

TMS320C549 를 이용한 4 채널 G.729A 의 실시간 구현

안도건, 최용수, 윤태인, 김혜진

LG 전자 디지털네트워크연구소, 서울시 금천구 가산동 60-39

Real-time Implementation of a 4 channel G.729A Using a TMS320C549

Do-Keun Ahn, Yong-Soo Choi, Tae-In Yoon, and Hye-Jin Kim

Digital Network System Lab., LG Electronics Inc.

E-mail: [dkahn,cando,tiyoon,hjkim]@lgic.co.kr

요약

본 논문에서는 TMS320C549 를 사용하여 4 채널 G.729A 음성 부호화기를 실시간 구현하였으며, 실제로 음성 사서함 서비스 시스템에 응용하였다. 구현된 G.729A 는 채널 당 부호화기와 복호화기에 각각 14.5 MIPS 와 3.6 MIPS 를 소요하였으며, 메모리는 코드와 데이터 부분에 각각 9.88 K 워드, 1.69 K 워드를 필요로 하였다. 결과적으로 개발된 VMS 시스템에는 두 개의 DSP 를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실험 결과, ITU-T 에서 제공된 모든 테스트 벡터 결과와 비트 단위로 동일하였다.

I. 서론

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

1995 년 2 월, ITU-T(International Telecommunication Union-Telecommunication)에서는 멀티미디어 DSVD(Digital Simultaneous Voice and Data) 응용 분야에 사용하기 위한 음성 부호화기의 표준화 작업을 시작하였다. 당시의 표준 음성 부호화기들은 단일 프로세서로 모델 알고리즘과 음성 부호화 알고리즘을 통합 처리하기에는 계산량이 과도하였으므로 계산량은 10 MIPS 이하, RAM 은 2K 워드, ROM 은 8K 워드 이하의 조건들이 요구되었다. 최종적으로 1995 년 11 월에 G.729 와의 부호화 비트열의 상호 호환성이 중요하게 고려되어 USH 에서 제안한 G.729[1][2]의 계산량 감소된 버전이 ITU-T 의 V.70 시리즈(DSVD)와 G.729A(G.729 Annex A)[3-5]의 음성 부호화기 표준으로 확정되었다. G.729A 는 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 관심이 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[6])의 표준 음성 부호화기인 G.723.1[7]에

비해서 G.729A 는 낮은 복잡도와 낮은 지연 등 인터넷 전화용 음성 부호화기로서 상당한 장점을 갖고 있다.

본 논문에서는 16 비트 고정 소수점 DSP(Digital Signal Processor)인 Texas Instruments 사의 TMS320C549[8]를 사용하여 4 채널 G.729A 음성 부호화기를 실시간 구현하였으며, 실제로 음성 사서함 서비스(Voice Mailing Service: VMS)[9] 시스템에 응용하였다. TMS320C549 는 100 MIPS 의 처리 용량과 DSP 내부에 32 K 워드 RAM, 8 K 워드 DARAM(Dual Access RAM)과 24 K 워드 SARAM(Single Access RAM)을 갖는다. VMS 시스템에는 두 개의 DSP 를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실제 응용 분야로 선택한 VMS 는 수신자가 부재시 송신자가 음성을 남길 수 있는 기능을 가진 서비스로서 시스템 구성은 주 처리기, DTMF(Dual Tone Multiple Frequency) 처리용 DSP 와 부호화된 음성 신호를 저장하기 위한 플래시 메모리를 장착하고 있고 두 개의 부호화기 용 DSP 를 장착하고 있어 동시에 8 채널을 지원한다.

본 논문의 구성은 다음과 같다. II 장에서는 G.729A 의 알고리즘에 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[3-5]의 기본 구성은 G.729[1][2]와 동일하며, G.729A 로 부호화된 비트열은 G.729 로 복호화될 수 있으며, 그 반대의 경우도 가능하다. G.729A 는 DSP 를 이용한 실시간 구현에 용이하도록 G.729 에 몇가지 고속 알고리즘을 도입함으로써 그 계산량을 약 50 % 정도 감소시켰으며, 음질 면에서는 약간의 음질 저하가 발생하는 3 탠덤(tandem)과 배경 잡음 환경을 제외한 대부분의 상황에서 G.729 과 32 kbps G.726 에 상응하는 음질을 갖는다.

G.729A[7-9]에서 고속 알고리즘이 적용된 주요한 부분을 요약하면 LSP 변환, 인지 가중 필터, 개회로 피치 분석, 가중 합성 필터의 충격파 응답 계산, 적응 코드북 검색, 고정 코드북 검색 및 하모닉

포스트필터이다. 계산량 감소를 위해 사용한 방법들을 구체적으로 살펴보면 다음과 같다.

LSP 변환: G.729 에서는 0 에서 π 까지 균등하게 분할된 60 개의 지점에서 체비셰프 다항식의 값을 계산하고 부호의 변환이 있으면 그 구간을 다시 4 개로 나누어 보다 정밀하게 근을 검색한다. G.729A 에서는 계산량 감소를 위하여 다항식의 근을 검색 지점의 수를 60 에서 50 으로 줄이고, 부호 변환시에 정밀 검색 횟수를 4에서 2로 줄인다.

인지 가중: G.729 와는 달리 일반적인 형태의 가중 함수를 사용한다.

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

여기서 가중 인자 $\gamma = 0.75$ 으로 고정되며 G.729 에서 사용되는 가중 인자의 적응 과정은 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.729 에서 달리 저역 통과 필터를 거친 가중된 음성 신호를 개회로 피치 분석의 입력으로 사용한다. 이 과정은 3 단계로 구성되는데, 그 첫 단계에서는 [20, 30], [40, 79], [80, 143]의 세 구간에서 상관 값을 최대로 하는 각 구간의 지연 값을 찾는다. 마지막으로, 정규화된 구간 최대치들은 낮은 지연에 선호도를 주어서 비교하여 최종 지연 값으로 선택한다. 결과적으로 상관 값을 계산하는 과정에서 단지 짝수 샘플만을 사용하고, 세 번째 구간 [80, 143]에서는 짝수 지연에 해당하는 상관 값을 계산한 후, 선택된 짝수 지연의 ± 1 에서의 지연만을 검색함으로써 계산량을 줄인다.

적응 코드북 검색: 적응 코드북 구조는 G.729 와 동일하다. 첫 번째 부프레임에서는 [19+1/3, 84+2/3] 구간에서는 실수 피치 지연이 사용되고, [85, 143] 구간에서는 정수 피치 지연이 사용된다. 두 번째 부프레임에서는 항상 $[\text{int}(T_1)-5-2/3, \text{int}(T_1)+4+2/3]$ 구간에서 1/3 정밀도를 갖는 실수 피치 지연이 사용된다. 여기서 $\text{int}(T_1)$ 는 첫 번째 부프레임의 실수 피치 지연 T_1 의 정수 부분이다

G.729A 에서는 계산량 감소를 위해서 다음 식과 같이 상호 상관 값을 최대로 하는 지연 검색 방법을 사용한다.

$$R(k) = \sum_{n=0}^{39} x(n)y_k(n) = \sum_{n=0}^{39} x_b(n)u_k(n) \quad (2)$$

여기서 $x(n)$ 은 목적 신호, $x_b(n)$ 은 역방향 필터링된 목적 신호, $y_k(n)$ 은 지연 k 에서 필터링된 과거의 여기 신호, $u_k(n)$ 은 지연 k 에서 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.

고정 코드북 탐색: 17 비트 고정 코드북 구조도 G.729 와 동일하다. 펄스 크기는 역방향 필터링된 목적 신호의 부호를 이용하여 코드북 검색 전에 미리 정하는 G.729 에서의 SSPA(Signal-Selected-Pulse-Amplitude)[2] 방법에 의해 결정한다. 그러나 펄스의 위치는 부프레임 당 총 2^{13} 중 최대 1440(17.5 %)가지의 가능한 위치의 조합에 대해서 검색하는 'focused nested-loop' 방식을 사용한 G.729 와는 달리 펄스마다 독립적인 루프를 사용하는 'depth-first' 방식을 사용하여 검색 횟수를 최대 320(3.9 %)으로 줄였다. 이러한 고정 코드북 검색의 간략화는 약 0.2 dB SNR 감소에 해당하는 음질 저하를 가져오지만 부호화기 전체 계산량 감소의 50%를 차지한다.

후처리 과정: 적응 포스트 필터는 장구간 포스트필터, 단구간 포스트필터, 경사 보상 필터의 세 필터의 직렬 연결로 구성된다. G.729 와의 주요한 차이는 하모닉 포스트필터링 과정이다. 하모닉 포스트필터의 장구간 피치 지연은 전송된 현재 부프레임에서 피치 지연의 정수 부분을 중심으로 ± 3 범위의 정수 값만을 검색하여 G.729 에서 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 DSP 는 TI(Texas Instruments)의 TMS320C549[8]로서 100 MIPS 의 처리 용량을 갖고 있으며 DSP 내부에 32 K 워드의 RAM 을 가지고 있고 8K 의 DARAM(Dual Access RAM)과 24 K 의 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. TMS320C54x 의 경우 프로그램 영역과 데이터 영역의 메모리 맵이 각각 존재하고 내부 RAM 의 경우 기본적으로 데이터 영역으로 지정되어 있으나 메모리 맵 모드를 변경하여 프로그램과 데이터의 영역을 동시에 내부 RAM 으로 지정할 수 있다. 본 논문에서도 실시간 구현 시스템의 처리 속도의 향상을 위하여 DSP 의 내부 RAM 만을 사용하였다.

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

3.1 알고리즘 및 C 소스 분석

부호화기의 효율적인 구현을 위하여 알고리즘을 분석하고 C 소스를 검토하여 각 기능 별로 분류하였고 각 모듈의 복잡도와 점유율을 조사하였다. G.729 는 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 1 은 G.729A 부호화기의 기능별 처리 부하(processing load)를 나타낸다.

표 1. G.729A 부호화기의 기능별 처리 부하.

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 조사하였다.

3.2 한 채널 실시간 구현 및 검증

다채널 음성 부호화기 구현에 앞서 한 채널 부호화기를 C 언어와 어셈블리어를 사용하여 구현하였으며 ITU-T 에서 제공되는 테스트 벡터를 적용하여 검증하였다. 본 논문에서 사용하고 있는 DSP 의 경우 프로그램 영역과 데이터 영역이 DSP 내부 RAM 에 존재해야만 최적의 성능을 발휘할 수 있다. 처리 속도 증가를 위한 또 다른 방법은 Dual-operand 를 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 DSP 의 처리 능력을 최적화하였다. 부호화기 검증용으로 사용한 테스트 벡터는 총 9 개로서 부호화기와 복호화기용으로 6 개의 벡터와 복호화기에만 사용하는 3 개의 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 선형 PCM 으로 변환 후 부호화기에 적용하였다. 표 2 는 본 논문에서 구현한 1-채널 G.729A 의 복잡도 및 사용한 메모리 크기이다.

표 2. G.729A 의 MIPS 및 메모리(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.729A 의 경우 10 ms 프레임 단위로 부호화 하고 한 프레임 시간 내에서 시분할하여 다채널 구현이 가능하다. 따라서

부호화기의 성능에 의해서 구현 가능한 채널의 수가 결정 되어지고 최대 처리 시간이 2.2 ms 인 경우의 예로서 10 ms 내에 4 채널의 부호화기를 구현할 수 있다. 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 1 은 본 논문에서 사용한 데이터 영역 메모리 구성도이다.

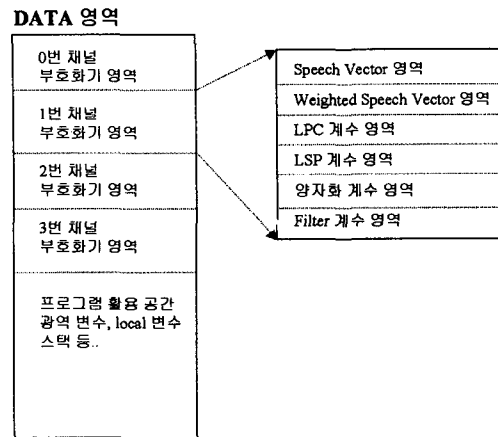


그림 1. DSP 데이터 영역 메모리 구성.

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

VMS 는 수신자가 부재시 송신자가 음성을 남길 수 있는 기능을 가진 음성 사서함 서비스로서 VMS 시스템의 구성은 주 처리기, DTMF(Dual Tone Multiple Frequency) 처리용 DSP 와 부호화된 음성 신호를 저장하기 위한 플래시 메모리를 장착하고 있고 부호화기 용 DSP 가 두 개 장착되어 있어 동시에 8 채널을 지원한다. 그림 2 는 VMS 시스템 구성도이다. 그림 2 에서 주처리기는 전달 받은 DTMF 를 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 DSP 의 이상 유무를 진단하는 기능을 가지고 있어 오류 발생시 빠른 복구를 통하여 안정된 서비스를 할 수 있다.

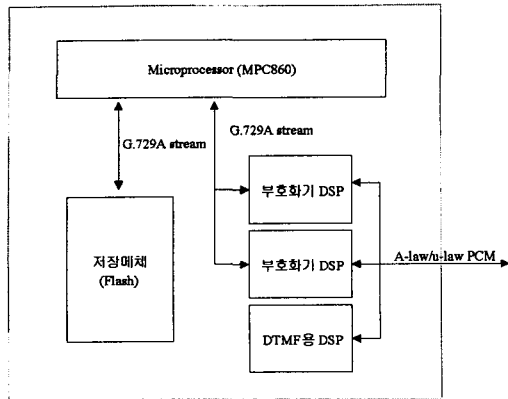


그림 2. VMS 시스템 H/W 블록도.

VMS 를 위한 DSP 의 동작은 기본적으로 주처리기에 의해서 그림 3 과 같이 제어 되고 있고 주처리기의 제어 명령 전달 방식은 폴링 방식을 사용하였고 직렬 포트로부터 들어오는 음성 데이터는 인터럽트 방식을 사용하였다. DSP 는 압축 및 복원을 하고 있지 않을 때는 항상 제어 명령 포트를 확인하고 주처리기로부터 음성 압축 명령을 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.

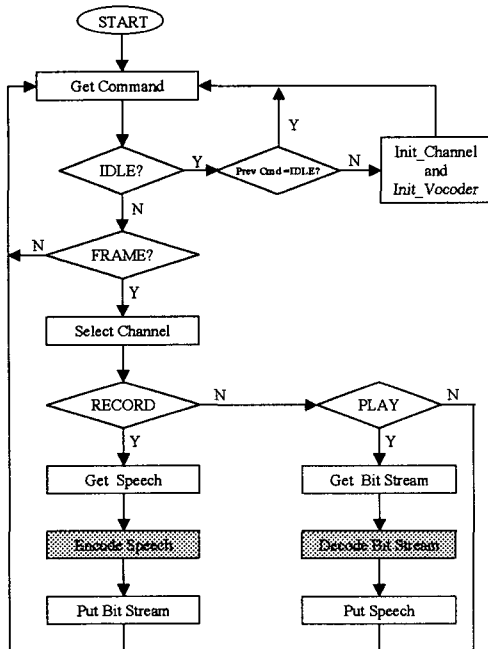


그림 3. VMS 시스템 S/W 흐름도.

IV. 결론

본 논문에서는 16 비트 고정 소수점 DSP 인 TI 사의 TMS320C549 를 이용하여 4 채널 G.729A 음성 부호화기를 실시간 구현하였으며, 최근에 관심이 되고 있는 멀티미디어 응용 서비스인 음성 사서함 서비스 시스템에 적용하였다. 실시간 구현된 다채널 G.729A 음성 부호화기는 채널 당 부호화기와 복호화기에 소요된 계산량이 각각 14.5 MIPS 와 3.6 MIPS 였으며, 필요한 메모리 크기는 코드가 9.88 K 워드, 데이터가 1.69 K 워드였다. 결과적으로 VMS 시스템에는 두 개의 DSP 를 사용하여 DSP 당 4 채널씩 총 8 채널을 수용하였다. 실험 결과, ITU-T 에서 제공된 테스트 벡터 샘플을 모두 통과하였다. 개발된 다채널 G.729A 음성 부호화기는 화상 회의 시스템이나 VOIP(Voice Over Internet Protocol)[10]에 응용이 가능하다.

참고문헌

- [1] ITU-T Recommendation G.729, *Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP)*, March 1996.
- [2] R. Salami, C. Laflamme, 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 8 kb/s Speech Coder," *IEEE Trans. on Speech and Audio Proc.*, Vol. 6, No. 2, pp.116-130, 1998.
- [3] ITU-T Recommendation G.729-Annex A, *Coding of Speech at 8 kbit/s Using Conjugate-Structure Algebraic-Code-Excited Linear-Prediction (CS-ACELP), Annex A: Reduced Complexity 8 kbit/s CS-ACELP speech Codec*, November 1996.
- [4] R. Salami, C. Laflamme, B. Bessette, and J- P. Adoul, "Description of ITU-T Recommendation G.729 Annex A: Reduced Complexity 8 kbit/s CS-ACELP Codec," *IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc.*, pp.775-778, 1997.
- [5] R. Salami, C. Laflamme, B. Bessette, and J- P. Adoul, "ITU-T G.729 Annex A: Reduced Complexity 8 kbit/s CS-ACELP Codec Digital Simultaneous Voice and Data," *IEEE Communications Magazine*, pp.56-63, September 1997.
- [6] ITU-T Recommendation H.324, *Terminal for low bit rate Multimedia Communication*. March 1996.
- [7] 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.
- [8] Texas Instruments Inc., *TMS320C54x CPU and Peripherals, Mnemonic Instruction Set.*, August 1997.
- [9] LG Information & Communications Ltd., *VIPS-HD System Manual*, August 1997.
- [10] Marcus Goncalves, *Voice over IP Networks*, McGraw-Hill, 1998.