

AMR 기반 저 전력 인공 대역 확장 기술 개발

Developing a Low Power BWE Technique Based on the AMR Coder

구 본 강*, 박 희 완*, 주 연 재*, 강 상 원*

(Bonkang Koo*, Heewan Park*, Yeon-Jae Ju*, Sangwon Kang*)

*한양대학교 전자전기제어계측공학과

(접수일자: 2011년 2월 28일; 수정일자: 2011년 4월 27일; 채택일자: 2011년 5월 9일)

대역폭 확장 (Bandwidth Extension)은 300–3400 Hz 대역의 협대역 음성 신호를 50–7000 Hz 대역의 광대역 음성신호로 확장하여 협대역 음성신호의 음질과 명료도를 높이는 기술이다. 본 논문에서는 협대역 음성 정보만을 이용해서 광대역 음성신호를 추정하는 인공 대역폭 확장 기술을 설계하여, ITU-T 협대역 표준 음성 코덱인 AMR (adaptive multi-rate) 복호화기에 내장시킴 (embedded)으로써, 대역폭 확장 모듈에서의 LPC 분석 및 LSP 해석과 관련된 계산량을 감소시켰고, 알고리즘 지연도 줄였다. 그리고 SDS (single distance search) 고속 탐색 방식을 대역폭 확장 시스템의 코드북 매핑에 적용하여, 최종적으로 저 전력 대역 확장 AMR 복호화기를 설계하였다. 제안된 대역폭 확장 방법은 AMR 복호화기 후단에 독립적으로 설치되는 기존 DTE (decode then extend)방식에 비해 28 % 정도의 계산량을 줄이고 알고리즘 지연도 20 msec 줄였다. 또한 제안방식은 피치정보를 이용한 classified 코드북 매핑 방식을 사용하여 스펙트럼 포락선을 확장 하였고, 코드 벡터 탐색 시 가중치를 적용하여 광대역 합성 음성의 성능을 향상시켰다.

핵심용어: 대역폭 확장, 음성코덱, 음질 개선, AMR 부호화기

투고분야: 음성처리 분야 (2,3)

Bandwidth extension is a technique to improve speech quality and intelligibility, extending from 300–3400 Hz narrowband speech to 50–7000 Hz wideband speech. This paper designs an artificial bandwidth extension (ABE) module embedded in the AMR (adaptive multi-rate) decoder, reducing LPC/LSP analysis and algorithm delay of the ABE module. We also introduce a fast search codebook mapping method for ABE, and design a low power BWE technique based on the AMR decoder. The proposed ABE method reduces the computational complexity and the algorithm delay, respectively, by 28 % and 20 msec, compared to the traditional DTE (decode then extend) method. We also introduce a weighted classified codebook mapping method for constructing the spectral envelope of the wideband speech signal.

Keywords: Bandwidth Extension, Speech Codec, Speech Enhancement, AMR Codec

ASK subject classification: Speech Signal Processing (2,3)

I. 서론

기존의 전화망과 이동 통신망의 음성은 300–3400 Hz 범위의 협대역 음성 신호로 제한되어져 있다 [1]. 이러한 협대역 음성신호는 50–7000 Hz 범위의 광대역 음성에 비해 이해력과 명료도 그리고 자연성이 부족 하다 [2–3]. 광대역 음성 코딩 표준안이 유용함에도 불구하고, 광대역 음성코덱은 원격회의와 VoIP (voice over IP)와 같은

응용 분야를 제외하면 널리 사용되고 있지 않다. 이동 통신 시스템에서, 광대역 음성 코덱 제공으로 인한 음질 향상 효과는 현재의 이동 통신을 사용하는 사람들에게 보다 나은 통신 서비스를 제공할 수 있다. 하지만 협대역에서 광대역 전화통신으로의 변화는 오랜 시간이 걸릴 것으로 예상되며, 이 기간 동안에 사용자들은 광대역과 협대역 음성 간의 상당한 음질 차이를 경험할 것이다. 변화 시기 동안에, 광대역과 협대역 음성 간의 음질 간극은 협대역 음성의 유용한 정보를 이용하여 고대역 (high-band) 음성신호를 추정 및 복원하고, 고대역에 대한 정보 없이 대역폭을 인공적으로 확장하는 인공 대역폭 확장 (ABE -

artificial bandwidth extension) 기술 [2]에 의해 감소될 수 있다. 음성신호의 대역폭 확장에 관한 대부분의 알고리즘들은 소스-필터 모델 [4]에 기초한다. 여기서 소스-필터 모델은 성도 필터와 여기 신호로 모델링 되는 사람의 음성 생성 메카니즘을 의미한다. 따라서, 대역폭 확장 문제는 성도 필터 확장 (즉, 스펙트럼 포락선 확장)과 여기 신호 확장으로 나눌 수 있다 [5]. 스펙트럼 포락선 확장에는 코드북 매핑 방식 [5], HMM (hidden markov model) [2], GMM (Gaussian mixed model) 방식 [6] 등이 있고, 여기신호의 확장에는 스펙트럼 folding이나 스펙트럼 shifting, 하모닉과 노이즈 모델링 등의 방법들이 있다 [7]. 이러한 대역확장 시스템을 AMR 디코더 출력에 독립적으로 위치시켜 AMR 협대역 음성 출력을 광대역 음성으로 확장하는 기술을 DTE (decode then extend) 방식이라 한다. DTE 방식은 대역확장 모듈에 대한 추가적인 알고리즘 지연과 계산량 증가를 필요로 한다. 본 논문에서는, AMR 음성 디코더에 대역 확장 시스템을 내장함 (embedded)으로써 기존의 DTE 방식에서 요구되는 알고리즘 지연을 없애고 계산량을 줄였다. 또한 대역확장 모듈 내 코드북 매핑 방식에 고속 코드북 탐색 방법을 도입하여 스펙트럼 확장 시 요구되는 계산량을 줄였다.

본 논문은, 2장에서 ABE 시스템이 내장된 AMR 디코더의 전체 구조에 대해 살펴보고, 성능 개선을 위해 가중치 적용 classified 코드북 매핑 방법과 후처리 필터를 설계한다. 3장에서 SDS 방식과 LSP 순서성질을 이용한 고속 탐색 코드북 매핑 방법을 제안한다. 4장에서 실험 및 성능 분석 결과를 보이고, 마지막으로 5장에서 결론을 맺는다.

II. ABE 시스템이 내장된 AMR 디코더 설계

2.1. 개요

본 논문에서 제안하는 AMR 기반 ABE 시스템은 AMR 디코더 내에 ABE 시스템을 내장 하였다. 그래서 AMR 디코더에서 복호화된 LSP계수와 여기신호를 ABE 시스템에 활용함으로써, ABE 시스템에서 요구되는 LPC 분석, LSP 변환 및 여기신호 발생과 관련된 계산량을 줄였다. ABE 시스템이 내장된 AMR 디코더의 구조는 그림 1과 같다. AMR 디코더는 수신된 ISP 인덱스를 사용하여 양자화된 ISP 벡터를 구성한다. 이러한 ISP 벡터들은 5 ms의 4개 부 프레임으로 보간된다. 각각의 부 프레임에서 보간된 ISP 벡터는 LP 필터를 통해 LPC로 변환되고 이를 이용하여 여기 신호를 생성한다.

ABE 모듈을 살펴보면, 먼저 AMR 디코더에서 복호화된 10차 협대역 LSP 계수를 사용하여 코드북 매핑 방식으로 18차 광대역 LSP 계수를 구한 후, 이를 18차 LPC로 변환한다. 그리고 AMR 디코더에서 추출된 여기신호를 스펙트럼 shifting 방식을 이용하여 광대역 여기신호로 확장한다.

스펙트럼 shifting 방식은 다음과 같이 수행된다. 먼저 협대역 여기신호를 2배로 up-sampling 해준 후, 저대역 필터를 통과시킨다. 이때 저대역 필터에 사용되는 cutoff 주파수 f_c 는 3400 Hz를 이용하였다. 그리고 식 (1)과 같이 코사인 함수로 협대역 여기신호를 shifting하여 광대역 여기신호를 구해준다.

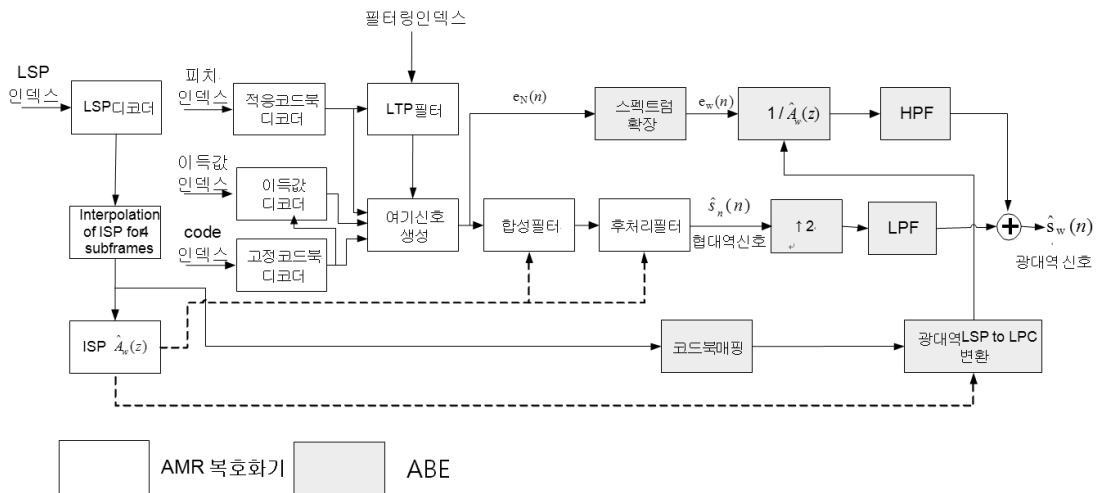


그림 1. ABE 시스템이 내장된 AMR 복호화기의 기본구조

Fig. 1. An artificial bandwidth extension (ABE) module embedded in the AMR decoder.

$$E_W(\omega) = E_N(\omega - \omega_c) + E_N(\omega + \omega_c) \quad (1)$$

$$\omega_c = \frac{2\pi f_c}{f_s} \quad (2)$$

스펙트럼 포락선의 확장을 위해 코드북 매핑 방식을 이용하는데 상세한 내용은 2.2절에서 언급한다. 코드북 매핑 과정은 LSP 계수의 스펙트럼 포락선만을 고려한 것이기 때문에 협대역 스펙트럼의 에너지와 광대역 스펙트럼의 협대역 부분의 에너지간의 차이가 발생할 수 있다. 이러한 차이를 10차 LPC계수 (A_N)와 18차 LPC계수 (A_W)의 비를 이용한 이득 값으로 보정해 줄 수 있다. 이러한 이득 값 계산은 아래 식 (3)과 같다. 본 논문에서는 계산량 감소를 위해 광대역 스펙트럼에 인접한 협대역 주파수 대역 (2400-3400 Hz)에서의 평균 에너지 비를 이득 값으로 사용하였다.

$$g_M = \frac{1}{(3400 - 2400)} \sum_{f_i=2400}^{3400} \frac{1/|A_N(e^{j2\pi f_i/8000})|}{1/|A_W(e^{j2\pi f_i/16000})|} \quad (3)$$

이와 같이 구해진 이득 값을 합성필터에 적용함으로써 저대역에서의 포락선 크기를 맞춘다. 마지막으로 AMR 디코더의 합성필터에서 출력된 신호에 2배 up-sampling 을 한 후 LPF를 통과시킨 저대역 부분과 ABE 합성필터에서 출력된 신호를 고대역 필터에 통과시킨 고대역 부분을 합해서 최종적인 광대역 음성신호를 생성한다.

2.2. 코드북 매핑

2.2.1. Dual 코드북을 이용한 18차 코드북 매핑

협대역에서는 10차 LSP를 이용하였고, 광대역에서는 18차 LSP를 사용하였다. 코드북 크기는 계산량 측면과 성능 관점을 고려하여 실험적으로 512를 사용하였다. 협대역과 광대역의 코드 벡터는 각 LSP 세트의 centroid 값에 의해서 결정되고 각 코드 벡터의 인덱스는 서로 대응된다. 이 방법은 비선형 보간 벡터 양자화 방식과 개념적으로 유사하다. 저대역 및 고대역 성분은 밀접한 상관관계에 있기 때문에 그림 2와 같이 협대역 스펙트럼 포락선 코드북은 이에 대응하는 광대역 스펙트럼 포락선 코드북을 동반한다. AMR 디코더에서는, 복호화된 10차 LSP 벡터를 이용하여, 협대역 코드북 내에서 왜곡이 최소가 되는 협대역 코드벡터 인덱스를 풀 서치 (full-search) 한 후 이 인덱스에 해당하는 광대역 코드북의 18차 LSP 벡터를 찾는다.

왜곡이 최소가 되는 코드벡터를 찾기 위해, 다음 식을

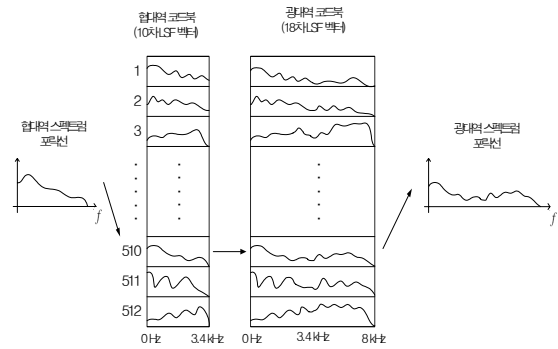


그림 2. Dual 코드북을 이용한 18차 코드북 매핑 방법
Fig. 2. 18th order codebook mapping method using dual codebook.

distortion measure로 사용하였다.

$$d_i(f, \hat{f}^i) = \sum_{j=1}^p (f_j - \hat{f}_j^i)^2 \quad (4)$$

여기서 f 는 디코더에서 복호화된 10차 LSP 벡터이며, \hat{f}^i 는 협대역 코드북 내 i 번째 인덱스에 대한 코드벡터이다. p 는 협대역 LSP 벡터의 차수로 10을 사용하였다.

2.2.2. 피치 값을 이용한 classified dual 코드북 매핑

사람의 음성은 성대의 진동 유, 무에 의해서 유성음 (voiced sound)과 무성음 (unvoiced sound)으로 나눌 수 있다. 공기가 성문을 통과할 때 성문이 거의 닫힌 상태에서 두 성대가 진동하며 만들어지는 음성을 유성음이라 하며, 반대로 성문이 열려져서 성대가 진동하지 않으며 만들어지는 음성을 무성음이라 한다. 유·무성음은 스펙트럼 포락선 특성이 서로 다르므로 단일 매핑 코드북을 이용 시 큰 매핑 에러를 초래할 수 있다. 그래서 본 논문에서는 스펙트럼 특성에 따라 매핑 코드북을 달리하는 classified 코드북을 사용하였다. 음성 분류를 위해, AMR 디코더에서 복호화 되는 피치정보를 이용하여 식 (5)의 voicing degree 계수 β 값을 구하였다.

$$\beta = \left| \frac{\sum_{n=0}^{N-1} s(n)s(n+\tau)}{\sum_{n=0}^{N-1} s^2(n+\tau)} \right| \quad (5)$$

여기서 τ 값은 피치 값이고, N 은 부 프레임 크기 값이다. 실험에 의해 β 값의 범위 (즉, $0.75 \leq \beta \leq 1.25$ 일 때 유성음, $\beta < 0.75$ 및 $\beta > 1.25$ 일 때 무성음)를 구하여 유성음과 무성음을 분류하였고, 각각에 대한 dual 코드

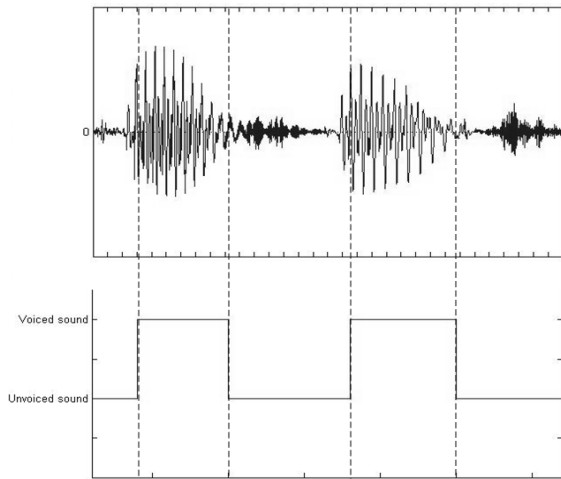


그림 3. Pitch 값을 이용한 유, 무성음 분류
 Fig. 3. Voiced, unvoiced sound classification using pitch value.

북을 별도로 만들어서 코드북 매핑을 수행하였다. 식 (5)의 β 값 범위에 따른 분류 결과는 그림 3과 같다.

2.2.3. 코드북 탐색시 가중치 적용

일반적으로, 코드북 탐색 시 입력 LSP 계수와 협대역 LSP 코드북 내 코드벡터들 간 Euclidean distortion measure를 이용하여 왜곡이 가장 적은 코드벡터를 찾는다. 하지만 이러한 distortion measure는 스펙트럼의 모든 주파수 성분에 동일한 가중치를 부여한 것이다. LSP 계수에는 포먼트 성분과 포먼트가 아닌 성분이 있는데, 음성 특성상 포먼트 성분이 포먼트가 아닌 성분보다 더 중요하기 때문에 이 점을 고려하여 가중치 적용 distortion measure인 다음 식 (6)를 본 논문에서 사용하였다.

$$d_w(f, \hat{f}) = \sum_j w_j (f_j - \hat{f}_j)^2 \tag{6}$$

여기서,

$$w_j = (f_{j-1} - f_j)^{-1} + (f_j - f_{j+1})^{-1}, \tag{7}$$

$j = 1, \dots, p$

식 (7)에서의 경계조건은 $f_0 = 1$ 및 $f_{p+1} = -1$.

2.3. Post 필터링을 통한 음질 개선

과거와 현재 프레임간의 상관도를 이용한 smoothing 필터를 설계함으로써 스펙트럼 포락선이 부드러워 지는 효과와 더불어 음질 향상의 결과를 얻었다. 사용된 smoothing 필터는 프레임마다 급변하는 LSP 계수의

변화를 자연스럽게 한다 [9]. 식 (8)에서와 같이 과거 프레임과 현재 프레임에 일정한 계수를 곱한 후 더하였다. α 값은 0부터 0.5까지 0.1 단위로 실험하여 PESQ (perceptual evaluation of speech quality) 음질 성능 관점에서 가장 최적 값인 0.3을 사용하였다.

$$\hat{f}_i(n) = \alpha \hat{f}_i(n-1) + (1-\alpha) \hat{f}_i(n) \text{ for } 1 \leq i \leq 18 \tag{8}$$

III. 고속 코드북 탐색에 의한 저전력 코드북 매핑

ABE 시스템 내 전체 계산량의 30% 이상이 코드북 매핑 모듈에서 요구되며, 코드북 매핑에서 요구되는 대부분의 계산량은 협대역 코드북 탐색에서 이용된다. 그래서 저 전력 ABE 시스템 설계를 위해서는 고속 코드북 탐색 방식의 도입이 필요하다. 본 논문에서는 single distance search (SDS)와 LSP의 순서성질을 이용한 고속 탐색 [5]을 적용한 저 전력 코드북 매핑 방식을 제안한다.

3.1. SDS 고속 탐색 방식을 이용한 고속 탐색 코드북 매핑

SDS 방식은 입력 LSP 벡터의 기본 요소값과 코드북 내 코드워드의 기본 요소값 간의 최대 왜곡 값을 제한함으로써, 제한된 범위 내에 있는 코드워드들만 탐색하여 코드북 탐색관련 계산량을 줄이는 방식이다. 제한 범위는 식 (9)과 같다.

$$p^j - T/\sqrt{w_j} < \hat{p}_i^j < p^j + T/\sqrt{w_j} \tag{9}$$

$$T = g\sqrt{d_{\max}}$$

여기서, j 는 LSP 벡터의 기본 요소로서, 본 논문에서는 실험적으로 결정된 5를 사용하였다. $\sqrt{d_{\max}}$ 값은 입력 LSP 벡터와 코드벡터의 최대 왜곡 값으로서, 10개의 음성 샘플들을 사용하여 실험한 결과 0.1268을 얻었다. 또한 폴 서지 시의 음질 성능에 비해 성능 열화가 없는 가중치 g 값들 중 최소 g 값을 실험에 사용하였다. w_j 는 가중치 계수로서 식 (7)을 사용하였다.

3.2. LSP 순서 성질을 이용한 고속 탐색 방식 코드북 매핑

본 절에서는 코드북 탐색 시 LSP계수의 순서 성질을

이용하여 탐색되는 코드벡터를 제한하는 고속 탐색 알고리즘을 제안한다. LPC계수의 전송계수로 널리 사용되고 있는 LSP 계수는 합성필터의 안정성과 관련된 순서 성질을 가지고 있다. 이러한 순서 성질로부터 주어진 대상 벡터 p 와 코드북 내의 코드벡터 \hat{p} 가 식 (10)와 식 (11)을 만족 한다는 것을 확인할 수 있다.

$$\hat{p}_l^n > p^{n+1}, \quad 1 \leq l \leq L \quad (10)$$

$$\hat{p}_l^n < p^{n-1}, \quad 1 \leq l \leq L \quad (11)$$

여기서, 식 (10)와 같이 순서성질의 만족여부를 결정하기 위해 어떤 코드벡터의 n 번째 열의 요소 값과 대상 벡터의 $(n+1)$ 번째 열의 요소 값을 비교하는 것을 순방향 비교라 하며, 식 (11)와 같이 코드벡터의 n 번째 열의 요소 값과 대상 벡터의 $(n-1)$ 번째 열의 요소 값을 비교하는 것을 역방향 비교라 한다 [9]. 그리고 이를 합쳐서 FBC (forward backward comparison) 방식이라 한다. 이러한 순방향 비교식과 역방향 비교식을 이용하여 코드북의 특정 열의 요소 값과 대상 벡터 내 해당 열 전후의 요소 값들을 비교하여 순서성질을 위배하는 코드벡터들을 탐색 대상에서 제외하는 개념으로 코드북의 탐색 범위를 결정한다. 이렇게 결정된 코드워드들만을 탐색함으로써, 코드북 내 전체 코드워드들을 모두 탐색하는 풀서치 (full search) 방식에 비해 계산량을 줄일 수 있다.

IV. 실험 및 결과

제안된 AMR 기반 ABE 시스템의 성능 평가를 위해 사용된 음성 데이터는, NTT에서 제작된 8분 분량의 광대역 영어 샘플들로서 남성 및 여성 화자 샘플들이 반반으로 구성되어 있다. 이러한 NTT 광대역 음성 샘플들을 3.4 kHz 컷오프 주파수의 16차 Chebyshev 저대역 필터에 통과 시킨 후 8 kHz로 다운 샘플링 함으로서 협대역 음성 샘플들을 구하였다. 이러한 협대역 및 광대역 음성 샘플들을 이용해서 각각 10차와 18차인 협대역 및 광대역 LSP값을 계산하였고, 이러한 LSP 벡터 값들을 이용하여 협대역 및 광대역 LSP 코드북을 설계하였다. 이렇게 설계된 코드북들을 사용하여, 저 전력을 위한 고속 코드북 탐색관련 실험과 음질 성능 향상을 위한 가중치 적용 classified 코드북 관련 실험을 수행 하였다.

위의 표 1에서 6명 화자 (남자 3명, 여자 3명)의 test 음성 샘플들에 대한 단일 코드북과 가중치 적용 classified

표 1. 단일 코드북과 가중치 적용 classified 코드북의 PESQ 음질 성능 비교

Table 1. PESQ performance of single codebook and weighted classified codebook.

	단일 코드북	가중치, 후처리 필터 적용 classified 코드북
PESQ	3.844	3.869

표 2. DTE 시스템과 저 전력 Embedded 시스템 간의 중요 함수 별 WMOPS값 비교

Table 2. WMOPS complexity comparison between the DTE system and the low power embedded system.

ABE내 모듈	DTE	저전력 Embedded
코드북 매핑	3.352	1.838
LPC 분석	0.280	0
LPC to LSP	1.505	0
파치값 계산	0.229	0
여기신호 계산	0.280	0
기타 모듈들	6.839	6.839
합계	12.485	8.957

표 3. 고속 탐색에 의한 계산량 감소율 비교

Table 3. Complexity reduction with fast search.

	코드벡터의 탐색 범위	탐색 범위 감소율 (%)
전체 탐색	512	0
SDS 고속 탐색	279	45
FBC 고속 탐색	430	16
SDS+FBC 고속탐색	271	47

코드북의 PESQ 음질 성능을 비교 하였다. 가중치, 후처리 필터 적용 classified 코드북을 도입함으로써 PESQ MOS 측면에서 0.025 향상 되었다. 이러한 PESQ 성능 차이를 주관적인 음질 비교분석에서 명확히 느낄 수 있었다.

표 2는 AMR 디코더 출력 단에 ABE를 위치시키는 “DTE 시스템”과 AMR 디코더 내부에 ABE를 위치시키고 코드북 매핑 모듈에 고속 탐색 알고리즘을 적용한 “저 전력 Embedded 시스템” 간의 중요 함수 별 WMOPS (weighted million operations per second) 값을 비교한 것이다. 저 전력 embedded 시스템은 DTE 시스템에 비해 LSP 계수 분석, LPC-LSP 변환 및 LP 해석 필터링을 AMR 디코더에서 활용할 수 있으므로 전체 계산량 기준으로 약 16 % 감소된다.

그리고 표 3에서 보듯이 ABE 시스템의 코드북 매핑 모듈에 본 논문에서 제안한 고속 탐색 알고리즘을 적용함으로써, 코드북 매핑 모듈 기준으로는 47 %, 전체 계산량 기준으로는 18 % 감소된다. 따라서 저 전력 Embedded

표 4. DTE 시스템과 저전력 Embedded 시스템의 알고리즘 지연 비교 (프레임 크기: 20 ms 가정)

Table 4. Algorithmic delay comparison of the DTE system and the low power embedded system.

	알고리즘 지연 (msec)
DTE	20+20
저전력 Embedded	20

표 5. AMR 코덱의 협대역 음성 출력과 제안된 저전력 Embedded ABE 시스템의 광대역 음성 출력 간의 A/B 선호도 비교

Table 5. A/B preference between the narrowband AMR synthesized speech and the wideband ABE band-extended speech.

	AMR 코덱 협대역 음성 출력	저전력 Embedded 시스템의 광대역 음성 출력	구분 인됨
A/B 선호도	0 %	90 %	10 %

시스템은 DTE 시스템에 비해 28 % 정도 계산량을 줄일 수 있다. 그런데 고속 코드북 탐색 도입으로 인한 최적 코드워드의 인덱스 변화는 없으므로 음질 성능 저하는 없다.

표 4는 DTE 시스템과 저 전력 Embedded 시스템의 알고리즘 지연을 비교하였다. 저 전력 Embedded 시스템의 경우 AMR 내부에서 복호화 되는 LPC계수를 ABE 시스템이 활용함으로써, LPC 계수를 ABE 내부에서 별도로 분석해야 하는 DTE 시스템에 비해 알고리즘 지연을 20 ms 줄일 수 있다.

표 5는 AMR 코덱 협대역 음성 출력과 저전력 Embedded 시스템의 광대역 음성 출력 간의 A/B 선호도 비교 결과를 나타내었다. 실험 방법은 10명의 청취자들에게 남녀 각각 10개의 샘플들에 대한 음성 출력을 들려주고 어떤 음성이 더 좋은지를 선택하도록 했다. 조사 결과 9명이 저 전력 Embedded 시스템의 광대역 음성 출력을 선택하였고, 1명이 구별 안됨을 선택하였다.

V. 결론

본 논문에서는, 협대역 표준 음성코덱인 AMR 시스템의 디코더 내부에 인공 대역 확장 시스템을 포함하고, 인공 대역 확장 시스템의 코드북 매칭 모듈에 고속 코드북

탐색 방식을 도입하여, 대역폭 확장에 요구되는 계산량과 알고리즘 지연을 줄인 “AMR 기반 저 전력 인공 대역 확장 시스템”을 제안 하였다. 제안된 시스템은 AMR 디코더 출력 단계에 ABE를 위치시키고 폴 서치 코드북 탐색을 사용하는 기존 시스템에 비해 28 % 정도 계산량을 줄일 수 있었고, 알고리즘 지연도 20 ms 줄일 수 있었다. 또한 가중치, 후처리 필터 적용 classified 코드북을 도입하여 확장되는 광대역 신호의 음질을 PESQ MOS 측면에서 0.025 향상시켰다.

감사의 글

이 논문은 2010년도 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (과제번호 : 2010-0021860).

참고 문헌

1. ITU-T Rec. G.722, *Performance characteristics of PCM channels between 4-wire interfaces at voice frequencies*, 1988.
2. Peter Jax and Peter Vary, "On Artificial Bandwidth Extension of Telephone Speech," *SIGNAL PROCESSING*, vol. 83, no. 8, pp. 1707-1719, August, 2003.
3. P. Noll, "Wideband speech and audio coding," *IEEE Communications Magazine*, pp. 34-44, Nov.,1993.
4. G. Fant, "Acoustics theory of speech production," *The Hague, The Netherlands: Mouton*, 1960.
5. U. Kornagel, "Techniques for artificial bandwidth extension of telephone speech," *Signal Processing*, vol. 86, no. 6, pp 1276-1306, 2006.
6. K. Park and H. Kim, "Narrowband to wideband conversion of speech using GMM based transformation," *proc. ICASSP*, pp. 1843-1846, 2000.
7. J. Marhouf and M. Berouti, "High-frequency regeneration in speech coding systems," *proc. ICASSP*, vol. 4, pp. 428-431, 1979.
8. R. Laroia, N. Phamdo, N. Farvardin, "Robust and efficient quantization of speech LSP parameters using structured vector quantizers," *ICASSP-91*, pp. 641-644, vol. 1, 1991.
9. S. Kang, Y. Shin, C. Son, and T. R. Fischer, "A Fast Encoding Technique for Vector Quantization of LSP Parameters," *IEICE Trans. on commu.*, vol. E88-B, no. 9, pp. 3750-3755, Sept. 2005.

저자 약력

●구 본 강 (Bonkang Koo)

2009년: 한양대학교 전자및통신공학전공 (학사)
2011년: 한양대학교 전자전기제어계측공학과 (석사)
2011년 ~ 현재: LG전자 MC사업부 연구원
※ 주관심 분야: 음성/오디오 부호화, 신호처리, 이동통신

●박 희 완 (Heewan Park)

2009년: 한양대학교 전자및통신공학전공 (학사)
2009년 ~ 현재: 한양대학교 전자전기제어계측공학과 (석사)
※ 주관심 분야: 음성/오디오 부호화, 신호처리, 이동통신

●주 연 재 (Yeon-Jae Ju)

2008년: 한양대학교 전자전기공학과 (학사)
2010년: 한양대학교 전자전기제어계측공학과 (석사)
2010년 ~ 현재: 트란소노 오디오기술연구소 주임연구원
※ 주관심 분야: 음성/오디오 부호화, 신호처리, 이동통신

●강 상 원 (Sangwon Kang)

1980년: 한양대학교 전자공학과 (학사)
1982년: 서울대학교 전자공학과 (석사)
1990년: Texas A&M대학교 전기공학과 (박사)
1982년 ~ 1994년: 한국전자통신연구원 신호처리 연구실
1994년 ~ 현재: 한양대학교 전자컴퓨터공학부 정교수
※ 주관심 분야: 음성/오디오 부호화, 신호처리, 이동통신