

TMS320VC5510 DSP를 이용한 AMR 음성부호화기의 실시간 구현

김 준(국방과학연구소), 배건성(경북대)

<차 례>

- | | |
|---------------------------------|-------------------------------|
| 1. 서론 | 3.1. TMS320VC5510 DSK 및 개발 환경 |
| 2. AMR 음성부호화 시스템 | 3.2. 실시간 처리를 위한 최적화 과정 및 실험결과 |
| 3. TMS320VC5510 DSK를 이용한 실시간 구현 | 4. 결론 |

<Abstract>

Real-Time Implementation of AMR Speech Codec Using TMS320VC5510 DSP

Jun Kim, Keun Sung Bae

This paper focuses on the real time implementation of an adaptive multi-rate (AMR) speech codec, that is a standard speech codec of IMT-2000, using the TMS320VC5510. The series of TMS320VC55x is a 16-bit fixed-point digital signal processor (DSP) having low power consumption for the use of mobile communications by Texas Instruments (TI) corporation. After we analyze the AMR algorithm and source code as well as the structure and I/O of TMS320VC55x, we carry out optimizing the programs for real time implementation. The implemented AMR speech codec uses 55.2 kbyte for the program memory and 98.3 kbyte for the data memory, and it requires 709,878 clocks, i.e. about 3.5 ms, for processing a frame of 20 ms speech signal.

* Keywords: AMR speech codec, TMS320VC5510, Real-time implementation.

1. 서 론

이동통신 시스템에서 음성을 전송하거나 디지털 녹음기에서 음성을 저장하기 위해서는 음성신호의 압축 및 복원 과정, 즉, 음성부호화 알고리즘이 필요하다. IMT-2000 서비스를 위한 음성부호화기 표준으로 European Telecommunications Standards Institute (ETSI) 및 Third Generation Partnership Project (3GPP)에서는 무선 채널의 환경 변화에서도 가변 전송률을 이용하여 양호한 음질을 유지할 수 있는 adaptive multi-rate (AMR) 방식을 선정하였다. 음성부호화기는 부호화 방식에 따라 파형부호화기, 파라미터부호화기 또는 보코더(vocoder), 혼합부호화기로 분류할 수 있으며, 또한 전송속도에 따라서 저속부호화기(low-rate codec), 중속부호화기(medium-rate codec), 고속부호화기(high-rate codec)로 분류할 수 있다. AMR 음성부호화기는 파형부호화 및 파라미터부호화 방식의 혼합 형태인 혼합부호화 방식으로, algebraic code-excited linear prediction (ACELP) 알고리즘을 기반으로 하는 8개의 전송모드를 갖는 multi-rate 음성부호화기이다.

음성부호화 알고리즘을 휴대단말기와 같은 실제 시스템에 적용하기 위해서는 digital signal processor (DSP)를 이용한 실시간 구현에 관한 연구를 필요로 한다. 또한 휴대단말기에서는 부동소수점 DSP에 비해 가격 대비 성능이 우수하면서 저전력 특성을 갖는 고정소수점 형태의 DSP를 주로 사용하게 된다. 본 연구에서는 AMR 음성부호화 알고리즘 및 ETSI에서 공개한 ANSI C 소스 코드[1]를 분석하고 최적화 작업을 수행하여 Texas Instruments (TI) 사의 고정소수점 DSP인 TMS320VC5510을 장착한 DSP starter kit (DSK)에서 실시간으로 구현하였다. DSP로 구현된 시스템의 결과가 ETSI에서 공개된 ANSI C 소스 코드를 수행시킨 결과와 일치함을 검증하였으며, 마이크와 스피커를 이용한 실시간 사운드 입/출력 실험을 통해서도 음질의 왜곡 없이 실시간 동작함을 확인하였다.

본 논문의 구성은 다음과 같다. 2장에서는 AMR 음성부호화 시스템의 구성과 음성부호화 알고리즘을 간략하게 서술한다. 3장에서는 TMS320VC5510 DSP를 이용한 실시간 구현 과정 및 실험결과를 제시하고, 마지막으로 4장에서 결론을 맺는다.

2. AMR 음성부호화 시스템

AMR 음성부호화기는 8개의 전송률을 갖는 음성부호화기, 입력 신호의 음성/비음성 여부를 판단하는 voice activity detector (VAD), 비음성 구간에서 전송률을 낮추기 위한 컴포트 노이즈(comfort noise) 생성을 포함하는 source controlled rate (SRC) 부분, 그리고 전송 에러나 패킷손실의 영향을 극복하기 위한 에러 복구 메카니즘(error concealment mechanism) 등으로 구성된다. AMR에서 사용되는 음성부

<표 1> AMR 음성 부호화기의 전송모드 및 비트율

Codec mode	Bit rate (kbit/s)
AMR_12.20	12.20 (GSM EFR) ^{a)}
AMR_10.20	10.20
AMR_7.95	7.95
AMR_7.40	7.40 (IS-641) ^{b)}
AMR_6.70	6.70 (PDC-EFR) ^{c)}
AMR_5.90	5.90
AMR_5.15	5.15
AMR_4.75	4.75
AMR_SID ^{d)}	1.80

a) GSM-EFR: 3GPP TS 26.090 Enhanced Full Rate

b) IS-641: TIA/EIA IS-641 TDMA Enhanced Full Rate

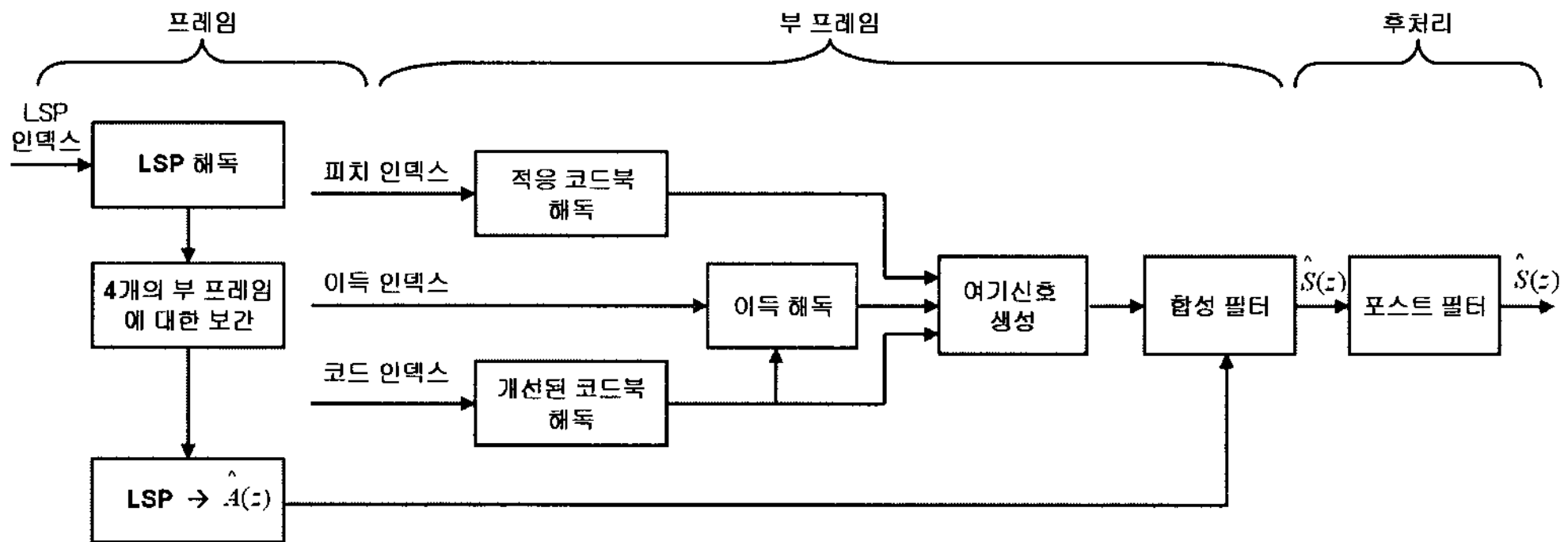
c) PDC-EFR: ARIB 6.7 kbit/s Enhanced Full Rate

d) SID: Silence Descriptor

호화기는 4.75 kbit/s에서 12.2 kbit/s 사이의 8가지 전송률에 silence descriptor (SID) 모드를 가진 하나의 통합된 음성부호화기이다. 8 kHz로 샘플링 되는 음성신호에서 20 ms의 한 프레임마다 무선 채널 환경에 따라 더 강인한 채널코딩을 할 수 있도록 전송모드를 바꾸어 줌으로써 양호한 음질을 유지할 수 있다. AMR 음성부호화기의 모드에 따른 전송률은 <표 1>과 같다.

SCR 부분은 음성의 비활성을 고려하여, 음성 프레임이 배경잡음만을 가진 경우에 보통의 비트율보다 낮은 비트율로 부호화되도록 하는 기능을 수행한다. 이는 단말기의 배터리 수명을 연장시킬 수 있으며, 부호화에 요구되는 평균 비트율을 감소시킬 수 있게 한다. 이러한 SCR의 동작을 위해서 송신측에서는 전송하는 프레임에 음성의 포함 여부를 판정하는 VAD 과정이 이루어져야 하며, 송신단의 배경잡음을 추정하여 수신단에 특징 파라미터 값으로 전달하여야 한다. 수신단에서는 일정한 간격으로 보내지는 SID 프레임의 정보를 바탕으로 음성이 포함되지 않은 프레임이 수신되는 동안에 컴포트 노이즈를 발생시켜 출력되는 음질의 자연스러움을 유지하게 만든다[2].

AMR 음성부호화기의 인코더 구조를 <그림 1>에 나타내었다[3]. 8 kHz로 샘플링된 음성 프레임(20 ms/프레임)에 대해서 3가지 파라미터, 즉 LP 필터 계수, 코드북 이득과 인덱스를 추출하게 된다. 이때 8개의 코덱 모드에 따라 인코딩 과정에 약간의 차이가 있다. 우선 각 파라미터에 할당되는 비트 수가 달라지는데, 한 프레임당 244 비트(12.2 kbit/s 모드)에서 95 비트(4.75 kbit/s 모드) 사이로 양자화되어 전송된다. 12.2 kbit/s 모드에서는 한 프레임마다 두개의 비대칭 창 함수를 이용하여 두 번의 LP 분석이 수행되고 나머지 7가지 모드에서는 한번의 LP 분석이 수행



<그림 2> AMR 디코더의 구조

AMR 디코더는 수신된 파라미터, 즉, LP 파라미터, 적응 코드북 벡터, 적응 코드북 이득, 고정 코드북 벡터, 고정 코드북 이득을 이용하여 음성합성 과정을 수행하게 되는데, 포스트 필터링(post filtering)과 업 스케일링(up-scaling) 과정을 거친 후 최종적인 합성음이 만들어지게 된다. <그림 2>에 AMR 음성부호화기의 디코더 구조를 나타내었다.

3. TMS320VC5510 DSP를 이용한 실시간 구현

3.1. TMS320VC5510 DSP 및 개발 환경

본 연구에 사용한 DSP는 TI 사의 저 전력 고성능 모델인 C55x계열 중 고정소수점 DSP인 TMS320VC5510이다. 200 MHz의 클럭 rate에서 최대 400 MIPS의 성능을 발휘할 수 있으며, 기본적으로 17-bit × 17-bit multiplier와 40-bit의 adder, non-pipelined single-cycle multiply-accumulate (MAC)로 구성된다. 또한 칩 내부에 64 kbyte의 DARAM과 256 kbyte의 SARAM이 존재하며 전체적으로 16 Mbyte에 해당하는 메모리 공간을 제공한다. 본 연구에서 사용한 DSP는 TMS320VC5510 DSP를 기반으로, AIC23 스테레오 오디오 코덱과 DMA 제어기, 외부 메모리로 8 Mbyte의 SDRAM과 512 kbyte의 플래쉬 메모리로 구성되어 있다. 시스템 개발을 위한 소프트웨어 툴로써 Code Composer Studio (CCS) version 2.12.07[5]을 이용하여 프로그램의 디버깅 및 실시간 분석을 수행하였다.

3.2. 실시간 처리를 위한 최적화 과정 및 실험결과

ETSI에서 제공되는 AMR 음성부호화 알고리즘의 ANSI C 소스 코드는 알고리즘 검증 및 프로그램 테스트를 위한 과정들이 포함된 범용 형태이므로 DSP를 이용한 실시간 구현을 위해서는 먼저 소스 코드 분석과정을 통하여 파일 입출력 부분이나, 연산 횟수, 함수의 호출을 계산하는 부분 등과 같이 실제 음성부호화 알고리즘과 관련 없는 부분을 프로그램 코드에서 제외하는 작업이 수행되어야 한다. 그런 다음, C-레벨에서의 소스 분석과 병행하여 TMS32VC5510 DSP 상에서 AMR 음성부호화 알고리즘을 실행하는데 소요되는 전체 클럭 수를 줄이기 위해 다음과 같은 최적화 작업을 수행하였다[6][7].

- 1) AMR 코드가 포함하고 있는 ETSI의 기본 산술연산 함수를 TMS320VC5510 DSP에서 가장 효율적으로 구성된 내장 함수로 대체하였다.
- 2) CCS에서 제공하는 컴파일 옵션을 최대한 활용하여 불필요한 코드의 사용을 최소화하였으며, 간단한 함수를 인라인(in-line) 함수로 대체하였다.
- 3) 내부 메모리의 접근 속도가 외부 메모리의 접근 속도보다 빠른 장점을 이용하기 위해 내부 메모리의 사용을 최대화하였다.
- 4) Loop trip counter가 작을 경우, loop 내부의 연산을 직접 반복 사용하여 for 문의 사용빈도를 최소화하였다.
- 5) TI사에서 제공하는 C55x DSP에 최적화된 신호처리 연산자를 포함하고 있는 DSP 라이브러리[8]를 이용하여 AMR 연산 과정의 autocorrelation과 convolution 등을 대체하였다.

위와 같은 최적화 과정을 통하여 구현된 AMR 음성부호화 시스템의 성능을 평가하기 위해 프로그램의 실행 속도와 메모리 사용량을 측정하였다. 프로그램의 실행 속도는 음성부호화 과정이 수행되는데 필요한 클럭 수를 측정하였으며, 프로그램 및 데이터 메모리의 측정은 COFF 링크시 생성한 메모리 맵 파일을 참조하여 측정하였다. <표 2>는 AMR ANSI C 소스 코드를 최적화 과정 없이 DSK에서 구현한 결과에서 전송률이 제일 높은 12.2 kbit/s 경우에 대한 것을 보인 것이며, <표 3>은 본 논문에서 구현한 AMR 음성부호화기에 대한 실험 결과를 나타낸 것이다. 묶음 구간을 비롯하여 유성음, 무성음 및 천이 구간을 포함한 음성 데이터를 이용하여 인코딩 및 디코딩 과정에서 프레임 당 최대로 소요되는 클럭 수를 측정하여 제시하였는데, 대응되는 수행시간은 식 (1)과 같이 계산하였다.

$$\text{프레임 당 최대 수행시간} = \text{최대 clock 수} \times \frac{1}{200\text{MHz}} \quad (1)$$

<표 2> 최적화 작업 수행전의 AMR 음성부호화기 성능

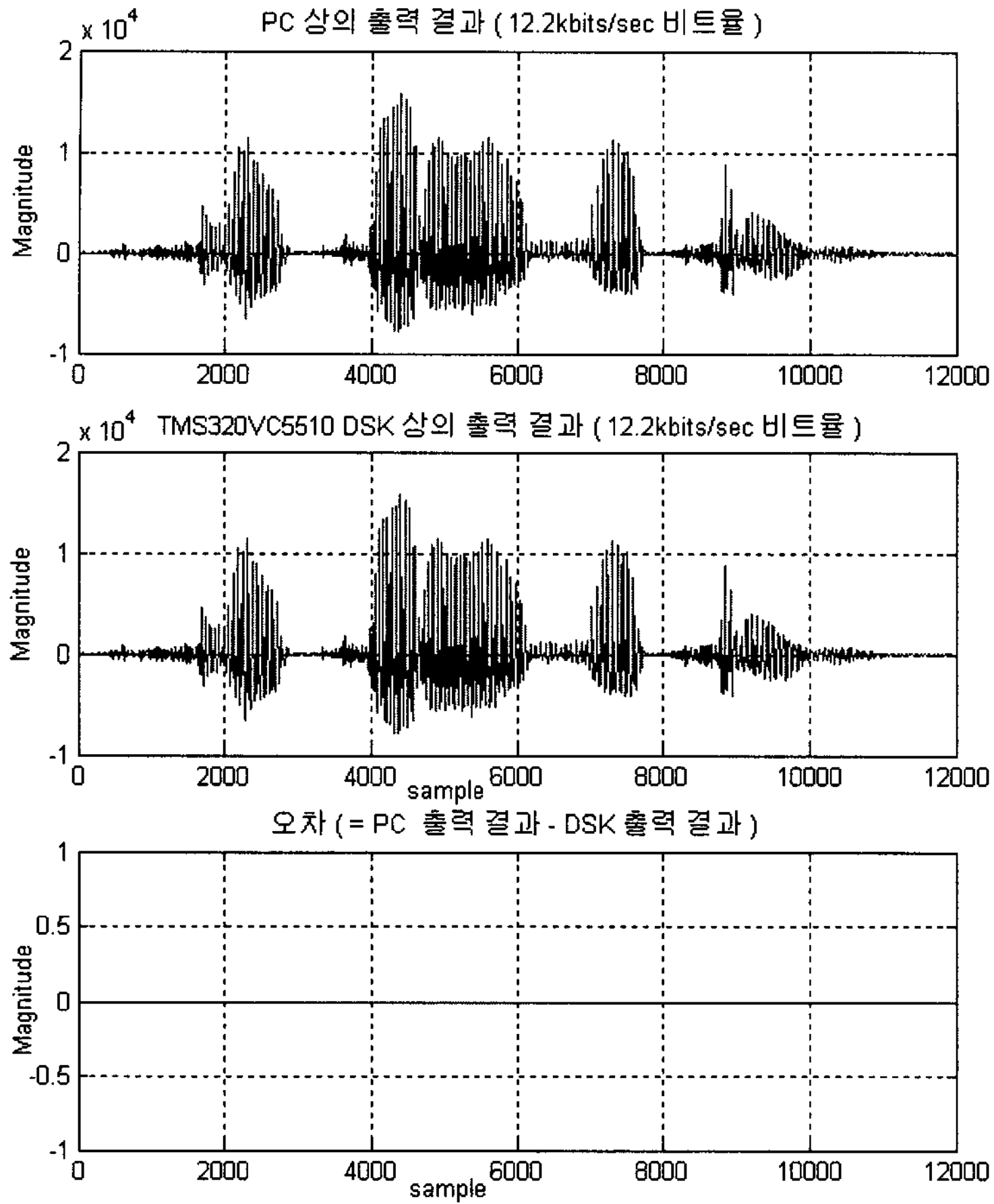
모드 (kbit/s)	수행속도 (Clock/Frame)			메모리 사용량 (kbyte)	
	부호화기	복호화기	전체	프로그램	데이터
12.2	4,196,920	787,370	5,130,908 (25.7 ms)	113.4	191.6

<표 3> DSK에 구현된 AMR 음성부호화기 성능

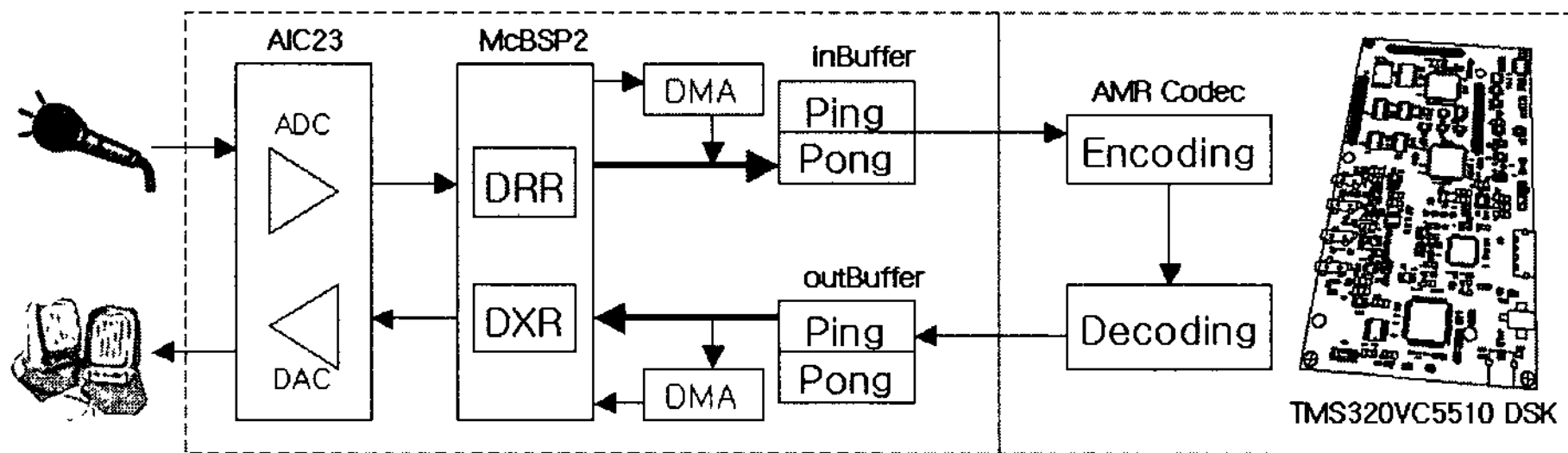
모드 (kbit/s)	수행속도 (Clock/Frame)			메모리 사용량 (kbyte)	
	부호화기	복호화기	전체	프로그램	데이터
12.2	611,325	74,515	709,878 (3.54 ms)	55.2	98.3
10.2	546,925	73,160	647,318	55.2	98.3
7.95	549,692	72,533	649,458	55.2	98.3
7.40	493,179	69,899	590,311	55.2	98.3
6.70	546,240	71,299	644,772	55.2	98.3
5.90	418,740	71,034	517,007	55.2	98.3
5.15	336,481	71,053	434,767	55.2	98.3
4.75	408,013	74,990	510,236	55.2	98.3

음성부호화기의 프레임 당 연산시간이 전송모드에 따라 다소 차이는 있으나 인코딩/디코딩 과정을 합쳐서 2.17~3.54 ms의 수행 시간이 필요함을 볼 수 있다. 따라서 TMS320C5510 DSK에서 최적화 작업을 거쳐 구현된 AMR 음성부호화기는 전이중방식 모드로 5~6개의 채널을 수용할 수 있다고 볼 수 있다.

실시간으로 구현된 AMR 음성부호화기가 올바르게 동작하는지를 확인하기 위해 임의의 음성신호에 대해 PC에서 얻어진 ANSI C 소스 코드의 출력과 DSP에 구현된 AMR 음성부호화기를 통해 얻어지는 출력신호를 비교하였다. <그림 3>은 각각의 결과를 나타낸 것으로서 출력 값이 오차 없이 서로 동일한 결과를 가짐을 확인하였다. 또한, <그림 4>와 같이 마이크 입력을 DSK 보드 상의 AIC23 stereo codec[9]을 통해 8 kHz로 샘플링하고, 샘플링된 음성신호를 DSK 상에서 구현된 AMR 음성부호화기의 부호화와 복호화 과정을 거쳐 스피커로 출력되도록 하여 실시간 동작을 통한 합성음을 확인하였다. 프레임 사이의 시간 간격으로 인한 음질의 저하를 방지하기 위해 더블-버퍼링 방식을 적용하였으며, 8개의 모드를 프레임 단위로 변화시키면서 테스트한 결과 시간적인 지연이나 음질의 왜곡 없이 실시간 구현됨을 확인할 수 있었다.



<그림 3> PC 및 DSK 상에서 얻어진 출력 신호 비교



<그림 4> 실시간 구현을 확인하기 위한 실험 구성도

4. 결 론

본 논문에서는 AMR 음성부호화기를 TMS320VC5510 고정소수점 DSP 기반의 DSK에서 실시간으로 구현하고, 그 결과를 제시하였다. 구현된 시스템은 프로그램 메모리 55.2 kbyte와 데이터 메모리 98.3 kbyte를 필요로 하며, 20 ms의 한 프레임을 처리하는데 사용되는 클럭 수는 최대 709,878로 200 MHz의 DSP 클럭을 고려할 때 약 3.5 ms의 수행시간이 소요되었다. 따라서 전이중방식 모드로 최대 5 채널을 수용할 수 있다. DSK상에서 구현된 AMR 음성부호화기의 출력과 ETSI에서 배포한 ANSI C 소스를 PC 상에서 구현한 결과를 비교하여 서로 일치함을 확인하였으며, DSK 상에서 실시간 음성 입출력을 통하여 AMR 음성부호화기가 지연이나 음질의 왜곡 없이 실시간으로 동작함을 확인하였다.

참 고 문 헌

- [1] 3GPP TS 26.073 V7.6, *Adaptive Multi Rate (AMR) Speech; ANSI-C Code for the AMR Speech Codec*, June 2007.
- [2] 3GPP TS 26.071, *Mandatory Speech Codec Speech Processing Function; AMR Speech Codec; General Description*, July 2007.
- [3] ETSI Draft EN 301 704, *Digital Cellular Telecommunication System; Adaptive Multi-Rate (AMR) Speech Transcoding*, May 2000.
- [4] Texas Instruments, *TMS320C55x DSP Functional Overview*, June 2000.
- [5] Texas Instruments, *Code Composer Studio Getting Started Guide*, Nov. 2001.
- [6] Texas Instruments, *TMS320C55x DSP Programmer's Guide*, Aug. 2001.
- [7] J. W. Seo, H. S. Kwon, M. H. Park, K. S. Bae, "Real-implementation of AMR speech coding algorithm using TMS320C6201", *Proc. International Conference on Electronics, Information and Communications*, pp. 267-270, 2000.
- [8] Texas Instruments, *TMS320C55x DSP Library Programmer's Reference*, Oct. 2002.
- [9] Texas Instruments, *TLV320AIC23 Stereo Audio CODEC Data Manual*, Apr. 2002.

접수일자: 2008년 2월 22일

게재결정: 2008년 3월 21일

▶ 김 준(Jun Kim)

주소: 대전 유성우체국 사서함 35-1호

소속: 국방과학연구소 무인자율화 연구실

전화: 042) 821-2839

E-mail: Wns3381@add.re.kr

▶ 배건성(Keun-Sung Bae) : 교신저자
주소: 702-701 대구광역시 북구 산격동 1370번지
소속: 경북대학교 전자전기컴퓨터학부
전화: 053) 950-5527
E-mail: ksbae@ee.knu.ac.kr