16 비트 고정 소수점 DSP를 이용한 다채널 G.729A 음성 부호화기의 실시간 구현

Real-time Implementation of a Multi-channel G.729A Speech Coder on a 16 Bit Fixed-point DSP

안 도 건*, 유 승 균*, 최 용 수*, 이 재 성*, 강 태 익*, 박 성 현*
(Do-Keun Ahn*, Seung-Kyun Ryu*, Yong-Soo Choi*, Jae-Sung Lee*, Tae-Ik Kang*, Sung-Hyun Park*)

요 약

본 논문에서는 16 비트 고정 소수점 DSP(Digital Signal Processor)를 사용하여 다채널 G.729A 음성 부호화기를 실시간 구현하였으며, 실제로 음성 사서함 서비스(Voice Mailing Service: VMS) 시스템에 응용하였다. DSP는 TI(Texas Instruments) 사의 TMS320C549 칩을 사용하였으며, 구현된 G.729A음성 부호화기는 채널 당 부호화기에 14.5 MIPS를, 복호화기에 3.6 MIPS를 소요하였으며, 메모리는 코드 부분에 9.88 K 워드, 데이터 부분에 1.69 K 워드를 필요로 하였다. 결과적으로 개발된 VMS 시스템에는 두 개의 DSP를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실험 결과, 구현된 다채널 부호화기는 ITU-T에서 제공된 테스트 벡터 샘플을 모두 통과하는 일관된 성능을 보였다.

핵심어: 음성 부호화기, G.729A, 다채널, 실시간 구현, 음성 사서함 서비스
투고 분야: 음성처리 분야(2.2, 2.1)

ABSTRACT

This paper describes real-time implementation of a multi-channel G.729A speech coder using a 16 bit fixed-point Digital Signal Processor (DSP) and its application to a Voice Mailing Service (VMS) system. TMS320C549 by Texas Instruments was used as a fixed point DSP chip and a 4 channel G.729A coder was implemented on the chip. The implemented coder required 14.5 MIPS for the encoder and 3.6 MIPS for the decoder at each channel. In addition, memories required by the coder were 9.88K words and 1.69K words for code and data sections, respectively. As a result, the developed VMS system that accommodates two DSP chips was able to support totally 8 channels. Experimental results showed that the our multi-channel coder passes all of test vectors provided by ITU-T.

Keywords: Speech coder, G.729A, Multi-channel, Real-time implementation, VMS.

I. 서 론

지난 십 여년간 낮은 복잡도와 전송률에서 우수한 음질을 얻으려는 연구의 결과로 현재 여러 가지 음성 부호화기 표준들이 국제 기구에 의해 채택되어 사용되고 있다.

지금까지 제안된 음성 부호화기들 중에서 우수한 음질을 갖는다고 알려져 있는 CELP(Code-Excited Linear Prediction)[1]는 표준 음성 부호화 알고리듬의 주류를 이루고 있으며, 현재도 성능 개선 연구가 계속되고 있다. CELP의 성능 개선 연구는 크게 계산량 감소와 음질 개선의 두 가지 분야로 나눌 수 있다. 제안될 당시의 초기의 CELP는 많은 계산량으로 인해 실시간 구현이 불가능하였지만 이후 많은 고속 알고리듬들이 제안되어 현재는 단일 범용 디지털 칩으로도 실시간 구현이 가능하게 되었다. 그 대표적인 예로 G.723.1[2], G.729[3][4], GSM-EFR(Global System for Mobile communications Enhanced Full Rate speech codec)[5] 등과 같은 최근의 음성 부호화기 국제 표준에 사용되고 있는 ACELP(Algebraic CELP)[6]를 들 수 있다.

최근에 멀티미디어 응용 분야에서 음성과 데이터를 동시에 전송하거나 저장하는 문제가 큰 관심사가 되었다. 이러한 응용 분야에서는 음질뿐만 아니라 복잡도가 음성 부호화기를 설계하는데 있어서 중요한 고려 사항이 된다.

* LG정보통신 디지털 네트워크연구소 미디어기기실
접수일자 : 2000년 1월 17일

1995년 2월, ITU-T(International Telecommunication Union-Telecommunication)에서는 멀티미디어 DSVD(Digital Simultaneous Voice and Data) 응용 분야에 사용하기 위한 음성 부호화기의 표준화를 위한 G.DSVD라는 전문가 그룹을 만들었다. 그 당시 G.728, G.723.1 그리고 G.729와 같은 표준 음성 부호화기들은 단일 프로세서로 모뎀 알고리듬과 음성 부호화 알고리듬을 통합 처리하기에는 계산량이 과도하였다. 따라서 이런 점을 고려하여 계산량은 10 MIPS 이하, RAM은 2K 워드, ROM은 8K 워드 이하라는 제한 조건들이 새로운 음성 부호화기 표준의 요구 사항으로 반영되었다. 1999년 여름, AT&T, AC/DSPG, NTT, Rockwell, USH 등에서 제안된 5 가지의 부호화기들이 DSVD 표준 후보로 제출되었다[7][8]. 최종적으로 1995년 11월에 G.729와의 부호화 비트열의 상호 호환성이 중요하게 고려되어 USH에서 제안된 G.729의 계산량 감소된 버전이 ITU-T의 V.70 시리즈(DSVD)와 G.729A(G.729 Annex A)[7-9]의 음성 부호화기 표준으로 확정되었다. G.729A는 멀티미디어 화상 회의, 이동 화상 전화 서비스, 대화형 게임, 원거리 강의 및 발표, 전자 상거래, 원거리 의료 검진, 자택 근무 등에 응용될 수 있다. 또 다른 응용 분야로는 최근에 큰 관심사가 되고 있으나 아직까지 음성 부호화 표준이 정해져 있지 않은 인터넷 전화와 인터넷 음성 메일 등을 들 수 있다. 현재 인터넷 전화에 사용되는 화상 전화(H.324[10])의 표준 음성 부호화기인 G.723.1에 비해서 G.729A는 낮은 복잡도와 낮은 지연 등 인터넷 전화용 음성 부호화기로서 상당한 장점을 갖고 있다.

본 논문에서는 16 비트 고정 소수점 연산을 갖는 범용 DSP(Digital Signal Processor)를 사용하여 다채널 G.729A 음성 부호화기를 실시간 구현하였으며, 실제로 음성 사서함 서비스(Voice Mailing Service: VMS)[11] 시스템에 응용하였다. 범용 DSP로는 100 MIPS의 처리 용량과 DSP 내부에 32K 워드 RAM, 8 K워드 DARAM(Dual Access RAM)과 24K 워드 SARAM(Single Access RAM)을 갖는 TI(Texas Instruments)사의 TMS320C549[12]를 사용하였으며, VMS 시스템에는 두 개의 DSP를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실제 응용 분야로 선택한 VMS는 수신자가 부재시 송신자가 음성을 남길 수 있는 기능을 가진 서비스로서 시스템 구성은 주 처리기, DTMF(Dual Tone Multiple Frequency) 처리용 DSP와 부호화된 음성 신호를 저장하기 위한 플래시 메모리를 장착하고 있고 두 개의 부호화기 용 DSP를 장착하고 있어 동시에 8 채널을 지원한다.

본 논문의 구성은 다음과 같다. II장에서는 G.729A의 알고리듬에 대해서 설명하고, III 장에서는 다채널 G.729A 실시간 구현 및 실제 응용에 대해서 자세히 살펴 보고, 마지막으로 IV 장에서 결론을 맺는다.

II. G.729 Annex A 음성 부호화기

G.729A[7-9]는 8kH로 샘플링 된 16 비트 PCM(Pulse Code Modulation) 신호를 입력으로 하고, 그 기본 구조는 G.729[3][4]와 같이 10 ms(80 샘플) 음성 프레임에서 동작하는 8 kbps CS-ACELP(Conjugate-Structured Algebraic CELP)[13][14] 알고리듬에 바탕을 두고 있다. 또한 G.729A로 부호화된 비트열은 G.729로 복호화될 수 있으며, 그 반대의 경우도 가능하다. G.729A는 DSP를 이용한 실시간 구현에 용이하도록 G.729에 몇가지 고속 알고리듬을 도입함으로써 그 계산량을 약 50% 정도 감소시켰으며, 음질 면에서는 약간의 음질 저하가 발생하는 3 탠덤(tandem)과 배경 잡음 환경을 제외한 대부분의 상황에서 G.729과 32 kbps G.726에 상응하는 음질을 갖는다. 다음으로 G.729와 G.729A의 기반이 되고 있는 CS-ACELP에 대해 설명한 후, G.729A의 특징이 되는 몇 가지 고속 알고리듬에 대해 설명한다.

8 kbps CS-ACELP는 기본적으로 합성을 통해서 분석을 하는 AbS(Analysis-by-Synthesis) 방식의 CELP 부호화기이며 기본 개념도는 그림 1과 같으며 비트 할당은 표 1과 같다.

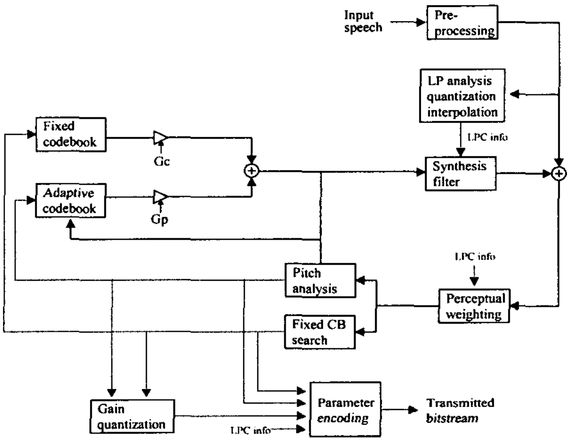


그림 1. ITU-T 8kbps CS-ACELP(G.729 & G.729A) 부호화기
Fig. 1. Block diagram of the ITU-T 8kbps CS-ACELP(G.729 & G.729A) encoder.

표 1. ITU-T 8kbps CS-ACELP(G.729 & G.729A)의 비트 할당
Table 1. Bit allocation of the ITU-T 8kbps CS-ACELP(G.729 & G.729A).

Parameter	Subframe 1	Subframe 2	Total
LSP coefficients			18
Pitch delay	8	5	13
Delay parity bit	1		1
Codebook position index	13	13	26
Codebook signs index	4	4	8
Gains VQ	3+4	3+4	14
Total			80

8 kbps CS-ACELP는 10 ms 크기의 프레임을 가지며 이것은 다시 5 ms의 부프레임으로 분할된다. 우선, 선형

예측(Linear Prediction: LP) 계수를 추출하기 위해 비대칭 창 함수를 사용함으로써 총 알고리듬 지연 시간을 감소시키며 전처리 과정으로 고역 통과 필터를 사용하여 입력 신호의 직류 성분(DC) 성분을 제거하고 오버플로우(overflow)를 방지하기 위하여 입력 신호의 크기를 절반으로 줄인다.

Durbin 알고리듬을 통해 얻어진 10차의 LP 계수는 양자화와 보간 특성이 우수하고 안정도 검사가 용이한 특성을 갖는 LSP(Line Spectrum Pair) 계수로 변환되고, 이 값과 4차의 MA(Moving Average) 예측기를 통해 예측된 값과의 차이는 2단계의 벡터 양자화(Vector Quantization: VQ) 및 분리 벡터 양자화(Split VQ: SVQ)를 거쳐서 부호화된다. 이때 서로 다른 계수 값을 갖는 2종류의 MA 예측기가 사용되며 그 중 각각의 예측기에 의해 구해진 예측 값과 LSF계수 값 사이의 가중된 오차 에너지(Weighted Mean Squared Error: WMSE)를 최소로 하는 것을 선택한다. 이렇게 양자화된 LSP 계수는 LP 합성 필터를 구성하기 위해 부프레임 단위로 보간된 후 다시 LP 계수로 복원된다.

피치 지연은 프레임 단위의 개회로 검색과 부프레임 단위의 폐회로 방식의 적응 코드북 검색의 두 과정을 순차적으로 거쳐 구해진다. 첫 번째 부프레임에서의 적응 코드북 검색은 개회로 피치 지연을 중심으로 몇 개의 지연 값에 대해서만 검색을 하여 최적의 실수 피치 지연을 구한다. 두 번째 부프레임에서는 검색의 효율을 높이기 위해 첫 번째 부프레임의 구한 피치 지연 값을 기준으로 한다. 적응 코드북 검색에 의해 입력 신호로부터 주기적인 피치 성분이 제거되면 여기 신호 추출의 마지막 단계인 고정 코드북 검색이 이루어진다. 고정 코드북은 절대 크기가 1인 다중 펄스의 부호와 위치로 구성되어 있으며 각 펄스는 트랙(track)이라 불리는 제한된 위치에서만 검색된다는 것이 특징이다. 적응 코드북과 고정 코드북의 이득 값은 켤레(conjugate) 구조를 갖는 2개의 코드북을 사용하여 폐회로 방식으로 양자화된다. 이득 코드북 검색은 효율을 높이기 위해서 각각의 코드북을 단독으로 개회로 검색하여 몇 개의 후보를 정하고, 이 후보들에 대해서만 대해서만 폐회로 검색을 수행한다. 결과적으로 이러한 방법은 전체 검색(full search)에 비해서 상당한 계산량의 감소를 얻게 한다.

복호화기에서는 복원된 합성음의 음질 향상을 위해 후처리 과정을 거치게 된다. 후처리 과정에서는 적응 포스트필터링(adaptive postfiltering)과 고역 통과 필터링과 신호 크기 증폭(up-scaling)을 수행한다.

G.729A[7-9]에서 고속 알고리듬이 적용된 주요한 부분을 요약하면 LSP 변환, 인지 가중 필터, 개회로 피치 분석, 가중 합성 필터의 충격과 응답 계산, 적응 코드북 검색, 고정 코드북 검색 및 하모닉 포스트필터이다. 계산량 감소를 위해 사용한 방법들을 구체적으로 살펴보면 다음과 같다.

- LSP 변환: G.729에서는 0에서 까지 균등하게 분할된 60개의 지점에서 체비쉐프 다항식의 값을 계산하고 부호의 변환이 있으면 그 구간을 다시 4개로 나누어 보다 정밀하게 근을 검색한다. G.729A에서는 계산량 감소를 위하여 다항식의 근을 검색 지점의 수를 60에서 50으로 줄이고, 부호 변환시에 정밀 검색 횟수를 4에서 2로 줄인다.
- 인지 가중: G.729와는 달리 양자화된 LP 계수를 사용한다.

$$W(z)=\frac{\hat{A}(z)}{\hat{A}(z/\gamma)}. \tag{1}$$

여기서 가중 인자 $\gamma=0.75$ 으로 고정되며 G.729에서 사용하는 가중 인자의 적응 과정은 사용되지 않는다. 이러한 가중 필터의 변화는 가중 필터와 LP 합성 필터의 결합된 형태인 가중 합성 필터를 $W(z)/\hat{A}(z)=1/\hat{A}(z/\gamma)$ 으로 간략화하고 결과적으로, 이것은 충격과 응답과 목적 신호 계산과 필터 상태를 갱신하는데 필요한 계산량을 감소시킨다. 가중 필터의 간략화는 평탄한 주파수 응답 특성을 갖는 입력 신호에 대해서는 약간의 음질 저하를 가져오지만 대부분의 경우에 대해서는 음질 차이는 거의 없다.

- 개회로 피치 분석: 최적의 적응 코드북 지연 검색 과정의 계산량을 줄이기 위해서 검색 범위는 개회로 분석한 피치 지연을 주위로 제한된다. 이러한 개회로 피치 분석은 한 프레임(10 ms) 당 한 번 수행되고, 가중된 신호를 바로 사용하는 G.729에서 달리 저역 통과 필터를 거친 가중된 음성 신호 $s_w(n)$를 입력으로 사용한다. 이 과정은 3 단계로 구성되는데, 그 첫 단계에서는 [20, 30], [40, 79], [80, 143]의 세 구간에서 다음 식으로 주어진 상관 값을 최대로 하는 각 구간의 지연 값을 찾는다.

$$R(k)=\sum_{n=0}^{39} s_w(2n)s_w(2n-k). \tag{2}$$

다음으로 이렇게 구해진 구간 최대치 $R(t_i), i=1,\ldots,3,$ 는 다음과 같이 정규화되고,

$$R'(t_i)=\frac{R(t_i)}{\sqrt{\sum_{n=0}^{39} s_w^2(2n-t_i)}}, i=1,\ldots,3, \tag{3}$$

마지막으로, 정규화된 구간 최대치들은 낮은 지연에 선호도를 주어서 비교하여 최종 지연 값으로 선택한다. 결과적으로 식 (2)로 주어진 상관 값을 계산하는 과정에서 단지 짝수 샘플만을 사용하고, 세 번째 구간 [80, 143]에서는 짝수 지연에 해당하는 상관 값만을 계산한 후, 선택된 짝수 지연의 ±1에서의 지연만을 검색함으로써 계산량을 줄인다.

- 적응 코드북 검색: 적응 코드북 구조는 G.729와 동일하다. 첫 번째 부프레임에서는 [19+1/3, 84+2/3] 구간에서는 실수 피치 지연이 사용되고, [85, 143] 구간에서는 정수 피치 지연이 사용된다. 두 번째 부프레임

에서는 항상 [$int(T_1)$-5-2/3, $int(T_1)$+4+2/3] 구간에서 1/3 정밀도를 갖는 실수 피치 지연이 사용된다. 여기서 $int(T_1)$는 첫번째 부프레임의 실수 피치 지연 T_1의 정수 부분이다. 적응 코드북 검색은 다음과 같이 주어진 식을 최대로 하는 지연을 선택함으로써 이루어진다.

$$R(k)=\frac{\sum_{n=0}^{39}x(n)y_k(n)}{\sqrt{\sum_{n=0}^{39}y_k^2(n)}}, \tag{4}$$

여기서 $x(n)$은 목적 신호이고, $y_k(n)$은 지연 k에서 필터링된 과거의 여기 신호이다. G.729A에서는 계산량 감소를 위해서 식 (4)의 분자항만을 고려하여 그것을 최대로 하는 지연 검색 방법을 도입한다.

$$R(k)=\sum_{n=0}^{39}x(n)y_k(n)=\sum_{n=0}^{39}x_b(n)u_k(n) \tag{5}$$

여기서 $x_b(n)$은 역방향 필터링된 목적 신호이고, 다시 말해서 입력 신호와 가중 합성 필터의 상호 상관이고, $u_k(n)$은 지연 k에서 과거의 여기 신호이다. 각 부프레임에서 최적 피치 지연 T_1 와 T_2 를 결정하는 과정을 살펴 보면, 만약 (5)를 사용하여 찾은 최적 정수 지연이 85 이상인 경우에는 실수 피치 검색 과정을 생략하고, 그 이하인 경우에는 최적 정수 지연을 주위로 하여 -1/3, 0, 1/3에서 과거의 여기 신호를 보간하여 (5)를 최대로 하는 실수 지연을 선택한다. 이와 같은 적응 코드북 검색의 간략화로 인해 G.729에 비해서 1/3 정도의 오차를 가져오지만 계산량은 1/2로 감소하고 주관적인 음질은 그 차이를 구별하기 어렵다.

- 고정 코드북 탐색: 17 비트 고정 코드북 구조도 G.729와 동일하다. 펄스 크기는 역방향 필터링된 목적 신호의 부호를 이용하여 코드북 검색 전에 미리 정하는 G.729에서의 SSPA(Signal-Selected-Pulse- Aamplitude) [3][4] 방법에 의해 결정한다. 그러나 펄스의 위치는 부프레임 당 총 2^{13} 중 최대 1440(17.5%)가지의 가능한 위치의 조합에 대해서 검색하는 nested-loop 방식을 사용한 G.729와는 달리 펄스마다 독립적인 루프를 사용하는 depth-first 방식을 사용하여 검색 횟수를 최대 320(3.9%)으로 줄였다. 이러한 고정 코드북 검색의 간략화는 약 0.2dB SNR 감소에 해당하는 음질 저하를 가져오지만 부호화기 전체 계산량 감소의 50%를 차지한다.
- 후처리 과정: 후처리 과정은 적응 포스트필터를 제외하고는 G.729와 동일하다. 적응 포스트 필터는 장구간 포스트필터, 단구간 포스트필터, 경사 보상 필터의 세 필터의 직렬 연결로 구성된다. G.729와의 주요한 차이는 하모닉 포스트필터링 과정이다. 하모닉 포스트 필터의 장구간 피치 지연은 전송된 현재 부프레임에서 피치 지연의 정수 부분 을 중심으로 ±3 범위의 정수 값만을 검색하여 G.729에서 1/8 단위로 실수 피치 지연을 찾는 과정을 제거함으로써 부가적인 계산량을 크게 줄였다.

III. 다채널 G.729A 실시간 구현 및 응용

G.729A 실시간 구현에 사용된 DSP는 TI(Texas Instruments)의 TMS320C549[12]로서 100 MIPS의 처리 용량을 갖고 있으며 DSP 내부에 32K 워드의 RAM을 가지고 있고 8K의 DARAM(Dual Access RAM)과 24K의 SARAM(Single Access RAM)으로 구성되어 있다. 외부 인터페이스로는 HPI(Host Port Interface), TDM(Time Division Multiplexed) 시리얼 포트와 두개의 버퍼 시리얼 포트(Buffered Serial Port)를 가지고 있다. TMS320C54x의 경우 프로그램 영역과 데이터 영역의 메모리 맵이 각각 존재하고 내부 RAM의 경우 기본적으로 데이터 영역으로 지정되어 있으나 메모리 맵 모드를 변경하여 프로그램과 데이터의 영역을 동시에 내부 RAM으로 지정할 수 있다. 본 논문에서도 실시간 구현 시스템의 처리 속도의 향상을 위하여 DSP의 내부 RAM만을 사용하였다.

다채널 실시간 구현은 알고리듬 및 C 소스 분석, 1 채널 부호화기 구현 및 검증, 다채널 부호화기 구현의 순으로 진행되었다.

3.1. 알고리듬 및 C 소스 분석

부호화기의 효율적인 구현을 위하여 알고리듬을 분석하고 C 소스를 검토하여 각 기능 별로 분류하였고 각 모듈의 복잡도와 점유율을 조사하였다. G.729는 부호화기와 복호화기로 구성 되어 있으며 부호화기의 경우 크게 전처리 과정, LP 계수 분석, 개회로 피치 검색, 적응 코드북 검색, 고정 코드북 검색으로 구분할 수 있다. 표 2는 G.729A 부호화기의 기능별 처리 부하(processing load)를 나타낸다.

표 2. G.729A 부호화기의 기능별 처리 부하
Table 2. Processing load of the G.729A encoder.

Pre processing	LP analysis	Open loop Pitch search	Adaptive codebook search	Fixed codebook search
1%	18%	10%	11%	52%

실제 구현에 앞서 시뮬레이션을 통하여 C 소스를 최적화하였으며 메모리의 효율적인 사용을 위하여 각 모듈에서 필요로 하는 스택(stack) 및 광역 변수의 양을 조사하였으며 광역 변수의 경우 다채널 구현을 위하여 별도의 관리가 필요하다.

3.2. 한 채널 실시간 구현 및 검증

다채널 음성 부호화기 구현에 앞서 한 채널 부호화기를 C 언어와 어셈블리어를 사용하여 구현하였으며 ITU-T에서 제공되는 테스트 벡터를 적용하여 검증하였다. 본

논문에서 사용하고 있는 DSP의 경우 프로그램 영역과 데이터 영역이 DSP내부 RAM에 존재해야만 최적의 성능을 발휘할 수 있다. 처리 속도 증가를 위한 또 다른 방법은 Dual-operand를 사용하는 방법이고, 이 경우 두개의 연산자가 DARAM 영역에 존재 하여야 한다. 본 논문에서는 DARAM 영역을 포함하는 어드레스의 하위 부분에 데이터 영역을 설정하고 나머지 상위 부분에 프로그램 영역을 설정하여 DSP의 처리 능력을 최적화하였다. 부호화기 검증용으로 사용한 테스트 벡터는 총 9개로서 부호화기와 복호화기용으로 6개의 벡터와 복호화기에만 사용하는 3개의 벡터로 구성되어 있고, 구현된 G.729A 부호화기 및 복호화기에 적용한 결과 모든 벡터가 비트 단위로 일치하였다. 입력 음성 신호는 시리얼 포트를 통하여 PCM(Pulse Code Modulation)방식의 a-law 또는 u-law PCM 데이터를 입력 받아 16 비트 선형PCM으로 변환 후 부호화기에 적용하였다. 표 3은 본 논문에서 구현한 1-채널 G.7294A의 복잡도 및 사용한 메모리 크기이다.

표 3. G.729A 의 MIPS 및 메모리(K words)
Table 3. Complexity of the G.729A in terms of MIPS and memory(K words).

G.729A	MIPS	CODE	DATA
Encoder	14.5	7.5	1.04
Decoder	3.6	3.4	0.92
Full	18.1	9.88	1.69

3.3. 다채널 확장 및 실제 시스템 응용

실시간 구현된 한 채널 음성 부호화기를 다채널로 확장하기 위해서 각 채널 별로 보관해야 하는 이전 프레임에 대한 정보들이 무엇인가를 분석하고, 음성 벡터, 여기 벡터, 양자화 이득 등을 채널 별로 분류하여 보관한다. 본 논문에서 구현된 다채널 부호화기에서는 이러한 정보들을 블록 형태로 분류하여 메모리에 위치시킨다.

G.729A의 경우 10ms 프레임 단위로 부호화 하고 한 프레임 시간 내에서 시분할하여 다채널 구현이 가능하다. 따라서 부호화기의 성능에 의해서 구현 가능한 채널의 수가 결정 되어지고 그림 2의 경우 부호화기의 최대 처리 시간이 2.2ms인 경우의 예로서 10ms 내에 4 채널의 부화화기를 구현할 수 있다

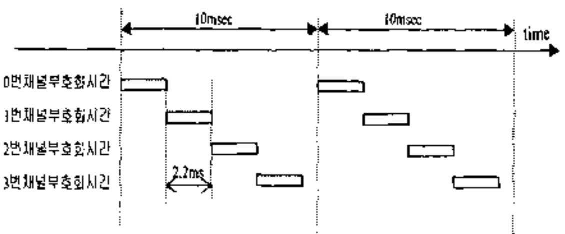


그림 2. 다채널 부호화기 타이밍도
Fig. 2. Timing diagram of a multi-channel coder.

다채널 부호화기의 동작 형태는 각 채널에 대한 우선 순위는 정해져 있지않고 순차적으로 각 채널 별 부호화 과정을 수행한다. 이것은 실제 구현에 있어서는 채널 분류기와 부호화기 두 가지의 기능으로 나누어 질 수 있다. 채널 분류기는 각 채널의 입력이 하나의 프레임으로 형성되면 입력된 프레임과 해당 채널의 이전 프레임에 대한 정보를 저장하고 있는 메모리의 시작 번지를 부호화기에 전달하고, 부호화기는 전달 받은 시작 번지를 기준으로 오프셋 번지를 이용하여 부호화한다. 채널 분류기와 부호화기를 분리하면 처리해야 할 채널의 수가 증가 할 때 이전 프레임에 대한 정보를 저장할 영역만 확장하면 되므로 확장이 용이하다. 그림 3은 본 논문에서 사용한 데이터 영역의 구성도이다.

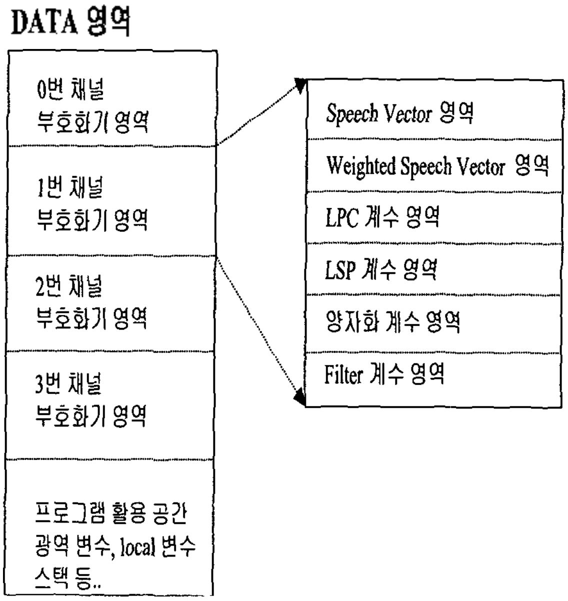


그림 3. DSP 데이터 영역 메모리 구성
Fig. 3. Data memory structure in a DSP.

본 논문에서는 4 채널 G.729A 부호화기를 실시간 구현하였으며 VMS(Voice Mailing Sevice) 시스템에 적용하였다.

VMS는 수신자가 부재시 송신자가 음성을 남길 수 있는 기능을 가진 음성 사서함 서비스로서 VMS 시스템의 구성은 주 처리기, DTMF(Dual Tone Multiple Frequency) 처리용 DSP와 부호화된 음성 신호를 저장하기 위한 플래시 메모리를 장착하고 있고 부호화기 용 DSP가 두 개 장착되어 있어 동시에 8 채널을 지원한다. 그림 4는 VMS 시스템 구성도이다. 그림 4에서 주 처리기는 전달 받은 DTMF를 이용하여 부호화기를 제어하고 압축된 음성 데이터를 플래시 메모리에 저장하거나 부호화기로 데이터를 전송한다. 음성 사서함 기능을 위하여 사용하는 제어 명령에는 녹음되어 있는 음성을 복원하기 위한 PLAY, 입력 음성을 압축하는 RECORD, 압축 또는 복원을 중지하는 IDLE, a-law / u-law PCM 선택 기능, 음성의 출력 레벨을

조정하는 기능 등이 포함 되어 있고 주 처리기에 의해 DSP의 이상 유무를 진단하는 기능을 가지고 있어 오류 발생시 빠른 복구를 통하여 안정된 서비스를 할 수 있다.

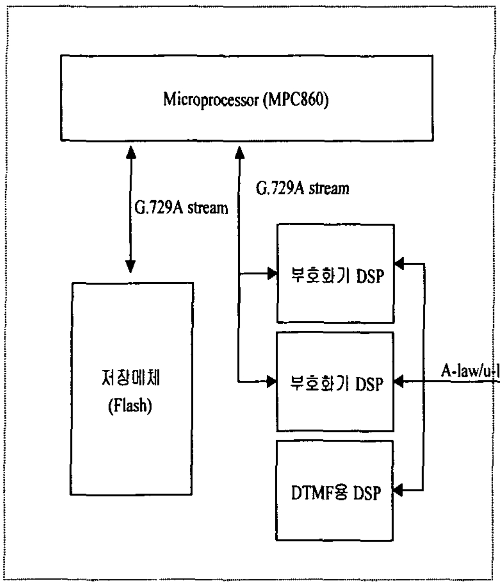


그림 4. VMS 시스템 H/W 블록도
Fig. 4. Block diagram of the VMS system H/W.

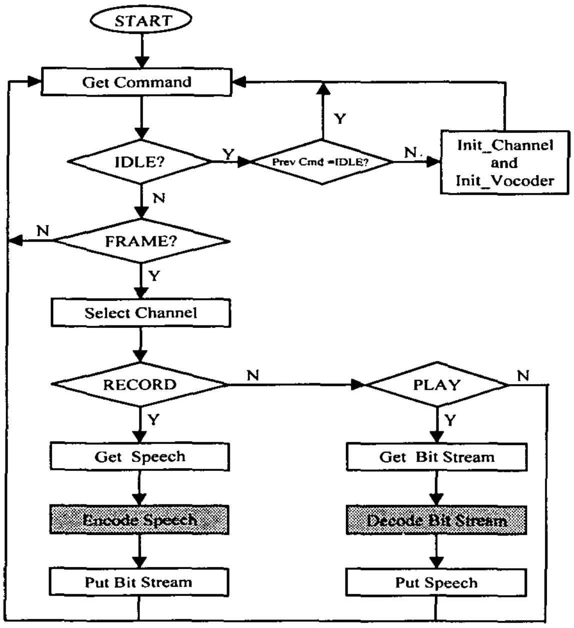


그림 5. VMS 시스템 S/W 흐름도
Fig. 5. Flowchart of the VMS system S/W.

VMS를 위한 DSP와 동작은 기본적으로 주처리기에 의해서 그림 5와 같이 제어 되고 있고 주처리기의 제어 명령 전달 방식은 폴링 방식을 사용하였고 직렬 포트로 부터 들어오는 음성 데이터는 인터럽트 방식을 사용하였다. DSP는 압축 및 복원을 하고 있지 않을 때는 항상 제어 명령 포트를 확인 하고 주처리기로부터 음성 압축 명령이 전달 되었을 경우 해당 하는 채널의 음성이 하나의 프레임을 형성할 때까지 IDLE 상태를 유지 하고 있다. 프레임이 형성되면 해당 채널의 이전 정보를 가지고 있는 메모리 블록을 선택하여 음성 압축 동작을 시작하고 압축된 데이터는 주처리기로 전달하고 주처리기는 압축된 음성을 파일 형태로 플래시 메모리에 저장한다. 음성 복원의 경우는 주처리기로부터 압축된 데이터를 전달 받아 음성으로 복원 후 직렬 포트를 사용하여 음성을 전달한다. 결과적으로 VMS 시스템에 G.729A를 적용함으로써 PCM에 비해 저장 용량이 8배 정도 향상되었다.

IV. 결 론

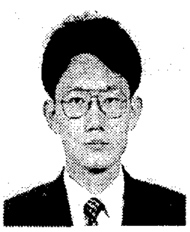
본 논문에서는 16 비트 고정 소수점 DSP를 사용하여 다채널 G.729A 음성 부호화기를 실시간 구현하였으며, 최근에 관심사가 되고 있는 멀티미디어 응용 서비스인 음성 사서함 서비스(VMS) 시스템에 적용하였다. 실시간 구현된 다채널 G.729A 음성 부호화기는 채널 당 부호화기와 복호화기에 소요된 계산량이 각각 14.5 MIPS와 3.6 MIPS였으며, 필요한 메모리 크기는 코드가 9.88K 워드, 데이터가 1.69K 워드였다. 결과적으로 VMS 시스템에는 두 개의 DSP를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실험 결과, ITU-T에서 제공된 테스트 벡터 샘플을 모두 통과하였다. 개발된 다채널 G.729A 음성 부호화기는 화상 회의 시스템이나 VOIP(Voice Over Internet Protocol)[15]에도 응용이 가능할 것으로 예상된다.

참 고 문 헌

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. Shroder, 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 300 726, *Digital Cellular Telecommunications System; Enhanced Full Rate (EFR) Speech Transcoding (GSM 06.60 Version 5.0.0)*, March, 1996.
6. J- P. Adoul, P. Mabilleau, 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. R. Salami, C. Laflamme, B. Bessette, and J- P. Adoul, "Description of ITU-T Recommendation G.729 Annex A:

Reduced Complexity 8kbit/s CS-ACELP Codec," *IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc.*, pp.775-778, 1997.
8. R. Salami, C. Laflamme, B. Bessette, and J- P. Adoul, "ITU-T G.729 Annex A: Reduced Complexity 8kbit/s CS-ACELP Codec Digital Simultaneous Voice and Data," *IEEE Communications Magazine*, pp.56-63, September, 1997.
9. ITU-T Recommendation G.729-Annex A, *Coding of Speech at 8kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP), Annex A: Reduced Complexity 8 kbit/s CS-ACELP speech Codec, November*, 1996.
10. ITU-T Recommendation H.324, *Terminal for low bit rate Multimedia Communication*. March 1996.
11. LG Information & Communications Ltd., *VIPS-HD System Manual*, August, 1997.
12. Texas Instruments Inc., *TMS320C54x CPU and Peripherals, Mnemonic Instruction Set.*, August, 1997.
13. R. Salami, C. Laflame, J-P. Adoul, A. Kataoka, S. Hayashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust, P. Kroon, and Y Shoham, "Description of the Proposed ITU-T 8 kb/s Speech Coding Standard," Proc. IEEE Workshop on Speech Coding for Telecomm., pp.3-4, 1995.
14. R. Salami, C. Laflame, J-P. Adoul, A. Kataoka, S. Hayashi, T. Moriya, C. Lamblin, D. Massaloux, S. Proust, P. Kroon, and Y Shoham, "Design and Description of CS-ACELP: A Toll Quality 8kb/s Speech Coder," IEEE Trans. on Speech and Audio Proc., Vol. 6, No. 2, pp.116-130, 1998.
15. Marcus Goncalves, *Voice over IP Networks*, McGraw-Hill, 1998.

▲ 안 도 건(Do-Keun Ahn) 1966년 1월 11일생



1993년 2월 : 동아대학교 전자공학과 (공학사)
1995년 2월 : 동아대학교 전자공학과 (공학석사)
1995년 6월~현재 : LG정보통신 디지털 네트워크연구소 주임연구원

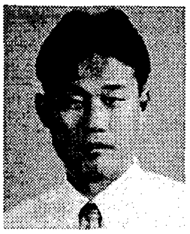
※ 주관심분야 : DSP 응용, 음성부호화, 영상신호처리

▲ 유 승 균(Seung-Kyun Ryu)

1989년 2월 : 연세대학교 전자공학과 (공학사)
1991년 2월 : 연세대학교 전자공학과 (공학석사)
1991년 2월~현재 : LG정보통신 디지털 네트워크선임 연구원
1997년 9월~현재 : 연세대학교 전자공학과(박사과정)

※ 주관심분야 : DSP 응용, 음성부호화

▲ 최 용 수(Yong-Soo Choi) 1970년 2월 25일생



1993년 2월 : 연세대학교 전자공학과 (공학사)
1995년 2월 : 연세대학교 전자공학과 (공학석사)
1999년 8월 : 연세대학교 전기컴퓨터공학과(공학박사)
1999년 9월~현재 : LG정보통신 디지털 네트워크연구소 선임연구원

※ 주관심분야 : 음성부호화, 음성신호처리, 실시간 DSP 응용

▲ 이 재 성(Jae-Sung Lee) 1969년 12월 23일생

1992년 2월 : 부산대학교 전자공학과(공학사)
1994년 2월 : 부산대학교 전자공학과(공학석사)
1994년~현재 : LG정보통신 디지털 네트워크연구소 선임연구원

※ 주관심분야 : DSP 응용, 음성인식, 음성신호처리

▲ 강 태 익(Tae-Ik Kang)

1979년 2월 : 연세대학교 전자공학과 (공학사)
1981년 2월 : 연세대학교 전자공학과 (공학석사)
1990년 1월 : Polytechnic University(Ph.D)
1981년 3월~1984년 6월 : 육군사관학교 교수요원
1990년 5월~현재 : LG정보통신 디지털 네트워크 책임연구원

※ 주관심분야 : DSP, 영상신호처리, VOIP, ATM

▲ 박 성 현(Sung-Hyun Park)

1978년 2월 : 한양대학교 전자공학과 (공학사)
1990년 2월 : 한국과학기술원 전기전자과(공학석사)
1978년 6월~1987년 6월 : 금성통신 선임연구원
1987년 7월~현재 : LG정보통신 디지털 네트워크연구소 연구위원

※ 관심분야 : 음성신호처리, VOIP, NGN