHz, 그리고 20.25KHz 등으로 표현되도록 총 5가지 형태로 SBR 대역폭을 조정하였다. 5가지 SBR 대역폭을 가진 부호화기로 부호화 된 결과는 MUSHRA[8] 테스트를 통해 그 성능들을 비교한다. 부호화 실험은 서로 다른 특성을 가진 4개의 실험 오디오 신호를 대상으로 하였고, 9명의 테스트에 익숙한 사람들이 그 성능을 평가하였다.

서론에 이어 2절에서는 MPEG USAC 부호화기를 설명한다. 3절에서는 SBR 부호화기 및 SBR 대역폭 결정 방법에 대해 설명한다. 그리고 4절과 5절에서는 각각 실험 결과와 본 논문의 결론을 제시한다.

## II. MPEG USAC

MPEG USAC은 음성과 오디오 신호를 통합하여 처리하는 부호화 기술로 현재 표준화 작업이 마무리 단계에 있다. USAC은 기존의 음성 및 오디오 부호화 기술을 효과적으로 융합할 뿐만 아니라 요소 기술들을 추가시켜 개선된 성능을 얻도록 하고 있다. USAC기술에는 MPEG surround, eSBR, ACELP, TCX, 그리고 AAC가 포함되어있다.

그림 1에서는 MPEG USAC의 부호화기와 복호화기를 나타내었다. 먼저 다채널 오디오 신호가 USAC 부호화기에 입력되면 MPEG surround에서 공간 파라미터(spatial parameter)를 추출하고, 다운믹스(down mix)된 모노 신호를 출력한다. eSBR에서는 입력된 모노 신호로부터 SBR 파라미터를 추출하고 모노 저주파 대역 신호를 출력한다. 추출

된 SBR 파라미터는 고주파 대역을 부호화할 때 필요한 보정 파라미터가 된다. 모노 저주파 신호는 LPD 혹은 FD(frequency domain) 유닛과 이에 따르는 양자화 및 산술 부호화(arithmetic coding)를 거쳐 부호화된다.

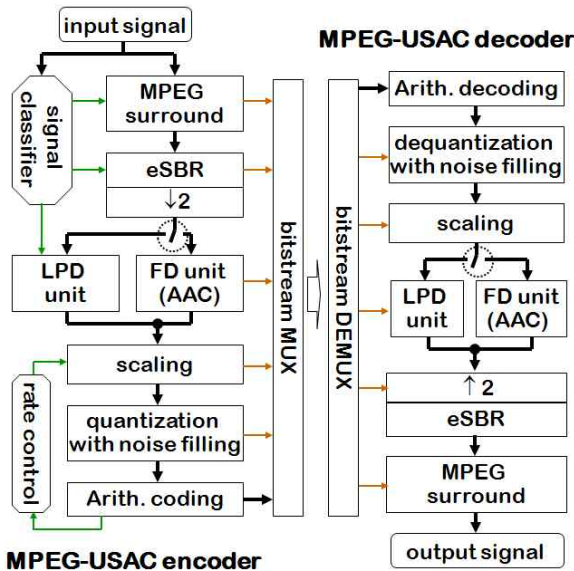


그림 1. MPEG USAC의 블록선도  
Fig. 1. Block diagram of the MPEG USAC

이와 같이 모노 저주파 신호를 부호화하는 장치를 주부호화기라고 한다. 주부호화기에 대해 좀 더 구체적으로 살펴보면 다음과 같다. 먼저 FD 유닛에서는 MDCT를 통해 시간영역의 신호를 주파수 영역으로 변환한다. 그런 다음, TNS(temporal noise shaping)를 이용하여 주파수 영역 개루프 DPCM을 수행함으로써 양자화 잡음이 시간영역에서 신호의 포락선(time envelope)에 비례하도록 조정한다. 마지막으로 심리음향모델(psychoacoustic model)을 도출하고 이를 참조하여 축적률 조정, 양자화, 산술 부호화 등을 거쳐 MDCT 계수들을 표현하여 전송한다. USAC FD 유닛은 MPEG-4 AAC-LC<sup>[6]</sup>와 유사하다. 그러나 잡음 채움(noise filling)을 도입하고, 호프만 부호화 대신 산술 부호화한다는 점 등에서 차이를 보인다.

LPD 유닛은 LPC를 기반으로 동작하며 ACELP와 TCX를 포함한다. 먼저 선강조(pre-emphasis)를 통해 고주파를

강조한 모노 저주파 대역 신호를 만들고 선형예측을 수행한다. 그 결과 얻어진 여기 신호(excitation signal)를 TCX와 ACELP를 선택적으로 사용하여 부호화한다. ACELP는 음성신호를 부호화하는데 적합한 기술로 피치(pitch)와 여기 신호를 각각 벡터 양자화하여 부호화한다. TCX는 혼합 신호와 저전송률 오디오 신호를 부호화하는데 적합한 기술로 입력된 여기신호에 대해 MDCT를 수행한다. 그런 다음 MDCT 계수를 양자화하고 산술 부호화하여 전송한다. AMR-WB+의 TCX와는 달리 USAC TCX에서는 DFT 대신 MDCT를 사용하고, 벡터 양자기 대신에 스칼라 양자기 및 산술부호화를 적용한다.

복호화기는 부호화기의 반대로 동작한다. 간단하게 설명하면 비트열을 입력받아 산술 복호화를 한 뒤 양자화 역과정과 축적률 역과정을 거친 다음 LPD 혹은 FD 유닛에서 모노 저주파 대역 신호가 복원된다. 그런 다음 eSBR에서 고주파 대역 합성을 통해 모노 전 대역 신호를 복원하고, MPEG surround에서 다채널의 전 대역 신호로 복원된다.

### III. SBR

SBR은 고주파 대역 신호를 효과적으로 부호화 하는 기술로서 보통 AAC-LC와 결합하여 HE-AAC 등으로 사용되었다. 그림 2에서는 SBR의 기본 원리를 나타내었다. SBR은 오디오 신호의 하모닉 성분으로 인해서 고주파 대역과 저주파 대역 사이에 높은 연관성이 존재한다는 가정에 기반을 두고 있다. 즉 저주파 대역에 있는 정보를 이용하면 고주파 대역의 정보를 추정해낼 수 있다는 것을 의미한다. 그림 2에 보인 것과 같이 SBR이 적용되는 시점을 SBR 시작 주파수(start frequency), SBR의 적용이 끝나는 점을 SBR 끝 주파수(stop frequency)라고 한다. SBR 시작 주파수 이하의 저주파 대역은 주부호화기에 의해 부호화되고, SBR에서는 시작 주파수와 끝 주파수 사이의 대역을 적절한 파라미터로 표현한다. 본 절에서는 먼저 SBR 부호화기에 대해 간략하게 살펴보고, SBR 대역폭 결정 방법을 설명한다.

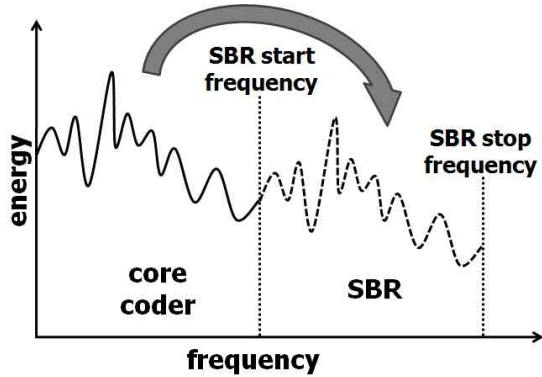


그림 2. SBR 기본 원리  
Fig. 2. Basic Principle of the SBR

### 1. SBR 부호화기

그림 3에 그 구조를 제시한 바와 같이<sup>[9]</sup> USAC에서 채택한 SBR의 기본 원리는 MPEG-4 SBR과 유사하다. 먼저 입력된 2048샘플은 QMF 분석을 통해 32(시간)×64(주파수)의 QMF 부대역 신호(subband signal)로 변환된다. 이 부대역 신호를 가지고 과도점 검색기(transient detector)에서 과도점이 있는지 검사한다. 프레임 분할기(frame splitter)에서는 과도점이 있는 경우 과도점을 기반으로 가변적으로 프레임을 나누고, 과도점이 없는 경우에는 프레임을 일정한 크기로 나눈다. 그런 뒤에 시간/주파수 표본 위치 생성기(T/F grid generator)에서 시간/주파수 표본 위치를 생성한다. 여기서 생성된 시간/주파수 표본 위치를 중심으로 SBR 정보들이 표현된다.

한편, 톤알리티 검색기(tonality detector)에서 2차 선형 예측을 통해 톤알리티를 검색한다. 이 결과를 바탕으로 추가적인 정보들을 생성한다. 여기서의 추가적인 정보는 고주파 대역의 특성을 유지하면서 그 신호를 복원하는데 필요한 파라미터들이 된다. 여기에는 에너지 포락선(energy envelope), 잡음바닥(noise floor), 추가적인 정현파 신호(additional sines), 역 필터링(inverse filtering) 등이 포함된다. 이 추가적인 정보를 SBR 파라미터라고 하며 이 파라미터는 적은 비트량으로 표현이 가능하므로 압축효율을 높여 준다.

한편 복호화기는 부호화기의 반대 순서로 동작된다. 먼

저 전송받은 비트열로부터 SBR 파라미터를 복원한다. 복원한 SBR 파라미터를 이용하여 주부호화기에서 복원된 저주파 대역 신호를 가공하여 고주파 대역신호를 합성한다. 합성된 고주파 대역 신호는 저주파 대역 신호와 함께 QMF 합성을 통해 전 대역 신호로 복원된다.

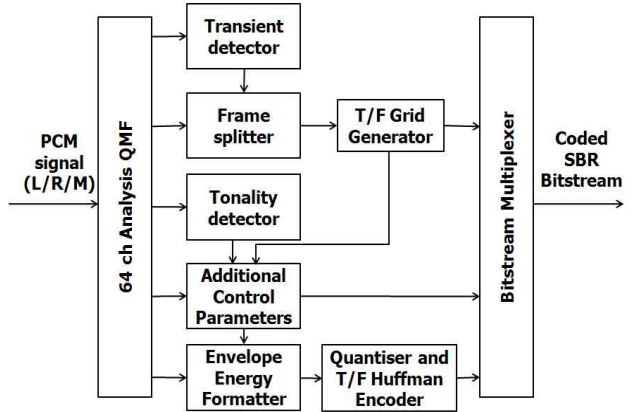


그림 3. SBR 부호화기의 블록선도  
Fig. 3. Block diagram of the SBR encoder

### 2. SBR 대역 결정

SBR 시작 주파수 이하의 저주파 대역 신호는 주부호화기에서 부호화하고 SBR 부호화기에서는 이 저주파 대역 신호를 토대로 SBR 시작 주파수에서 끝 주파수까지의 고주파 대역 신호를 복원한다. SBR은 전 대역을 64 QMF 대역으로 분석하기 때문에 SBR 시작 주파수와 끝 주파수는 각각 비균일 간격으로 나타내어지는 4비트 인덱스로 표현되어 전송된다.

일반적으로 SBR 시작 주파수와 끝 주파수는 비트 전송률, 채널 수, 그리고 표본화율(sampling frequency)에 따라 결정된다. SBR 시작 주파수는 주부호화기에서 부호화되는 신호의 대역폭을 결정하므로 주부호화기의 부호화 충실도와 관계된다. 만약 SBR 시작 주파수가 낮다면 주부호화기에서 처리해야 하는 저주파 대역 신호의 대역폭이 좁아진다. 이 때, 재표본화(resampling) 혹은 제한된 범위의 계수들만 부호화하는 영역부호화(zonal coding)등의 방법을 통

해 단위 시간당 부호화되는 샘플 수를 줄일 수 있다. 따라서 주어진 비트 전송률에서 대역폭이 좁아지면 부호화되는 샘플 수가 줄어들어 샘플 당 비트량이 많이 할당되므로 부호화 충실도가 높아진다. 한편, SBR 시작 주파수와 SBR 끝 주파수 사이의 대역, 즉 SBR 대역은 SBR 부호화기에 의해 표현되는 대역이다. 주부호화기에 의해 수행되는 저주파 대역 부호화와는 달리 SBR 대역은 적은 양의 SBR 파라미터로 표현되므로 효과적으로 넓은 대역을 표현할 수 있다는 장점이 있다. 일반적으로 대역폭이 넓어지면 공간감이 향상되고 풍부한 음색의 영향을 재현할 수 있다. 그러나 대역폭이 넓어지면 SBR 파라미터로 표현한 고주파 대역과 원래의 고주파 대역의 차이가 커질 수 있다. 또한 SBR 대역이 넓어짐에 따라 SBR 파라미터의 정보량이 많아지고 이에 따라 주부호화기에서 소비하는 정보량이 줄어들어 저주파 대역의 부호화 품질이 저하될 우려가 있다. 본 논문에서는 SBR 대역폭에 따른 부호화 성능을 조사하고자한다.

먼저, 표본화 주파수를 48KHz라고 할 때 주부호화기에서 부호화되는 신호의 대역폭은 표본화 주파수의 1/4보다 약간 작은 5.625KHz로 제한한다. 이는 SBR 시작 주파수  $k_0=15$ 에 해당하고  $bs\_start\_freq=7$ 이다. 그런 다음, 전체 부호화 대역폭이 10~20KHz 범위에서 표현되도록 SBR 끝 주파수를 조정한다. 즉, 비트열 구성 요소인  $bs\_stop\_freq$ 를 4, 6, 8, 10, 11로 변화시켜 SBR 끝 주파수  $k_2$ 를 각각 29, 35, 41, 49, 54로 설정하여 SBR 대역폭을 가변시킨다. 이때 전체 부호화 대역폭은 각각 10.875KHz, 13.125KHz,

15.375KHz, 18.375KHz 및 20.25KHz가 된다. 본 논문에서는 5개의 전체 부호화 대역폭에 대해 비트량과 음질 평가를 통해 최적의 SBR 대역폭을 알아보하고자 한다.

표 1에서는 실험신호 *twinkle\_ff51*을 24kbps로 부호화하였을 때 각 SBR 파라미터 및 주부호화기에서 소비되는 평균 비트량을 제시하였다. SBR 끝 주파수 색인이 4에서 11로 증가함에 따라 패치의 수가 2에서 4로 증가하고, 주파수 표본위치의 개수 또한 증가한다. 따라서 에너지 포락선, 잡음바닥, 그리고 역 필터링 등을 위한 비트량이 증가하여, SBR 비트량은 끝 주파수 색인이 2만큼 증가할 때마다 약 10~15비트 증가함을 알 수 있다. SBR 비트량이 증가하면, 그만큼 주부호화기의 비트량이 감소하여 주부호화기의 부호화 성능 저하를 야기시킨다. 이러한 상황에서 전체 부호화 성능을 실험을 통하여 비교하고자 한다.

#### IV. 실험 및 결과

##### 1. 실험 방법

본 논문에서는 주부호화기에서 부호화되는 신호의 대역폭을 고정한 상태에서 SBR 대역폭에 따른 음질평가를 수행한다. 실험에 사용한 오디오 신호는 48KHz의 표본화율을 갖는 모노(mono) 신호로 *Arirang\_speech*, *Music\_1*, *te15*, *twinkle\_ff51*이다. *Arirang\_speech*는 남성 목소리 신

표 1. *twinkle\_ff51*의 SBR 대역폭에 따른 파라미터들의 비트 전송률  
Table 1. Bit rates of the parameters on *Twinkle\_ff51* according to the SBR bandwidth (단위, bits/frame)

Parameter		stop freq stop 4 10.875Hz	stop 6 13.125Hz	stop 8 15.375Hz	stop 10 18.375Hz	stop 11 20.25Hz
SBR	SBR_ivec	4	4	6	6	8
	SBR_env	49.95	60.02	71.15	82.56	82.89
	SBR_noise	6.36	6.26	8.67	8.51	10.8
	etc	29.52	29.64	29.49	29.47	29.56
	SBR_total	89.83	99.92	115.31	126.54	131.25
core		907.25	891.33	875.42	859.5	859.5
etc		29.17	29.63	30.48	31.15	31.6
total		1026.25	1020.88	1021.21	1017.19	1022.35

호로 저주파 대역 신호가 상대적으로 강한 특성을 가지고 있다. Music\_1은 락(rock) 음악으로 전자 기타소리와 드럼 소리가 강한 특성을 갖는다. te15는 타악기 소리가 강한 오케스트라 연주이고 마지막으로 twinkle\_ff51은 드럼, 베이스, 색소폰의 연주와 여성의 내레이션으로 이루어져 있다.

실험에서는 위의 오디오 신호를 14Kbps, 18Kbps, 24Kbps의 전송률에서 2절과 3절에서 설명한 MPEG USAC과 SBR 부호화 방법으로 부호화하였다. 단, 주부호화기는 FD를 사용하지 않고, LPD 유닛만으로 6KHz의 저주파 대역폭을 부호화하였다. SBR은 3절에서 설명한 5가지 SBR 대역폭을 적용하여 부호화하고, 각 성능을 비교하였다. 성능 비교는 서로 상이한 SBR 대역폭으로 부호화한 5가지 부호화 결과, 3.5KHz 저주파 대역 통과필터로 얻은 앵커신호, 그리고 원본신호 등 7개의 신호를 가지고 수행하였다.

본 논문에서는 대표적인 주관적 음질평가방법인 MUSHRA 테스트로 성능을 비교하였다. 평가에는 MUSHRA 테스트에 경험이 있는 7명의 전자전기 전공 학생과 2명의 음악 전공자가 참여하였다. 각 평가자는 각 실험 신호에 대해 0-100점의 점수를 부여하였고, 95% 신뢰구간으로 성능을 비교한다.

## 2. MUSHRA 테스트 결과

4가지 실험 오디오 신호를 14Kbps의 비트 전송률로 부호화하였을 때 주관적 음질 평가 결과는 그림 4에 제시하였다. 그림 4에서는 3.5KHz 앵커(A)와 5개의 전체 부호화 대

역폭 등 6개의 서로 다른 부호화 결과를 9명의 평가자가 평가하고, 평가 결과의 95% 신뢰구간을 도시한 것이다.

Arirang\_speech는 저주파 대역 신호가 강한 남성의 목소리 오디오 신호이다. 따라서 전체 부호화 대역폭이 15.375KHz 이상에서는 고주파 대역이 넓어 졌음에도 음질에 큰 차이를 보이지 않고 있다. 또한 13.325KHz와 15.375KHz 결과는 신뢰구간이 많이 겹치는 것으로 나타나, 두 경우의 성능 차이가 크지 않음을 알 수 있다.

Music\_1은 전체 부호화 대역폭 18.375KHz보다 작은 경우 드럼의 하이햇소리가 들리지 않거나 둔탁하게 들리며, 기타소리도 자연스럽게 못하다. 전체 부호화 대역폭이 18.375KHz 이상인 경우 고주파 대역 신호가 점점 자연스럽게 들린다. te15는 Music\_1에서와 마찬가지로 전체 부호화 대역폭 18.375KHz보다 작은 경우에는 탬버린 같은 타악기의 소리가 들리지 않거나 둔탁하게 들린다. 전체 부호화 대역폭을 18.375KHz 이상으로 갖는 경우에는 타악기음이 선명해지는 않지만 어느 정도 표현해 주고 있다. twinkle\_ff51은 SBR 대역폭이 증가하면서 고주파 대역 신호의 부재로 인한 둔탁한 소리는 점점 사라진다. 전체 부호화 대역폭이 18.375KHz, 20.25KHz인 경우는 거의 구분하기가 힘들다.

그림 4의 결과를 종합하면, (a)와 같은 음성신호에 대해서는 전체 부호화 대역폭이 13KHz 이상으로 요청되고 (b)-(d)와 같은 음향신호에 대해서는 전체 부호화 대역폭이 18KHz 이상으로 요청된다.

다음으로 18Kbps와 24Kbps의 비트 전송률로 부호화된

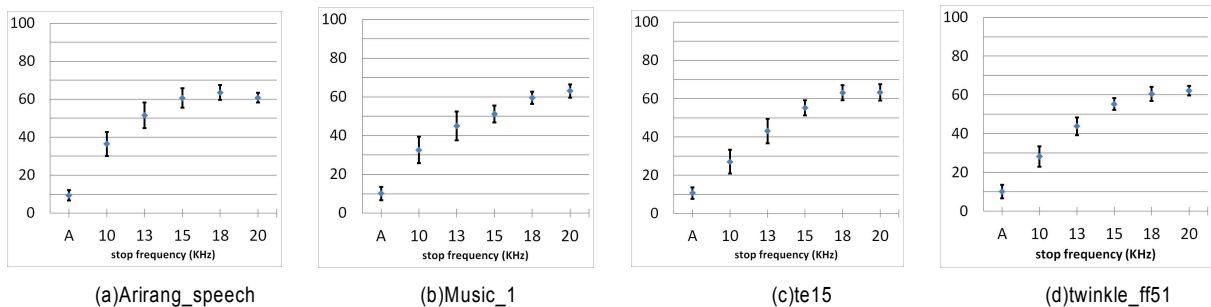


그림 4. 주관적 음질 평가 결과(14Kbps)

Fig. 4. Results of the subjective tests at 14Kbps

오디오의 주관적 음질 평가 결과는 각각 그림 5와 6에 제시하였다. 그림 5와 6을 참조하면 비트 전송률이 증가하면 전체 부호화 대역폭 증가에 따른 주관적 음질의 차이가 더욱 커지고, 음질 평가 결과 또한 높아진다. 즉, 비트 전송률이 14Kbps, 16Kbps, 24Kbps로 증가함에 따라 주관적 음질 평가의 최대치는 각각 65, 72, 78로 커짐을 알 수 있다. 그리고 Arirang\_speech, Music\_1 등에서 14Kbps의 비트 전송률에서는 전체 부호화 대역폭이 13.125KHz일 때와 15.375 KHz일 때의 신뢰구간이 서로 겹쳐지는 현상이 나타났다. 반면에 18Kbps, 24Kbps로 비트 전송률이 높아짐에 따라 전체 부호화 대역폭에 따른 주관적 음질 평가의 신뢰 구간이 더 이상 겹쳐지지 않고, 분명하게 구별되었다.

### V. 결론

본 논문에서는 MPEG USAC 부호화 기법에서 SBR이

담당하는 대역폭을 가변시킴으로써 전체 부호화 대역폭을 조정하였을 때 부호화 성능을 평가하였다. 주부호화는 LPD만을 사용하고, 6KHz 이하의 저주파 대역 신호를 부호화하도록 하였다. 고주파 대역 부호화는 SBR 부호화로 처리되는데, SBR 끝 주파수를 조정함으로써 SBR이 담당하는 대역폭을 가변시켰다. 부호화 성능은 MUSHRA 테스트를 통해 평가되었는데, 9명의 평가자들이 0-100점의 점수로 평가하고, 95% 신뢰구간을 측정하여 각 부호화 방법의 성능을 비교하였다. 비교 결과, Arirang\_speech와 같은 저주파 대역 음성이 포함된 신호에 대해서는 전체 부호화 대역폭이 13~15KHz 이상일 때 부호화 성능이 동일한 것으로 측정되었다. 반면에 음향신호의 경우 전체 부호화 대역폭이 18KHz 이상일 때 동일한 부호화 성능을 보이는 것으로 측정되었다. 따라서 6KHz 이하의 저주파 대역 신호를 주부호화기에서 처리할 때는 전체 부호화 대역폭을 18KHz로 설정함으로써 우수한 성능을 얻을 수 있을 것으로 판단된다. 즉, 주부호화기 대역폭의 2배 즉 전체 부호화 대역폭을

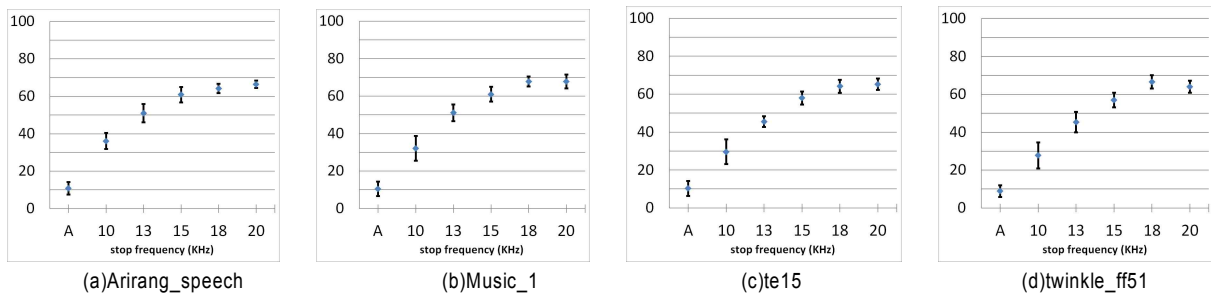


그림 5. 주관적 음질 평가 결과(18Kbps)  
 Fig. 5. Results of the subjective tests at 18Kbps

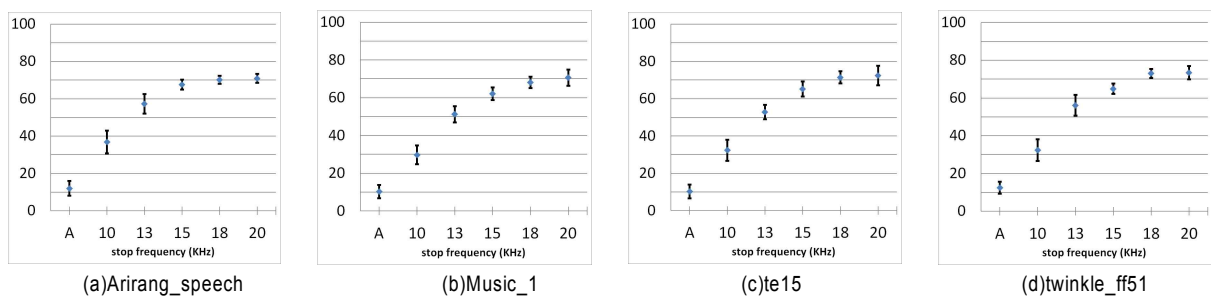


그림 6. 주관적 음질 평가 결과(24Kbps)  
 Fig. 6. Results of the subjective tests at 24Kbps

12KHz로 부호화하는 일반적인 방법 대신에, 주부호화기 대역폭의 3배 이상 즉 전체 부호화 대역폭을 18KHz 이상으로 부호화함으로써 보다 우수한 성능을 얻을 수 있음을 보였다.

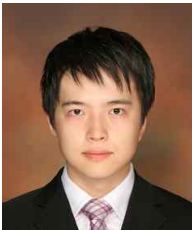
본 논문에서는 주부호화기에서 6KHz 이하의 저주파 대역 신호를 부호화할 때 MPEG USAC의 성능을 살펴보았다. 만약 주부호화기에서 부호화하는 저주파 대역 신호의 대역폭을 가변시켰을 때 최적의 성능을 얻을 수 있는 방법을 조사한다면, MPEG USAC 성능을 향상시키는데 큰 도움이 되리라 기대한다.

### 참 고 문 헌

- [1] 박호중, 박영철, "MPEG 오디오 표준 기술: USAC," 한국방송공학회지 제 14권 제 2호, pp. 63-74, 2009년.
- [2] ISO/IEC DIS 23003-3:2011(E), "Information technology-MPEG audio technologies-Part 3: Unified speech and audio coding," 2011.
- [3] 3GPP TS 26.290: Extended AMR-WB+ Codec; Transcoding Functions, v7.0.0, 2007.
- [4] M. Wolters, K. Kjörling, D. Homm, and H. Purnhagen, "A closer look into MPEG-4 high efficiency AAC," in Proceedings of the 115th Convention of the Audio Engineering Society, preprint 5871, Los Angeles, CA, USA, Oct. 2003.
- [5] ISO/IEC FDIS 23003-1:2006(E), "MPEG audio technologies-Part 1: MPEG surround," 2006.
- [6] Tsung-Han Tsai, Shih-Way Huang, and Liang-Gee Chen, "Design of a low power psycho-acoustic model co-processor for MPEG-2/4 AAC LC stereo encoder," International Symposium on Circuits and Systems, vol. 2, pp. II-552-II-555, 2003.
- [7] M. Schug, A. Groschel, M. Beer, and F. Henn, "Enhancing audio coding efficiency of MPEG layer-2 with spectral band replication (SBR) for DigitalRadio (EUREKA 147/DAB) in a backwards compatible way," in Proceedings of the 114th Convention of the Audio Engineering Society, preprint 5850, Amsterdam, Netherlands, Mar. 2003.
- [8] ITU-R Recommend. BS.1534, "Method for the subjective assessment of intermediate quality level of coding systems (MUSHRA)," 2001.
- [9] 3GPP TS 26.404: Enhanced aacPlus General Audio Codec; Enhanced aacPlus Encoder SBR Part, v7.0.0,2007.

---

### 저 자 소 개



안 경 준

- 2009년 : 서울시립대학교 전자전기컴퓨터공학부 학사 졸업
- 2011년 : 서울시립대학교 전자전기컴퓨터공학부 석사 졸업
- 주관심분야 : 신호처리, 오디오 코덱



정 유 선

- 2011년 : 서울시립대학교 전자전기컴퓨터공학부 학사 졸업
- 2011년 ~ 현재 : 서울시립대학교 전자전기컴퓨터공학부 석사과정
- 주관심분야 : 신호처리, 오디오 코덱



---

저 자 소 개

---



백 승 권

- 1999년 2월 : 한국항공대학교 항공전자공학과 (공학사)
- 2001년 2월 : 한국정보통신대학교 전자공학 (공학석사)
- 2005년 8월 : 한국정보통신대학교 전자공학 (공학박사)
- 2005년 8월 ~ 현재 : ETRI 실감음향연구팀, 선임연구원
- 주관심분야 : 오디오 신호처리 및 부호화, 3D 오디오



강 경 옥

- 1985년 2월 : 부산대학교 물리학과 (이학사)
- 1988년 2월 : 부산대학교 대학원 물리학과 (이학석사)
- 2004년 2월 : 한국항공대학교 대학원 항공전자공학과 (공학박사)
- 2006년 4월 ~ 12월 : 영국 Southampton University, ISVR 방문연구원
- 1991년 2월 ~ 현재 : ETRI 실감음향연구팀장, 책임연구원
- 주관심분야 : 오디오 신호처리 및 부호화, 3D 오디오



김 인 철

- 1985년 : 서울대학교 제어계측공학과 학사 졸업
- 1987년 : 서울대학교 제어계측공학과 석사 졸업
- 1992년 : 서울대학교 제어계측공학과 박사 졸업
- 1992년 ~ 1994년 : 대우전자(주) 영상연구소 선임연구원
- 1994년 ~ 1999년 : 한성대학교 정보전산학부 조교수
- 1999년 ~ 현재 : 서울시립대학교 전자전기컴퓨터공학부 교수
- 주관심분야 : 실시간 신호처리, 오디오 신호처리, 오디오/비디오 압축