

# 16 비트 고정 소수점 DSP를 이용한 GSM-EFR 음성 부호화기의 실시간 구현

## Real-time Implementation of a GSM-EFR Speech Coder on a 16 Bit Fixed-point DSP

최민석\*, 변경진\*, 김경수\*

(Min Seok Choi\*, Kyung Jin Byun\*, Kyung Soo Kim\*)

\*한국전자통신연구원

(접수일자: 2000년 8월 8일; 채택일자: 2000년 10월 10일)

본 논문에서는 DSP Group사의 16비트 고정 소수점 DSP(Digital Signal Processor)인 OakDSP Core를 사용하여 유럽의 이동통신에서 표준으로 사용되고 있는 음성 부호화기 알고리즘인 GSM-EFR (Global System for Mobile communications-Enhanced Full Rate) 을 실시간으로 구현하였다. 실시간 구현된 GSM-EFR 음성 부호화기의 계산량은 약 24MIPS가 소요되며, 7.06K 워드의 코드 메모리와 12.19K 워드의 데이터 메모리를 사용하였다. 구현된 음성 부호화기는 ETSI에서 제공하는 시험 벡터 샘플을 모두 통과하였으며, 객관적 평가 틀을 이용하여 지각 평가를 수행한 결과, 32kbps ADPCM과 비슷한 음질을 보였다. 본 논문에서 실시간으로 구현된 GSM-EFR 음성 부호화기는 IMT2000 비동기 방식의 음성 부호화기 표준인 GSM-AMR의 최상위 전송률 모드로서, 앞으로 IMT-2000 비동기식 단말기용 모뎀 ASIC에 탑재할 GSM-AMR 음성 부호화기의 구현을 위한 기본 구조로 이용될 예정이다.

**핵심용어:** 음성부호화기, GSM-EFR, 실시간 구현, OakDSP

**투고분야:** 음성처리 분야 (2.2)

This paper describes a real-time implementation of a GSM-EFR (Global System for Mobile communications Enhanced Full Rate) speech coder using OakDSP core; a 16bit fixed-point Digital Signal Processor (DSP) by DSP Group, Inc. The real-time implemented speech coder required about 24MIPS for computation and 7.06K words and 12.19K words for code and data memory, respectively. The implemented GSM-EFR speech coder passes all of test vectors provided by ETSI (European Telecommunication Standard Institute), and perceptual speech quality measurement using MNB algorithm shows that the quality of the GSM-EFR speech coder is similar to the one of 32kbps ADPCM. The real-time implemented GSM-EFR speech coder which is the highest bit-rate mode of the GSM-AMR speech coder will be used as the basic structure of the GSM-AMR speech coder which is embedded in MODEM ASIC of IMT2000 asynchronous mode mobile station.

**Key words:** Speech coder, GSM-EFR, Real-time implementation, OakDSP

**Subject classification:** Speech signal processing (2.2)

### I. 서 론

지난 십 여년간 낮은 복잡도와 전송률 하에서 우수한 음질을 얻으려는 연구의 결과로 현재 여러 가지 음성 부호화기 표준들이 국제 기구에 의해 채택되어 사용되고 있다. 지금까지 제안된 음성 부호화 방식들 중에서 우수한 음질을 갖는다고 알려져 있는 CELP (Code-Excited Linear

Prediction)[1]는 표준 음성 부호화 알고리즘의 주류를 이루고 있으며 현재에도 성능 개선 연구가 계속되고 있다. CELP의 성능 개선 연구는 크게 계산량 감소와 음질 개선의 두 가지 분야로 수행되고 있다. 제안될 당시의 초기 CELP는 많은 계산량으로 인해 실시간 구현이 불가능하였지만, 이후 많은 고속 알고리즘들이 제안되어 현재는 단일 범용 디지털 프로세서로도 실시간 구현이 가능하게 되었다. 또한 최근에 개발된 DSP (Digital Signal Processor) 들은 과거에 비해 향상된 성능과 많은 내부 메모리를 가지고 있어서 기존의 많은 계산량을 필요로 하던 음성 부

책임저자: 최민석 (mschoi@etri.re.kr)  
305-350 대전광역시 유성구 가정동 161 한국전자통신연구원  
(전화: 042-860-1317; 팩스: 042-860-6108)

호화 방식의 실시간 구현이 가능해졌을 뿐만 아니라, 다채널 서비스를 할 수 있는 음성 부호화기를 하나의 DSP에 구현하는 것이 가능하여 디지털 이동통신 시스템의 소형화를 통한 비용 절감 효과를 기대할 수 있게 되었다. G.723.1[2], G.729[3][4], GSM-EFR[5] 등과 같은 최근의 음성 부호화기 국제 표준에서도 CELP기반의 음성 부호화 알고리즘인 ACELP (Algebraic CELP)[6]를 사용하고 있다. 디지털 이동통신 시스템에서 음성 부호화기는 통신 품질에 직접적으로 영향을 주는 부분 중의 하나로서 낮은 전송률과 무선 환경에 대한 강인성이 요구되며, 유선 전화의 음질(Toll Quality)을 보장해야 한다. 초기의 유럽의 이동통신 표준인인 GSM (Global System for Mobile communications) 시스템에서는 음성 부호화 방식으로 13kbps RPE-LTP (Regular Pulse Excitation-Long Term Prediction)을 사용하였으나, 최근에는 열악한 채널 환경이나 잡음 환경에서 보다 나은 음질을 제공하기 위해 1996년 GSM-EFR (Enhanced Full Rate speech codec) 음성 부호화 표준안 (GSM 06.60)[7]을 채택하였다. 이 방식은 ITU-T G.729 표준안과 유사하며, ACELP 음성 부호화 방식에 기반을 두고 있는데, 음성 부호화를 위해 12.2 kbps, 채널 부호화를 위해 10.6kbps의 비트율을 할당하여 전체 채널 전송률은 22.8kbps가 된다. 또한, GSM-EFR 음성 부호화 방식은 차세대 이동통신 표준안으로서, IMT-2000에서 채택된 AMR (Adaptive Multi-Rate) 음성 부호화 방식의 기초가 되고 있다. 이 방식은 기존의 FR (Full-Rate) 음성 부호화 표준안 (GSM 06.10)[8]을 쓰고 있는 GSM 네트워크에서 전송속도를 증가 시킬 필요가 없으며, 현재 사용되고 있는 음성 부호화 방식의 표준안을 변형시킬 필요도 없다는 장점을 가지고 있다.

본 연구에서는 IMT2000 비동기 방식 단말기에서 사용되는 GSM-AMR 음성부호화기의 개발의 일환으로서 최상위 전송률을 가지는 GSM-EFR 음성 부호화 방식을 DSP Group 사의 OakDSP EVM (Evaluation Module) 보드 상에서 실시간으로 구현하였다.

본 논문의 구성은 다음과 같다. II장에서는 GSM-EFR 음성 부호화 알고리즘의 구조와 개요를 개괄적으로 분석하고, III장에서는 GSM-EFR 음성 부호화 알고리즘의 실시간 구현 및 성능평가에 대해 상세히 설명하고, 마지막으로 IV장에서는 결론을 맺는다.

## II. GSM-EFR 음성 부호화 알고리즘

GSM-EFR 음성 부호화기는 20ms의 프레임을 사용하는데, 한 프레임은 5ms의 서브프레임으로 나뉘어 분석된다. 음성 부호화 과정에서 분석 프레임별로 할당되는 비트수는 표 1과 같다. 부호화 과정에서 한 프레임 당 2세트의 선형 예측 계수를 구하여 LSP (Line Spectral Pair) 계수로 변환한다. 또한 각 서브프레임 당 적응 코드북과 고정 코드북의 인덱스와 이득을 구하고, 복호화 과정에서는 위의 파라미터들을 이용하여 서브프레임 단위로 음성 신호를

합성하고 마지막 단에서 후처리 필터링 과정을 통해 음질을 개선하게 된다.

표 1. GSM-EFR 음성 부호화기의 비트 할당

Table 1. Bit allocation of the GSM-EFR speech codec for 20ms frame.

Parameter	1 <sup>st</sup> & 3 <sup>rd</sup> subframe	2 <sup>nd</sup> & 4 <sup>th</sup> subframe	Total per frame
2 LSP sets			38
Pitch delay	9	6	30
Pitch gain	4	4	16
Algebraic code	35	35	140
Codebook gain	5	5	20
Total(bit)			244

GSM-EFR 음성 부호화 알고리즘의 부호화 과정은 그림 1과 같다. 입력 음성은 직류 성분 제거를 위한 고역통과 필터 등의 전처리 과정을 거친 후, 30ms의 길이를 가지는 서로 다른 두개의 비대칭 창함수를 적용하여 한 프레임 (20ms) 당 두 번의 10차 선형 예측 분석을 수행하게 된다. 이렇게 구해진 두 세트의 선형 예측 계수 들은 각각 LSP 파라미터로 변환되고, 두 세트의 LSP 계수들은 양자화를 위해 1차의 MA (moving average) 예측 과정을 거치게 된다. 여기서 LSP 잔여 벡터들은 2x2 영역으로 구성된 5개의 부행렬을 가지는 분할 행렬 양자화 (Split Matrix Quantization (SMQ))를 사용하여 양자화 된다. 이 부행렬들은 순차적으로 7, 8, 9, 8, 6 비트를 사용하여 각각 양자화되는데, LSP의 양자화에는 총 38비트가 사용된다. 1번째와 3번째 서브프레임에서의 LP 파라미터는 이웃하는 두개의 서브프레임의 값을 선형보간하여 구해진다.

적용 코드북 탐색을 위한 피치 예측 과정은 고정된 지연 영역인 [17 3/6, 143] 사이에서 개루프 탐색과 페루프 탐색을 수행함으로써 이루어진다. 먼저 매 10ms 마다 가중 필터를 거친 원 음성 신호로부터 최적의 정수 지연,  $T_0$  값을 구하기 위해 개루프 탐색을 하게 된다. 이렇게 구해진  $T_0$  값을 이용하여 페루프 탐색 과정을 수행하는데, 1번째와 3번째 서브프레임에서는  $T_0 \pm 3$  범위 내에서 최적의 페루프 지연  $T_1$ 을 찾게 된다. 이때, 구해진 지연 값이 [17 3/6, 94 3/6]의 범위에 있을 경우에는 소수 지연이 1/6분해능을 가지도록 탐색하며, 반면에 [95, 143]의 범위에서는 단지 정수 지연 만을 찾게 된다. 또한, 2번째와 4번째 서브프레임에서는 위에서 구한  $T_1$  값을 이용하여 계산되는데 [ $T_1 - 5$  3/6,  $T_1 + 4$  3/6]의 범위 내에서 1/6 분해능으로 지연 값이 구해진다. 이렇게 구해진 지연 은 1번째와 3번째 서브프레임에서는 9비트로, 2번째와 4번째 서브프레임에서는 6비트로 양자화되고, 코드북 이득 값은 4비트로 양자화된다.

위의 두 과정을 거친 잔차신호를 모델링하기 위해 매 서브프레임마다 대수 코드북 (algebraic codebook)을 사용하게 된다. 이 코드북에서는 각 여기벡터들은  $\pm 1$ 의 크기를 가지는 10개의 non-zero 펄스열로 구성된다. 따라서, 각 서브프레임에서 40개의 펄스 위치는 두 개의 펄스열을 가지는 5개의 트랙으로 나뉘어져 모델링된다. 각 트랙의 펄스 위치는 표 2와 같다. 각 펄스의 위치는 3비트로 양자화되고, 각 트랙의 첫번째 펄스의 부호를 나타내기 위해 1비트가 사용되므로, 한 서브프레임의 고정 코드북을 전송하기 위해서는 총 35비트가 사용된다.

최적의 펄스 위치는 적응 코드북 탐색 과정을 거쳐 나온 잔차 신호와 대수 코드벡터 사이의 상관이 최대가 되는 값을 찾음으로써 구해진다. 이때, 대수 코드북의 구조에서는 5개의 검색 루프를 통해 연속적으로 최적의 펄스 위치를 찾게 되므로 계산량의 감소와 우수한 음질을 얻을 수 있게 된다. 고정 코드북의 이득 값은 MA 예측을 이용하여 5비트로 양자화된다.

GSM-EFR 음성 부호화기의 복호화 과정에서는 전송된 파라미터들로부터 LP 파라미터, 피치, 코드북 인덱스, 그리고 이득 값을 추출하여, 이 파라미터들을 이용하여 합성음을 복원하게 된다. 또한, 복원된 합성음의 음질을 향상시키기 위해 후처리 필터링을 거치게 된다.

표 2. 고정(대수) 코드북의 각 트랙에 대한 펄스 위치  
Table 2. Potential positions of individual pulses in the algebraic codebook.

track	pulse	Positions
1	i0, i5	0, 5, 10, 15, 20, 25, 30, 35
2	i1, i6	1, 6, 11, 16, 21, 26, 31, 36
3	i2, i7	2, 7, 12, 17, 22, 27, 32, 37
4	i3, i8	3, 8, 13, 18, 23, 28, 33, 38
5	i4, i9	4, 9, 14, 19, 24, 29, 34, 39

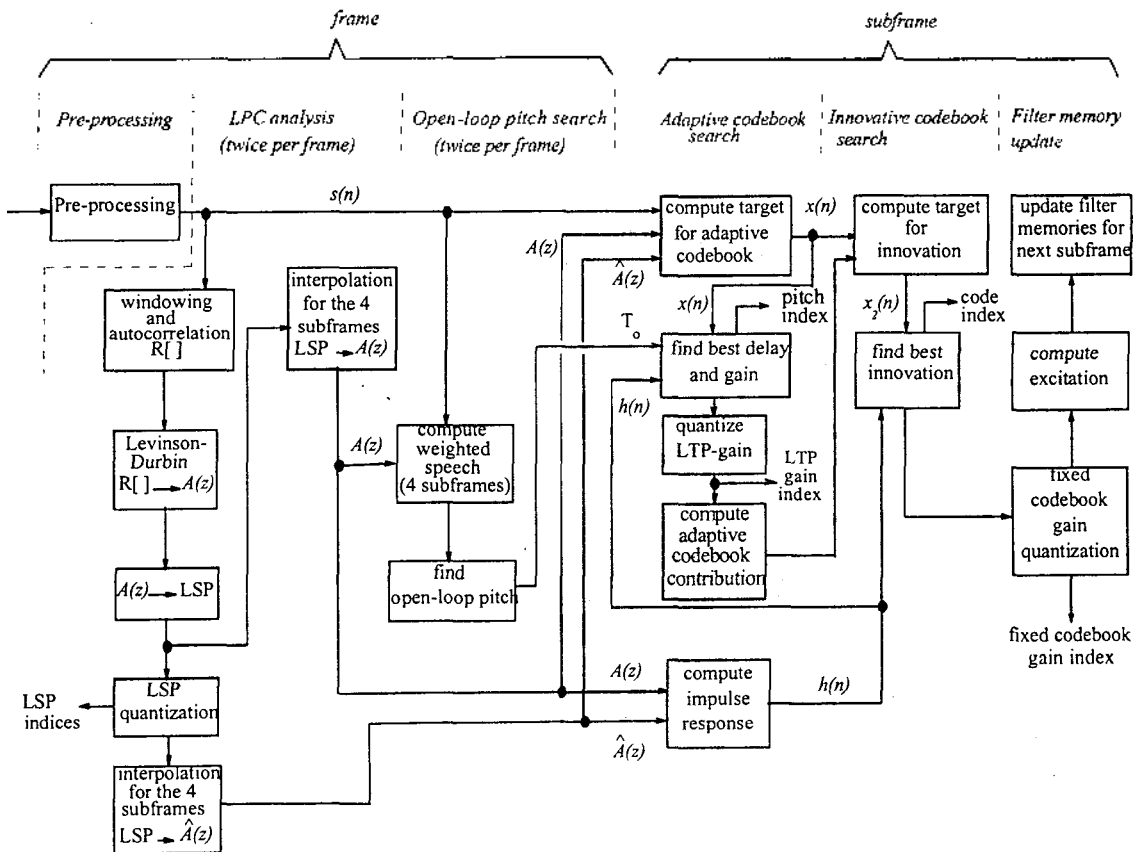


그림 1. GSM-EFR 부호화기의 기본 블록도  
Fig. 1. Simplified Block diagram of the GSM-EFR encoder.

### III. GSM-EFR 음성 부호화기의 실시간 구현

본 연구에서 사용된 OakDSP Core는 DSP Group, Inc. 의 16비트 고정 소수점 신호처리 프로세서로서 40MHz의 동작 주파수와 40MIPS의 계산 능력을 가지며, 16비트 데이터와 프로그램 버스를 지닌다. 또한, 저전력으로 설계되어 있어 통신과 전자 응용분야에서 다양하게 사용되고 있다. OakDSP Core는 일반 범용 DSP와는 달리 대부분의 ASIC 공정에서 구현 가능한 IP 형태로 제공되기 때문에 향후 IMT2000 이동기 방식 단말기의 ASIC 구현에 용이하다. OakDSP Core의 주요 특징은 다음과 같다.

- 40MHz 동작 주파수 (40 MIPS)
- Single cycle multiply-accumulate instruction
- 36bit ALU
- 4개의 36비트 누산기
- Single cycle exponent evaluation
- Max/Min single cycle instruction with pointer latching and modification
- 4K word on-core memory, 6K word on-chip memory
- 64k word maximum program memory space

GSM-EFR의 실시간 구현은 표준안(GSM 06.60)과 함께 제공되는 고정 소수점 C 프로그램(GSM 06.53)을 바탕으로 OakDSP 어셈블리 언어로 구현하였다.

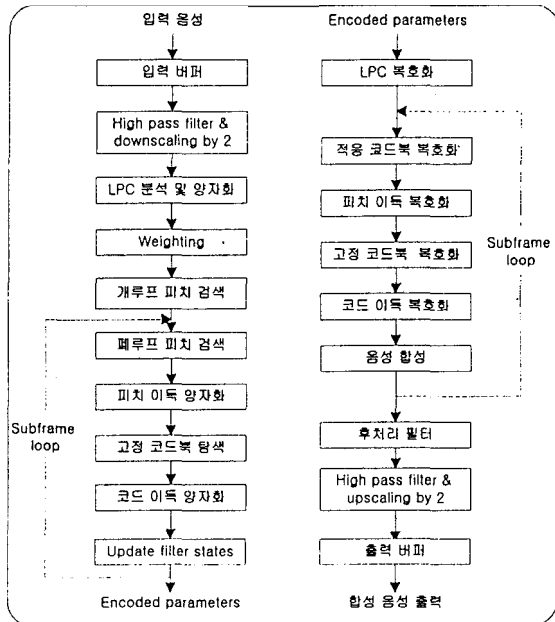


그림 2. GSM-EFR 음성 부호화기의 흐름도  
Fig. 2. Flow chart of the GSM-EFR speech coder.

C의 각 모듈을 OakDSP 어셈블리 코드로 작성하고, C의 결과와 구현된 결과가 일치하는지를 검증하기 위해 시험 벡터를 C에 인가했을 때 얻어진 입/출력 데이터를 구현

된 어셈블리 코드에 인가하여 bit-exact 결과가 같음을 확인하였다. 위와 같은 방법으로 각 모듈 들을 구현한 후, 메인 모듈과 I/O 모듈을 완성하여 그림 2와 같은 소프트웨어 구조를 갖는 프로그램을 완성하였다.

GSM-EFR 음성부호화기를 OakDSP EVM 보드 상에서 실시간 구현함에 있어서 표준 알고리즘을 그대로 구현하기 위해서 알고리즘 상의 최적화는 고려되지 않았으며, OakDSP Core의 구조적 특성을 이용하여 계산량을 감소 시키는데 집중하였다. 표 3은 구현된 GSM-EFR 음성 부호화기의 주요 모듈의 계산량을 나타낸 것으로 ACELP (고정 코드북) 모듈이 상대적으로 많은 계산량을 요구함을 알 수 있다. 따라서, 본 연구에서도 ACELP의 최적화에 많은 노력을 기울였다. 구현된 음성 부호화기는 ETSI에서 제공하는 모든 테스트 벡터[9]에 대해서 C 프로그램의 수행 결과와 완전히 일치된 결과를 보였다. 표 4에서는 GSM-EFR의 실시간 구현에 사용된 메모리 양을 나타내었다.

표 3. 구현된 주요 모듈의 계산량

Table 3. Computational load of the major module of the implemented GSM-EFR speech coder.

부호화기 전체(22.187 MIPS)		
전처리부	0.313 MIPS	1.3%
LPC 분석 및 양자화	5.396 MIPS	22.4%
피치 검색(개루프 검색)	1.714 MIPS	7.1%
적용코드북 및 이득양자화	3.308 MIPS	13.8%
고정코드북 및 이득양자화	11.068 MIPS	46.1%
기타 모듈	0.388 MIPS	1.6%
복호화기 전체(1.844 MIPS)		
파라미터 추출부	1.049 MIPS	4.4%
후처리 필터	0.760 MIPS	3.2%
기타 모듈	0.035 MIPS	0.1%
전체	24.031 MIPS	100%

표 4. GSM-EFR 구현 결과

Table 4. Result of implementation of GSM-EFR.

	부호화기	복호화기	전체
Program ROM	5204 words	1856 words	7060 words
Data RAM	6394 words	823 words	7217 words
Data ROM	4842 words	127 words	4969 words

실시간으로 구현한 GSM-EFR음성 부호화기의 객관적 성능 평가를 위해 G.721(32kbps ADPCM)과 EVRC (Enhanced Variable Rate Codec) 음성 부호화기의 시뮬레이션 결과와 비교한 결과를 표 5에 나타내었다. 객관적 평가 톨로 널리 사용되는 것으로는 SNR과 segSNR, CD (Cepstral Distance), BSD(Bark Spectral Distance) 등이 있으나, 최근의 음성 부호화기 들은 파형 부호화기(waveform coder)와는 달리 파형 매칭 보다는 지각적 성능에 중점을 두고 개발되었기 때문에 SNR과 같은 객관적 톨에 의한

평가는 표 5에서 볼 수 있듯이 적당하지 않다. 그러므로 본 논문에서는 최근에 제안된 지각 평가(perceptual measure) 방법인 MNB (Measuring Normalizing Block) 알고리즘[10]을 사용하여 그림 3과 같이 객관적 평가를 수행하였다. 코딩 SNR과 지각 평가(Perceptual Measure)의 측정에 사용한 입력은 부호화기 블록의 전처리를 거친 신호이고, 출력 신호는 복호화기 블록에서 후처리 필터링을 거치기 전의 신호이다. 표 5에서 보면, 실제 주관적인 평가의 결과와는 상이하게 코딩 SNR이 ADPCM에 비해 상당히 저하됨을 알 수 있다. 그러므로 이러한 SNR 측정은 최근의 음성 부호화기 들의 객관적 평가 틀로 사용하기에 적합하지 않다. MNB 알고리즘을 사용한 지각 평가에서는 GSM-EFR가 EVRC에 비해 좋은 성능을 보이며, ADPCM과도 비슷한 성능을 보이는 것을 알 수 있다. 이를 통해 EVRC음성 부호화기에 비해 GSM-EFR이 좀 더 지각적 특성을 잘 반영하는 것을 알 수 있다.

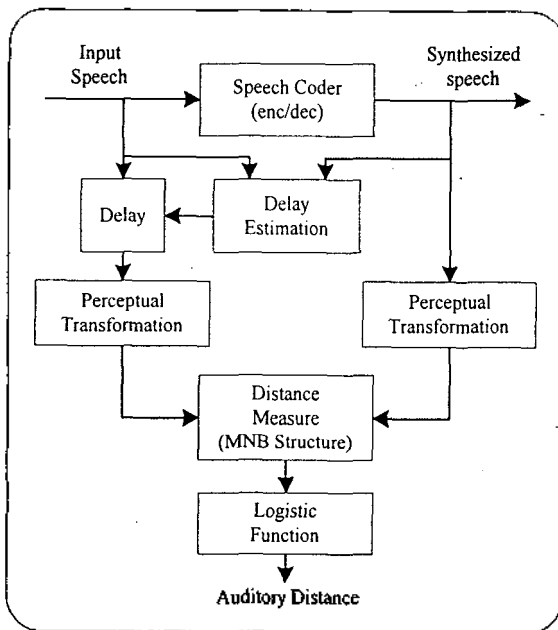


그림 3. MNB 알고리즘의 기본 블록도  
Fig. 3. Block diagram of the MNB algorithm.

표 5. Perceptual Measure(PM) 및 SNR 측정 결과  
Table 5. Measurement of Perceptual Measure(PM) and SNR(dB).

	PM(0~1.0)			SNR(dB)		
	G.721	EFR	EVRC	G.721	EFR	EVRC
Male1	0.814	0.832	0.701	16.115	4.142	-1.121
Male2	0.824	0.828	0.689	16.575	4.125	-1.960
Male3	0.829	0.831	0.676	18.397	3.359	-1.568
Female1	0.838	0.837	0.738	21.815	10.487	-3.450
Female2	0.794	0.815	0.630	23.213	7.233	-1.983
Female3	0.819	0.816	0.692	21.384	7.907	-3.046
Average	0.820	0.827	0.690	19.583	6.209	-2.188

### V. 결론

본 논문에서는 비동기 방식의 IMT2000 이동통신 단말기에서 사용되는 GSM-AMR 음성 부호화기의 최상위 전송률 모드인 GSM-EFR 음성 부호화기를 16비트 고정 소수점 DSP인 OakDSP Core에서 실시간으로 구현하였다. 실시간 구현된 GSM-EFR 음성 부호화기에 소요된 계산량은 부호화기와 복호화기에 각각 22.187MIPS와 1.844 MIPS이었으며, 필요한 메모리 크기는 코드 메모리가 7.06K words, 데이터 메모리가 12.19K words였다. 실험 결과, 구현된 GSM-EFR 음성 부호화기는 ETSI에서 제공하는 시험 벡터 샘플을 모두 통과하는 일관된 성능을 보였으며, MNB 알고리즘을 이용한 32kbps ADPCM, EVRC 음성 부호화기와의 지각 평가 비교를 통해 GSM-EFR 음성 부호화기의 우수한 성능을 확인하였다. 개발된 GSM-EFR 음성 부호화기는 IMT-2000 비동기 방식 음성 부호화기의 표준인 GSM-AMR 음성 부호화기의 개발의 기본적인 구조로 사용될 예정이다.

### 참고 문헌

1. M. R. Schroeder, and B. S. Atal, "Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates," *IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc.*, pp. 25.1.1-25.1.4, 1985.
2. ITU-T Recommendation G.723.1, *General Aspects of Digital Transmission Systems: Dual Rate Speech Coder For Multimedia Communications Transmission at 5.3 and 6.3 kbit/s*, March 1996.
3. G. Schroeder, and M. H. Sherif, "The Road to G.729: ITU-T 8-kb/s Speech Coding Algorithm with Wireline Quality," *IEEE Communications Magazine*, pp. 48-55, September 1997.
4. ITU-T Recommendation G.729, *Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear Prediction (CS-ACELP)*, March, 1996.
5. ETS Draft prETS 200 726, *Digital Cellular Telecommunications System: Enhanced Full Rate (EFR) Speech Transcoding (GSM 06.60 Version 5.0.0)*, March, 1996.
6. J. P. Adoul, P. Mabilieu, M. Delprat, and S. Morisette, "Fast CELP Coding Based on Algebraic Codes," *IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc.*, pp. 1957-1960, 1987.
7. GSM 06.60 (ETSI EN 300 726), *EFR Speech Transcoding*, ETSI, 1998.
8. GSM 06.10 (ETSI 300 961), *FR Speech Transcoding*, ETSI, 1997.
9. GSM 06.54 (ETSI EN 300 725), *Test sequences for the GSM EFR speech codec*
10. Stephen Voran, "Objective Estimation of Perceived Speech Quality Part I: Development of the Measuring Normalizing Block Technique," *IEEE Transactions on Speech and Audio Processing*, Vol.7, No.4, July 1999.

▲ 최 민 석(Min Seok Choi)



1997년 8월 : 한국과학기술원 전기 및 전자공학과(공학사)  
1999년 8월 : 한국정보통신대학원 공학부(공학석사)  
1999년 8월 ~ 현재 : 한국전자통신연구원 통신디지털회로팀 연구원  
※ 주관심분야 : 음성 신호처리, 음성 인식 및 합성, DSP 설계

▲ 변 경 진(Kyung Jin Byun)



1987년 2월 : 국민 대학교 전자공학과 (공학사)  
2000년 2월 : 한국정보통신대학원대학교 공학부(공학석사)  
2000년 3월 ~ 현재 : 한국정보통신대학원대학교 공학부 박사 과정 재학 중

1987년 3월 ~ 현재 : 한국전자통신연구원 통신디지털회로팀 선임연구원

※ 주관심분야 : DSP 설계, 음성신호 코딩

▲ 김 경 수(Kyung Soo Kim)

1977년 2월 : 서강대학교 전자공학과(공학사)  
1977년 2월 ~ 1985년 : 한국전자기술연구소  
1986년 ~ 현재 : 한국전자통신연구원 책임연구원