

정규논문 (Regular Paper)

방송공학회논문지 제18권 제1호, 2013년 1월 (JBE Vol. 18, No. 1, January 2013)

http://dx.doi.org/10.5909/JBE.2013.18.1.106

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

가변 핵심 대역 하향 표본화 비를 가진 MPEG USAC 성능 평가

이재화^{a)}, 김인철^{a)†}

A Performance Evaluation of the MPEG USAC with Variable Core-Band Down-Sampling Ratio

Jae Hwa Lee^{a)} and Rin Chul Kim^{a)†}

요 약

본 논문에서는 내부 표본화 주파수와 핵심 대역 하향 표본화 비의 조정이 MPEG USAC의 전체 성능에 끼치는 영향을 다룬다. 여기서 내부 표본화 주파수는 실제 부호화될 때의 내부적인 표본화 주파수를 말한다. 핵심 대역 하향 표본화 비는 전체 부호화 대역폭에 대한 핵심 대역폭의 비를 나타낸다. 성능은 6개의 실험 음원에 대해 10명이 참가한 MUSHRA 검사를 통해 평가되었다. 실험에서는 기존의 1/2에 비해 1/3 혹은 1/4의 핵심 대역 하향 표본화 비를 사용하면 저전송률에서 부호화 성능을 향상시킬 수 있음을 보였다.

Abstract

This paper deals with the effect of the internal sampling frequency and core band down sampling ratio on the overall performance of the MPEG USAC. Here, the internal sampling frequency is the sampling frequency of a signal actually coded. The core band down sampling ratio is the ratio of the width of the core band over that of the coded band. The performance was measured on 6 different test sound sources by the MUSHRA test with 10 subjects. The experiments showed that 1/3 or 1/4 core band down sampling ratio could yield the better performance than the conventional 1/2 ratio, especially at low rates.

Keyword : down-sampling ratio of core band, USAC, internal sampling frequency

1. 서 론

휴대용 멀티미디어 및 통신 기기의 보급이 확대됨에 따

라 음성 혹은 오디오 전송의 중요성이 강조되고 있다. 음성 코덱의 대역폭이 넓어지고, 오디오 코덱의 전송률이 낮아짐에 따라 두 코덱들이 처리하는 신호의 대역폭 및 전송률의 공통 영역이 발생하게 되었다. 이에 따라 오디오와 음성을 통합적으로 처리할 수 있는 오디오/음성 통합코덱(USAC: unified speech and audio codec)^{[1][2]}에 대한 필요성이 자연스럽게 대두되었고, 2011년 표준^[3]이 제정되었다.

MPEG USAC은 12-64Kbps의 전송률에서 기존의 AMR-WB^[4] 및 HE-AAC v2^[5]에 포함된 기술들을 확장 및 수정

a) 서울시립대학교 전자전기컴퓨터공학부(Dept. of ECE, Univ. of Seoul)

† Corresponding Author : 김인철 (Rin Chul Kim)

E-mail: rin@uos.ac.kr

Tel: +82-2-2210-2661 Fax: +82-2-2249-6802

※ 이 논문은 2011년도 서울시립대학교 연구년교수 연구비에 의하여 연구되었음.

· Manuscript received November 30, 2012 Revised January 14, 2013

Accepted January 14, 2013

하여 향상된 성능을 얻도록 제정한 표준이다. MPEG USAC에서는 주파수 대역별로 부호화가 수행된다. 먼저, 저주파 대역 즉 핵심 대역은 AAC (advanced audio coding), ACELP (algebraic code-excited linear prediction) 혹은 TCX (transform coded excitation)^[6]를 선택적으로 적용하여 부호화한다. 핵심 대역을 부호화하는 방법들은 시간 혹은 주파수 영역에서 잉여 정보를 감축한 다음 양자화 하는 것이다. 따라서 핵심 대역 부호화는 비트량과 부호화 성능이 밀접하게 연관되기 때문에 일반적으로 많은 비트량이 요구된다. 반면에 고주파 대역 혹은 확장 대역은 SBR (spectral band replication)^[5] 혹은 eSBR^[3]을 통해 부호화한다. 그리고 다채널 확장을 위해 MPEG surround 기법^[7]을 적용한다. 확장 대역 부호화와 다채널 부호화 방법들은 이미 부호화된 저주파 대역 혹은 모노 신호를 참조하여 적은 비트량의 파라미터만으로 신호를 표현한다.

MPEG USAC 부호화 성능은 핵심 대역과 확장 대역의 합으로 구성되는 부호화 대역의 폭에 의해 영향을 받는다. 부호화 대역이 넓은 경우 공간감이 향상되고 풍부한 음으로 재현된다. 그러나 저 전송률에서는 부호화 대역, 특히 부호화에 많은 비트량이 요구되는 핵심 대역이 넓으면 부호화 충실도가 크게 저하된다. 따라서 저 전송률에서 효율적인 부호화를 위해서는 부호화 대역 혹은 핵심 대역의 폭을 줄여야 한다. 이러한 목적에 부합되는 방법으로는 내부 표본화 주파수 조정과 핵심 대역 하향 표본화 비 (core-band down sampling ratio) 조정 등이 있다. 첫 번째로 내부 표본화 주파수는 입력 신호의 표본화 주파수와는 달리 실제 부호화되는 신호의 표본화 주파수를 말한다. 내부 표본화 주파수로의 변환은 부호화기 앞단에 재표본화기 (resampler)를 설치함으로써 수행된다. 내부 표본화 주파수를 낮추면 일정한 전송률에서 단위 샘플당 비트량이 증가하므로, 좁은 부호화 대역폭에서 부호화 충실도를 높일 수 있다. 두 번째로 핵심 대역은 1/2 하향 표본화를 통해 얻는데, 이 하향 표본화 비를 1/2 이외에 1/3, 1/4로 다양화하는 것이다. 핵심 대역 하향 표본화 비를 1/3 혹은 1/4로 설정하면 기존의 1/2로 설정하는 것에 비해 핵심 대역폭이 고정되어 있다고 할 때 부호화 대역폭을 크게 늘릴 수 있게 된다. 혹은 부호화 대역폭이

고정되어 있다고 하면 핵심 대역폭을 줄여 부호화에 필요한 비트량을 감축할 수 있다.

본 논문에서는 내부 표본화 주파수 및 핵심 대역 하향 표본화 비를 조정하여 핵심 대역폭 및 부호화 대역폭을 변경하였을 때 모노 신호에 대한 MPEG USAC의 부호화 성능을 평가한다. 본 논문에서 핵심 부호화기로는 ACELP와 TCX를 선택적으로 사용하는 LPD (linear prediction domain) 장치를 사용한다. 고주파 대역 부호화기로는 SBR 부호화기를 사용한다. 성능 평가는 2가지 실험으로 진행된다. 첫 번째 실험은 부호화 대역폭을 12KHz로 고정한 후 핵심 대역 표본화 비를 변경하여 3KHz, 4KHz 그리고 6KHz의 핵심 대역폭을 가질 때 부호화 성능을 비교한다. 두 번째 실험은 핵심 대역을 모두 6KHz로 고정한 후 핵심 대역 표본화 비 및 내부 표본화 주파수를 조정하여 전체 부호화 대역폭이 12KHz 18KHz 그리고 24KHz일 때 부호화 성능을 비교한다. 2가지 실험 모두 MUSHRA 테스트를 통해 그 성능들을 비교한다. 성능 비교는 서로 다른 특성을 가진 6개의 실험 오디오 신호를 대상으로 하였고, 10명의 테스트에 익숙한 사람들이 그 성능을 평가하였다.

서론에 이어 2절에서는 MPEG USAC 부호화기를 설명한다. 3절에서는 내부 표본화 주파수 및 핵심 대역 하향 표본화 비 조정에 관해 설명한다. 그리고 4절과 5절에서는 각각 실험 결과와 본 논문의 결론을 제시한다.

II. MPEG USAC

MPEG USAC은 기존의 음성 및 오디오 부호화 기술을 효과적으로 융합하고 선택적으로 사용하는 방법이다. MPEG USAC의 기본적인 동작을 그림 1에 제시한 블록선도를 참조하여 소개하면 다음과 같다. 먼저 스테레오 신호가 USAC 부호화기에 입력되면 MPEG surround에서 공간 파라미터를 추출하고, 입력 신호를 모노 신호로 다운 믹스한다. SBR에서는 다운 믹스된 모노 신호를 분석하여 SBR 파라미터를 추출하고 1/2 하향 표본화된 신호를 출력한다. 이 하향 표본화된 신호는 LPD 장치에 의해 선형예측 영역 혹은 FD (frequency domain) 장치에 의해 주파수 영역으로

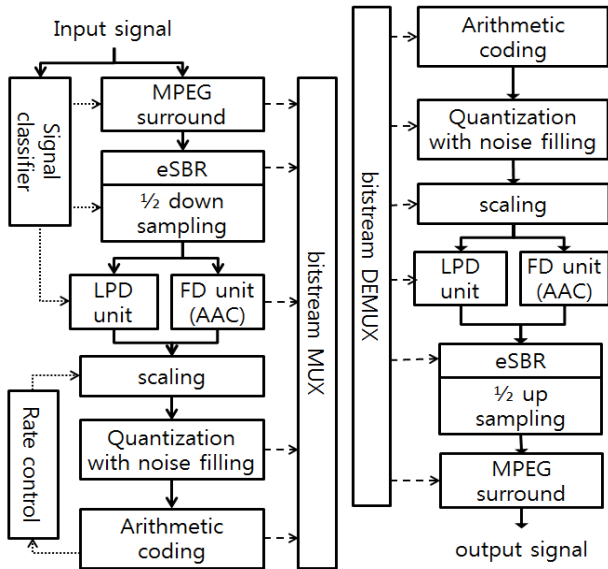


그림 1. MPEG USAC의 블록선도
Fig. 1. Block diagram of the MPEG USAC

변환된다. 변환된 값들은 축척, 양자화 및 산술 부호화 과정을 거쳐 부호화된다. LPD 혹은 FD 장치는 신호 분류기에서 분석된 신호의 특성에 따라 선택된다. 각 부분을 구체적으로 살펴보면 다음과 같다.

먼저 FD유닛에서는 MDCT를 통하여 시간 영역의 신호를 주파수 영역으로 변환한다. 그런 다음, TNS를 이용하여 양자화 잡음을 신호의 포락선에 비례하도록 만든다. 마지막으로 심리음향 모델을 참조하여 축척을 조정, 양자화, 산술 부호화 등을 거쳐 변환 계수들을 전송한다.

ACELP 혹은 TCX를 사용하는 LPD 장치는 선형예측을 기반으로 하여 부호화를 진행한다. 먼저 선 강조를 통해 고주파가 강조된 모노 신호에 대해 선형예측을 수행하며 그 결과 얻어진 여기신호(residual signal)를 TCX와 ACELP를 선택적으로 사용하여 부호화 한다. TCX는 혼합신호 및 저 전송률 오디오 신호를 주 처리 대상으로 한다. TCX에서는 먼저 MDCT를 통해 여기 신호를 변환 계수로 변환한다. 변환 계수는 축척, 양자화 및 산술 부호화를 거쳐 전송된다. 또한 양자화 과정에서 발생하는 오차의 크기를 잡음 지수로 전송한다. 이 잡음 지수는 복호화기에서 배경 잡음을 생성하는 데에 이용된다. USAC에서 사용하는 TCX를 AMR-

WB+의 TCX와 비교하면, DFT 대신 MDCT를 사용하고, AVQ (algebraic vector quantizer) 대신에 스칼라 양자기와 산술부호화를 사용한다는 점에서 크게 구별할 수 있다.

ACELP는 음성신호를 효과적으로 부호화하기 위한 기술로써 USAC에서는 다음과 같은 과정으로 수행된다. ACELP에서 음성 신호는 선형예측 계수, 피치, 그리고 여기 신호로 모델링된다. 먼저 선형 예측을 통해 선형 예측 계수를 얻고 그 계수를 LSP (line spectral pair)로 표현하고 AVQ 부호화한다. 그리고 피치 정보는 가변길이 스칼라 양자화되고, 여기신호는 이득 VQ 및 대수 부호화를 이용하여 양자화된다.

일반적으로 LPD 혹은 FD 장치를 선택적으로 사용하는 부호화기를 핵심 부호화기 (core codec)라고 하고, 핵심 부호화기에서 처리되는 저주파 대역을 핵심 대역 (core band)이라고 한다. SBR 부호화기^[8]에서는 이 핵심 대역을 참조하여 고주파 대역을 복원할 수 있도록 SBR 파라미터를 추출한다. 본 논문에서는 SBR 파라미터를 이용하여 복원되는 고주파 대역을 확장 대역 (extended band)라고 하겠다. 확장 대역과 핵심 대역을 결합하면 연속된 하나의 대역, 즉 USAC이 부호화하는 전체 부호화 대역 (coded band)을 이룬다.

부호화기에서 얻어진 비트열은 복호화기에 의해 부호화의 역 과정을 통해 원 신호로 복원된다. 먼저 비트열의 일부는 산술복호화, 역양자화, 축척 등을 거친 다음 LPD 혹은 FD 장치를 통해 핵심 대역 신호로 복원된다. 복원된 핵심 대역 신호는 SBR 파라미터들과 함께 SBR 복호화기에 의해 전체 신호로 확장된다. 마지막으로 수신된 공간 파라미터를 이용하여 MPEG surround 복호화기에서 스테레오 신호로 복원된다.

III. 부호화 대역 조정

MPEG USAC에서 전체 부호화 대역은 핵심 대역과 확장 대역으로 나뉘어 부호화된다. 핵심 대역은 ACELP, TCX, 혹은 AAC로 부호화되고, 확장 대역은 SBR 부호화기로 부호화된다. MPEG USAC에서 SBR 부호화는 SBR 파라미

터만으로 표현되므로 적은 비트 량이 필요하다. 반면에 핵심 대역 부호화는 충실한 부호화를 위해 매우 많은 비트 량이 필요하여 USAC의 대부분의 비트 량이 핵심 대역 부호화에서 사용된다.

부호화 및 복호화를 거쳐 복원된 음질은 효율적 인영 정보 감축 및 낮은 양자화 잡음 등으로 성취되는 부호화의 충실도뿐만 아니라 전체 부호화 대역폭에 따라 영향을 받는다. 그러나 낮은 전송률에서는 넓은 부호화 대역폭에서 높은 부호화 충실도를 얻기 어렵다. 따라서 낮은 전송률에서는 부호화 대역폭을 줄여 좁은 대역폭에 대해 충실히 부호화할 수 있도록 한다. 본 절에서는 전체 부호화 대역폭을 줄이는 방법인 내부 표본화 주파수 조정에 대해 먼저 살펴본다. 그런 다음, 전체 부호화 대역폭은 어느 정도 유지하면서 많은 비트 량이 요구되는 핵심 대역폭을 감축시킬 수 있는 핵심 대역 하향 표본화 비 조정에 대해 살펴보겠다.

1. 내부 표본화 주파수 설정

MPEG USAC에서는 7.35부터 96KHz까지 다양한 표본화 주파수를 지원하고 있다. 표본화 주파수가 높으면 신호의 대역폭이 넓어져 풍부한 음을 표현할 수 있게 된다. 표본화 주파수가 낮으면 표현할 수 있는 대역폭은 좁아지나, 전송률이 일정할 때 단위 샘플당 더 많은 비트 량을 할당할 수 있어 주어진 대역폭을 더 충실히 부호화할 수 있다는 장점이 있다. 예를 들어 48KHz의 표본화 주파수를 갖는 음원의 경우 낮은 비트 전송률에서 부호화 하면 매우 큰 양자

화 오차를 보인다. 만약 신호의 표본화 주파수를 36KHz 혹은 24KHz로 낮추면 신호의 대역폭은 각각 18 혹은 12 KHz로 줄어들어 음의 풍부성이 감소된다. 그러나 단위 샘플당 비트 량은 각각 1.5배 혹은 2배 증가하므로 양자화 오차를 줄여 좁아진 대역폭을 충실히 부호화할 수 있다. 이러한 점을 참조한다면 높은 전송률과는 달리 낮은 전송률에서는 신호의 표본화 주파수를 낮추면 보다 충실한 부호화를 기대할 수 있을 것이다.

이러한 점을 참조하여 부호화기의 앞단에 신호의 표본화 주파수를 바꾸는 재표본화기(resampler)를 설치한다. 재표본화기는 입력 음원의 특성 및 전송률 등을 참조하여 입력의 표본화 주파수를 부호화에 적합한 표본화 주파수로 변환한다. 이때 변환된 주파수를 부호화 내부에서 사용되는 표본화 주파수이므로 AMR-WB+[4]에서와 칭한 바와 마찬가지로 내부 표본화 주파수 (ISF; internal sampling frequency)라고 부르겠다.

2. 핵심 대역 하향 표본화 비의 조정

MPEG USAC 부호화기에서 대역별 신호처리 방법은 그림 2에 제시한 바와 같다. 먼저 입력 신호를 내부 표본화 주파수로 재표본화한 다음, 핵심 대역과 확장 대역으로 분리하여 처리한다. 핵심 대역은 재표본화된 신호를 1/2 하향 표본화(down sampling)하여 얻고, 이는 핵심 부호화기에 의해 부호화된다. 반면에 확장 대역은 핵심 대역에 속하지 않는 고주파 대역으로 SBR 부호화기에 의해 부호화된다.

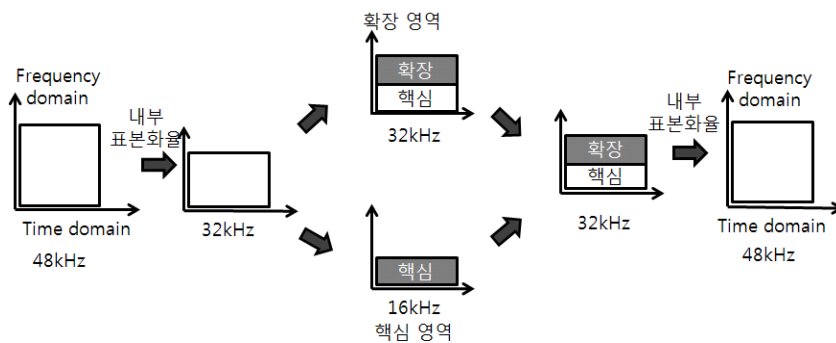


그림 2. MPEG USAC에서의 신호 처리
 Fig. 2. Signal processing in the MPEG USAC

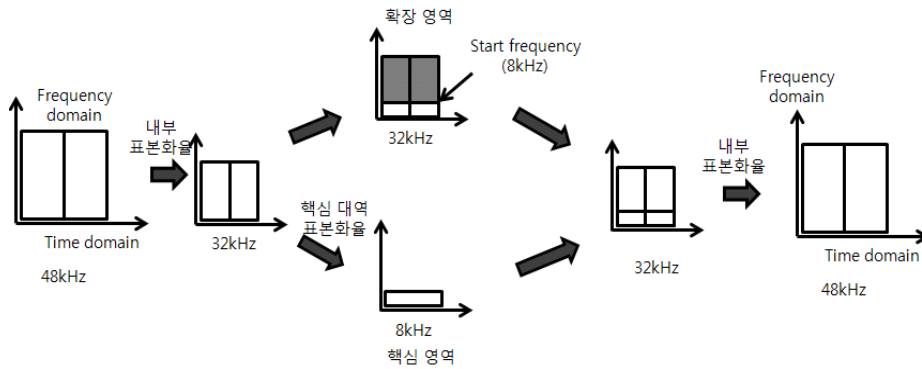


그림 3. 가변 핵심대역 표본화 주파수 적용 시 신호처리
 Fig. 3. Signal processing when variable core band sampling frequency is employed

복호화기에서는 핵심 복호기를 통해 전송된 비트열로부터 핵심 대역을 복원한다. 그런 다음 SBR 복호기에서 복원된 핵심 대역과 전송된 SBR 파라미터를 이용하여 확장 대역을 복원한다. 복원된 두 대역 신호를 합쳐서 전체 대역 신호로 복원되고, 최종적으로 내부 표본화 주파수를 입력 신호의 표본화 주파수로 재표본화하여 원 신호로 복원한다.

이러한 방식에서 핵심 대역은 LPD 혹은 FD 장치 등 핵심 부호화기에서 효율적인 부호화를 수행함에도 불구하고 많은 비트 량이 요구된다. 반면에 확장 대역은 적은 비트 량의 SBR 파라미터만으로 표현할 수 있다^[8]. 낮은 전송률에서 충실한 부호화 결과를 얻으려면 3.1절에서 설명한 바와 같이 내부 표본화 주파수를 낮추면 된다. 내부 표본화 주파수를 낮추면 단위 시간당 비트 량이 많아져 부호화 충실도가 높아지지만 전체 부호화 대역폭은 좁아져 복호화된 음의 풍부성은 약화된다.

이러한 단점들을 해결할 수 있는 한 방법은 핵심 대역의 하향 표본화 비를 기존의 1/2뿐만 아니라 1/3, 혹은 1/4로 다양화하는 것이다. 그림 3은 핵심 대역 하향 표본화 비를 기존의 1/2에서 1/4로 변경하였을 때의 신호처리 방식을 나타내었다. 그림 3과 같이 핵심 대역 하향 표본화 비를 1/4로 하면, 전체 부호화 대역폭은 동일하게 유지하면서 핵심 대역의 대역폭을 4KHz로 작게 만들 수 있다. 이때 SBR 부호화기가 담당하는 확장 대역폭은 4~16KHz로 넓어지는데, 이는 매우 작은 비트 량의 SBR 파라미터만으로 표현할 수 있다. 핵심 대역을 임계 표본화 (critical sampling)한다고

할 때, 핵심 대역 표본화 비를 1/2에서 1/4로 변경하면 핵심 대역의 단위 시간 당 샘플 수는 1/2로 줄게 된다. 따라서 주어진 전송률에서 핵심대역 표본화 비를 1/2에서 1/4로 변경하면 핵심 대역 샘플당 비트 량은 약 2배 증가한다. 이는 핵심 대역폭은 감축되지만, 감축된 핵심 대역을 보다 충실히 부호화할 수 있음을 의미한다.

표 1에서는 한 예로 내부 표본화 주파수가 48, 36, 24KHz일 때, 각 내부 표본화 주파수에서 핵심 대역 하향 표본화 비를 1/2, 1/3, 1/4로 하였을 때 핵심 대역 및 확장 대역의 크기를 제시하였다. 실제로는 표 1에 제시한 예 이

표 1. 내부 표본화 주파수와 핵심 대역 하향 표본화 비에서의 대역폭의 예
 Table 1. Examples of bandwidths under several ISFs and core band down sampling ratios

ISF (KHz)	핵심대역 하향 표본화 비	핵심대역 (KHz)	확장대역 (KHz)
48.0	1/2	0~12.0	12.0~24.0
	1/3	0~8.0	8.0~24.0
	1/4	0~6.0	6.0~24.0
36.0	1/2	0~9.0	9.0~18.0
	1/3	0~6.0	6.0~18.0
	1/4	0~4.5	4.5~18.0
24.0	1/2	0~6.0	6.0~12.0
	1/3	0~4.0	4.0~12.0
	1/4	0~3.0	3.0~12.0

외에도 다양한 내부 표본화 주파수에 대해 3가지 핵심 대역 하향 표본화 비율 적용할 수 있다. 표 1을 참조하면 핵심 대역과 확장 대역을 포함하는 전체 부호화 대역폭은 임계 표본화에 의해 내부 표본화 주파수의 1/2로 결정된다. 이때 핵심 대역폭은 전체 부호화 대역에서 핵심 대역 하향 표본화 비만큼 감축되어 설정된다. 나머지 고주파 대역은 확장 대역으로 설정된다. 예를 들어 내부 표본화 주파수가 36KHz이면 전체 부호화 대역폭은 18KHz이다. 만약 핵심 대역 표본화 비가 1/3이라면 핵심 대역은 0~6KHz이고 나머지 고주파 대역 6~18KHz는 확장 대역으로 설정된다. 이 경우 핵심 대역 대역폭은 6KHz이므로 프레임당 비트량을 높일 수 있어 핵심 대역 부호화를 충실히 수행할 수 있다. 또한, 확장 대역이 넓은 주파수 범위를 수용하여 전체 부호화 대역폭은 18KHz까지 넓힐 수 있어 음의 풍부성을 확보할 수 있다는 장점이 있다. 여기서, 핵심 대역 표본화 비를 1/2로 유지하면서 max_sfb 등의 비트열 요소를 통해 핵심 대역폭을 6KHz로 유지하는 방법도 생각할 수 있다. 그러나 이 방법은 프레임 당 비트량이 앞의 방법에 비해 2/3으로 축소되므로 핵심 대역 부호화 성능이 낮아지고, 따라서 전체 부호화 성능이 저하된다. 따라서 핵심 대역 표본화 비를 1/2 이 아니라 1/3 혹은 1/4로 설정하는 것은 저전송률에서 우수한 성능을 보일 것으로 판단된다.

핵심 대역 하향 표본화 비를 다양화하는 경우 다음과 같은 장점이 있다. 첫째, 확장 대역 부호화기가 작은 비트 량

으로 넓은 대역을 부호화하기 때문에 낮은 비트 전송률에서도 부호화 대역을 넓힐 수 있다. 둘째, 핵심 대역의 하향 표본화 비를 낮춤으로써 핵심 대역 부호화 시 프레임 당 비트 수를 증가시킬 수 있어 보다 충실한 핵심 대역 부호화가 가능하다. 또한 SBR의 특징에 의해 충실한 핵심 대역 복호화는 확장 대역 신호의 복호화의 충실도 향상을 가능하게 한다.

IV. 실험 및 결과

본 절에서는 핵심대역폭과 확장대역폭 변화에 따른 MPEG USAC의 성능을 주관적 품질 측면에서 평가한 결과를 제시한다. 본 실험에서 핵심 부호화기로는 ACELP와 TCX를 선택적으로 사용하는 LPD 장치를 사용한다. 고대역 부호화기로는 SBR 부호화기를 사용한다. 본 실험에서는 6개의 모노 실험 음원들을 대상으로 12와 16Kbps라는 낮은 전송률에서 성능을 비교하였다. 실험 음원 중에서는 es01, HarryPotter, louis는 사람 목소리로 구성된 음성적 특성을 가진다. 반면에 phi7, Music_1, Music_3은 악기 연주로 구성되어 음향적 특성을 가진다. 부호화 성능은 대표적인 주관적 음질 평가 방법인 MUSHRA 검사 방법으로 10 명이 참여하여 평가하였다. 평가자들은 MUSHRA 검사에 대해 경험이 있는 4명의 전자전기 전공자와 6명의 비전공

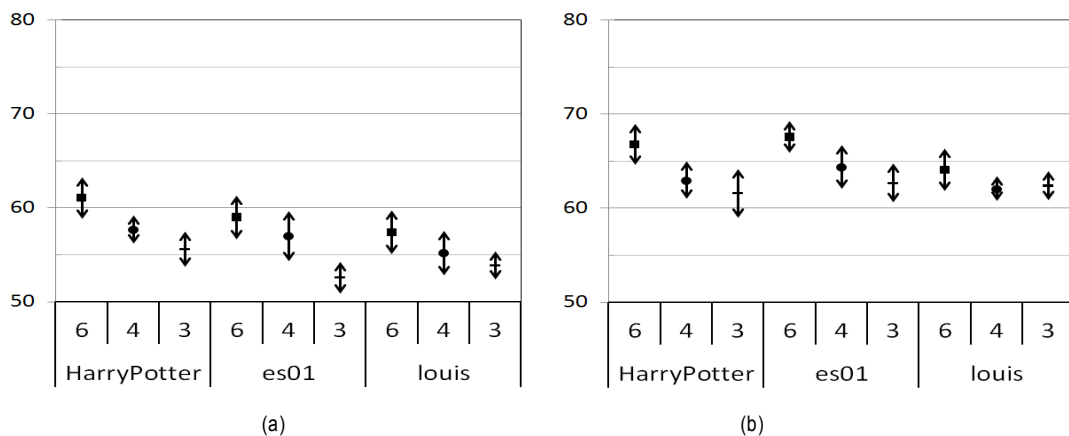


그림 4. 전송률이 (a) 12Kbps, (b) 16Kbps 일 때 음성적 특성을 가진 파일에 대한 실험 1의 MUSHRA 검사 결과
 Fig. 4. Result of the MUSHRA test of the experiment 1 on the sound sources with voice at (a) 12Kbps and (b) 16Kbps

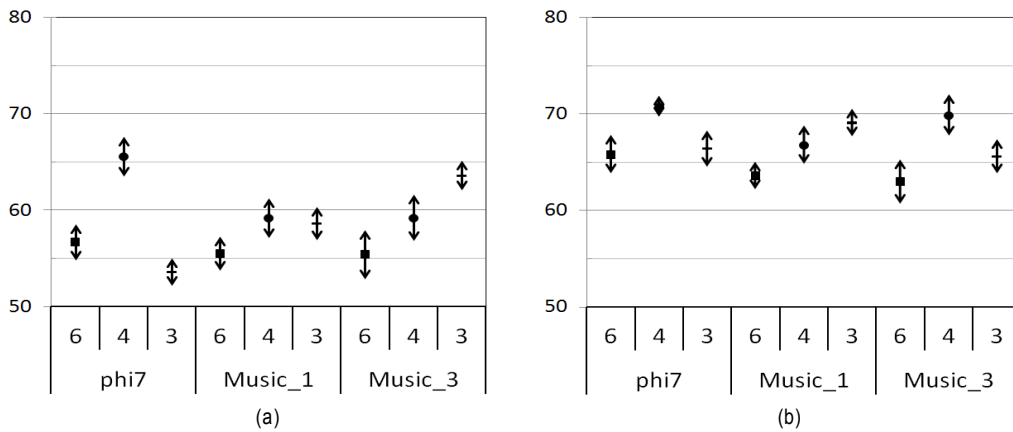


그림 5. 전송률이 (a) 12Kbps, (b) 16Kbps 일 때 음향적 특성을 가진 파일에 대한 실험 1의 MUSHRA 검사 결과
 Fig. 5. Result of the MUSHRA test of the experiment 1 on the sound sources with instrumental sound at (a) 12Kbps and (b) 16Kbps

자들로 구성되었다. 주관적 음질 평가 결과는 음질 평가의 경험과 숙련도에 따라 달라질 수 있다. 따라서 실험은 음질을 평가한 경험이 있는 사람들과 함께 진행하였다.

본 논문에서는 2가지 실험을 수행하였다. 첫 번째 실험에서는 전체 부호화 대역폭을 12KHz로 고정시킨 상태에서 핵심 대역폭을 가변시켰을 때 부호화 성능을 비교하였다. 두 번째 실험에서는 핵심부호화 대역폭을 6KHz로 고정시킨 상태에서 전체 부호화 대역폭을 가변시켰을 때 부호화 성능을 비교하였다. 두 실험에서 주어진 핵심 대역폭 및 확장 대역폭은 표 2를 통해 그에 맞는 내부 표본화 주파수 및 핵심 대역 하향 표본화 비로 대응된다.

그림 4과 5에는 첫 번째 실험을 수행하였을 때 각각 음성적 특성을 가진 파일들과 음향적 특성을 가진 파일들의 성능을 비교한 결과를 제시하였다. 첫 번째 실험에서는 내부주파수를 24KHz로 고정시키고, 핵심 대역 하향 표본화 비를 1/2, 1/3, 1/4로 변화시켜 핵심 대역폭을 각각 6, 4, 3KHz로 조정하였다. 여기서 핵심 대역폭에 포함되지 않는 12KHz까지의 고주파 대역은 전부 SBR에 의해 확장된다.

먼저 전송률이 12 및 16Kbps일 때, 음성적인 특성을 가진 es01, HarryPotter, louis에 대한 첫 번째 실험 결과를 그림 4에 제시하였다. 그림 4를 참조하면, 3개의 음성파일 모두 핵심 대역이 넓을수록 주관적 품질이 높게 나타났다. 세

파일 모두 음성적 특성을 가지고 있어 핵심 부호화기로 TCX 보다는 ACELP가 주로 선택되었다. 그리고, 핵심 대역폭이 좁아지면 배경 잡음이 커지고, 음의 풍부성이 축소되는 경향을 보였다.

다음으로 음향적인 특성을 보이는 Phi7, Music_1, Music_3에 대한 두 번째 실험 결과를 그림 5에 제시하였다. 3개의 실험 파일들 모두 ACELP 보다 TCX가 핵심 부호화 방법으로 주로 선택되는 특성을 보인다. 그림 5를 살펴보면, Phi7의 경우 12 및 16Kbps 등 두 가지 전송률에서 핵심 대역이 4KHz일 때 가장 높은 평가를 얻었다. Music_1과 Music_3의 경우에는 핵심 대역폭이 3 혹은 4KHz일 때가 핵심 대역폭이 6KHz일 때보다 좋은 성능을 보였다. 핵심 대역폭이 축소되면 임계 표본화에 의해 단위 샘플당 비트량이 많아진다. 이에 따라 양자화 오차가 감소하고, spectral hole이 사라져서 저주파 대역을 효과적으로 부호화할 수 있다. 그리고 SBR 부호화를 통해 확장 대역 또한 비교적 충실히 표현되고 있다. 따라서 음향적 특성을 가진 파일에 대해서는 핵심 대역폭이 3 혹은 4 KHz일 때 성능이 우수함을 알 수 있다.

두 번째 실험에서는 내부 주파수와 핵심대역 하향 표본화 비를 조정하여 핵심 대역폭은 6KHz로 동일하게 유지하면서 확장 대역폭을 가변시켜 전체 대역폭이 12, 18, 24KHz가 되도록 한다. 이 3가지 부호화기에 대한

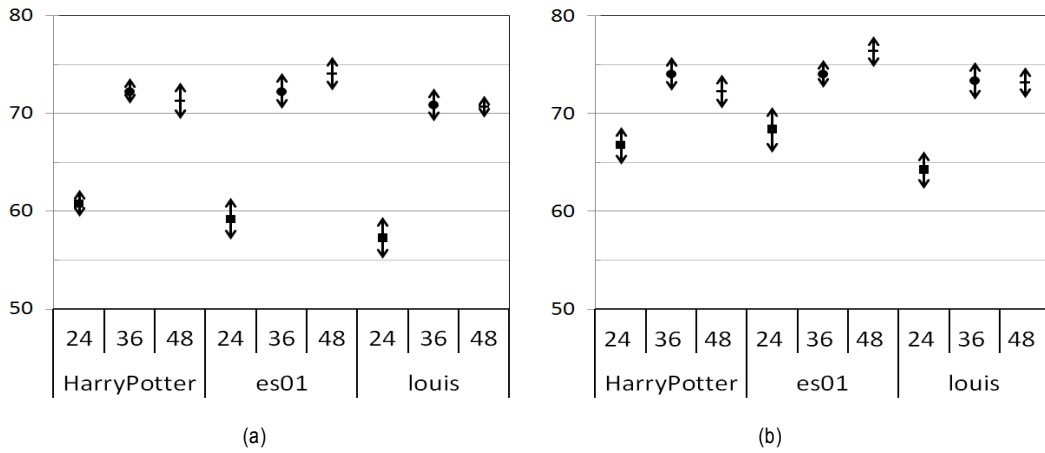


그림 6. 전승률이 (a) 12Kbps, (b) 16Kbps 일 때 음성적 특성을 가진 파일에 대한 실험 2의 MUSHRA 검사 결과.
 Fig. 6. Result of the MUSHRA test of the experiment 2 on the sound sources with voice at (a) 12Kbps and (b) 16Kbps.

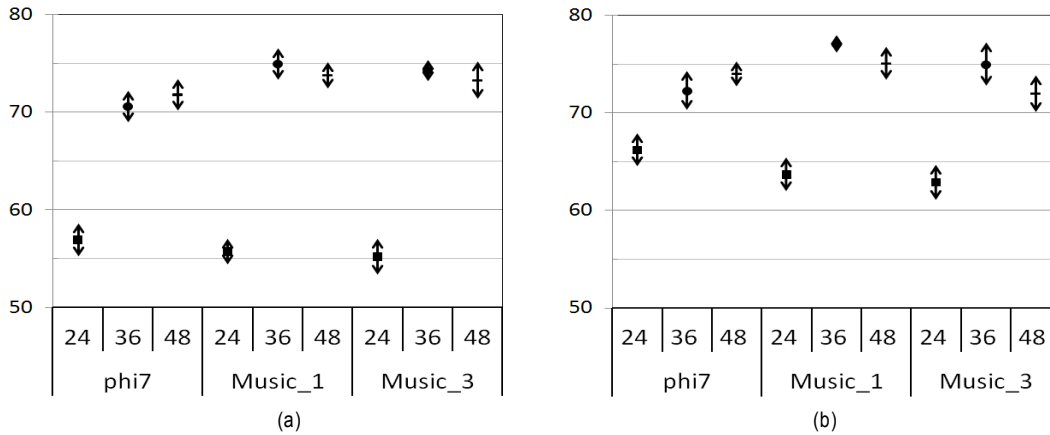


그림 7. 전승률이 (a) 12Kbps, (b) 16Kbps 일 때 음향적 특성을 가진 파일에 대한 실험 2의 MUSHRA 검사 결과
 Fig. 7. Result of the MUSHRA test of the experiment 2 on the sound sources with instrumental sound at (a) 12Kbps and (b) 16Kbps

MUSHRA 평가 결과는 그림 6과 7에 제시하였다.

그림 6에는 음성적인 특성을 가진 es01, HarryPotter, louis에 대한 성능 평가 결과를 제시하였다. 여기서 세 파일 모두 부호화 대역폭이 18 혹은 24KHz일 때가 12KHz일 때 보다 우수한 평가 결과를 보였다. 특히 고주파 대역에 존재하는 파일음 등의 특징들이 부호화 대역폭이 12KHz일 때는 나타나지 않지만 18KHz 이상일 때는 확인할 수 있었다.

그림 7에는 음향적 특성을 보이는 Phi7, Music_1, Music_3에 대한 주관적 음질 평가 결과를 제시하였다. 그림 7을 참조하면, 세 음원 모두 부호화 대역폭이 18 혹은

24KHz일 때가 12KHz일 때보다 더 우수한 음질을 보여주었다. 즉, 세 파일 모두 고주파 영역에 존재하는 소리가 부호화 대역폭이 12KHz일 때는 나타나지 않지만 18 혹은 24KHz일 때는 나타난다. 이에 따라 부호화 대역폭이 18KHz 이상에서는 음질이 풍부해지고 보다 정확해짐을 알 수 있었다. 한편, 부호화 대역폭이 18KHz일 때와 24KHz일 때의 성능은 6개의 음원에 대해 모두 주관적 음질이 유사한 것으로 나타났다. 그 이유는 두 경우 모두 피 실험자의 가청 주파수 대역을 대부분 포함하고 있어 두 경우의 음질 차이를 거의 분별할 수 없기 때문으로 판단된다.

V. 결론

본 논문에서는 MPEG USAC 부호화 기법에서 핵심 대역폭 및 확장 대역폭에 따른 주관적 음질을 평가하였다. 첫 번째로 전체 부호화 대역폭을 12KHz로 고정하였을 때, 음성적 특성을 가진 음원에 대해서는 핵심 대역폭이 넓을수록 우수한 성능을 보였다. 반면에 음향적 특성을 가진 음원에 대해서는 핵심 대역폭이 3 혹은 4KHz일 때 보다 우수한 성능을 보였다. 이를 통해 음성적 특성과 음향적 특성을 가진 파일들을 분류하여 부호화한다면, 더 효과적인 부호화가 가능함을 알 수 있다. 두 번째로 핵심 대역폭을 6KHz로 고정하였을 때는 음성 및 음향적 특성을 가진 파일 모두 부호화 대역폭을 18KHz 이상으로 하였을 때 우수한 성능을 보임을 알 수 있었다.

본 논문에서 제시한 부호화 구조는 음원의 표본화 주파수를 내부 주파수로 변경하는 표본화 주파수 변화기를 추가로 설치함으로써 가능하다. 그리고 핵심대역 표본화를 변화는 MPEG USAC의 eSBR 4:1 구조를 통해 구현할 수 있다. 즉 본 논문에서의 결과는 표본화 주파수 변화기와 eSBR 4:1 구조가 MPEG USAC 부호화 성능을 향상시키는데 기여할 수 있음을 보였다. 또한, MPEG USAC에서 최적의 부호화 파라미터를 추출하는데 크게 기여할 것으로 기대된다.

참고 문헌

- [1] Taejin Lee, Kyeongok Kang, and Whanwoo Kim, "MPEG Audio New Standard: USAC Technology" JBE, Vol. 16, No. 5 pp. 705-713, Sept. 2011.
- [2] H.Park, young cheol Park, "MPEG audio standard technology: USAC," Korea Society Broadcast Engineers Magazine, Vol. 14, No. 2, pp. 63-74, June 2009.
- [3] ISO/IEC 23003-3:2011(E), "MPEG audio technologies-Part 3: Unified speech and audio coding," 2011.
- [4] J. Makinen, et al. "AMR-WB+: a new audio coding standard for 3rd generation mobile audio services," IEEE ICASSP, vol. 2, pp. 1109 - 1112, Mar. 2005.
- [5] ISO/IEC 14496-3:2008(E), "Coding of audio-visual objects - Part 3: Audio, Subpart 4: General audio coding (GA)-AAC, TwinVQ, BSAC," 2008.
- [6] B. Bessette, R. Salami, C. Laflamme, R. Lefebvre, "A wideband speech and audio codec at 16/24/32 kbits/s using hybrid ACELP/TCX techniques," IEEE Speech Coding Workshop, 20 - 23 June 1999, pp. 7- 9.
- [7] ISO/IEC FDIS 23003-1:2006(E), "MPEG audio technologies - Part 1: MPEG surround," 2006.
- [8] Kyung Jun An, Yoosun Jung, Seungkwon Beack, Kyeongok Kang, and Rin Chul Kim, "Performance evaluation of the MPEG USAC according to the spectral band replication bandwidth," JBE, Vol. 16, No. 5, pp. 705-713, Sept. 2011.

저 자 소 개



이 재 화

- 2010년 : 서울시립대학교 전자전기컴퓨터공학부 학사 졸업
- 2012년 : 서울시립대학교 전자전기컴퓨터공학부 석사 졸업
- 2012년 ~ 현재 : 삼성전자(주) 연구원
- 주관심분야 : GSM, LTE, 오디오 코덱



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