主題

대화형 멀티미디어 서비스를 위한 기술 : 데이터 부호화 기술

한국전자통신연구원 김상범, 최장식, 김기현, 민병기

차 례

1. 서 론
2. 비디오 압축/복원 표준
3. 오디오 압축/복원 표준 G.723.1
4. H.263 VLSI 구현
5. 영상전화 응용시스템의 구조 설계 및 구현
6. 결 론

1. 서 론

멀티미디어 서비스에 대한 정의를 "상호 교환되는 정보가 한 가지 이상의 정보 타입, 예를 들어 영상, 데이터, 음성, 그래픽을 갖는 서비스"로 규정하고 있으며, 이는 텍스트, 음성, 그래픽, 비디오, 오디오 등과 같은 다수의 미디어 중에서 두 가지 이상의 미디어로 구성되는 복합적인 서비스를 말한다. 멀티미디어 응용 서비스에서 대화형 멀티미디어 서비스는 사용자 사이의 양방향의 실시간 통신을 제공하는 서비스를 말하는 것으로 멀티미디어 회의, 비디오 폰, 비디오 회의, 비디오 감시, 원격 학습 등의 응용분야가 있다.

이들 서비스들은 정보통신 기술의 발전으로 네트워크를 통한 분산 환경에서 제공되게 되었고, 서비스의 질도 고품질의 영상 및 음성 서비스가 제공되고 있다. 고속망에서의 멀티미디어 회의 서비스는 망에 접근이 가능한 특정 사용자들만이 서비스를 이용할 수 있는 제한으로 인하여 일반인에게 서비스를 제공하기 위하여서는 전화망을 이용한 서비스의 제공이 가장 유용하다. 이처럼 기존의 공중 전화 회선을 사용하여 상대방의 얼굴을 보면서 통화할 수 있는 전화가 영상전화, 비디오폰이다.

영상전화는 최근에 와서는 표준화 작업의 완료와 인터넷의 확산 및 컴퓨터, 특히 PC의 성능 향상으로 인해 다양한 종류가 등장하였으며, 상용 시장이 확대되고 있다. 영상전화에 관련된 국제표준의 종류는 64Kbps 이상의 협대역 종합정보통신망(ISDN) 기반의 H.320과 28.8Kbps의 일반 전화선에서의 H.324, 그리고 인터넷 및 LAN까지 확장된 H.323[2] 표준 규격으로 나누어 진다. 이들 규격내에는 〈표1〉에서 처럼 각각의 비디오/오디오 코덱, 멀티플렉싱, 전송 프로토콜 방식등 다양한 세부 규격을 포함하고 있으며 이들 각 규격들은 상호 연결된다[8].

〈표1〉 영상전화 및 영상회의 관련 핵심 표준

표준 규격 구분			설 명
H.320	H.324	H.323	영상전화 응용시스템 표준
H.261	H.263, 261	H.263, 261	비디오 압축/복원 표준
G.711, 722, 728	G.723.1	G.711, 722, 728, G.723.1, 729	오디오 코덱 표준
H.221	H.223	H.225	오디오/비디오 멀티플렉싱
H.242	H.245	H.245	제어 프로토콜
I400	V.34	LAN, Internet	네트워크 인터페이스

이 세부 규격 중에서 특히 비디오, 오디오 코덱이 바로 핵심 기술이며, 시스템 성능의 90% 이상을 차지한다. 이들에 대한 세부 내용은 다음 장에서 설명하고 있다.

ITU의 전신 CCITT는 ISDN을 이용한 영상전화 H.320[5]을 구상하였고 이 영상전화기의 비디오 압축 알고리즘으로 H.261표준을 제안하였고, ISDN의 기본 인터페이스인 64Kbps 기본 채널 대역폭에서 서비스되는 이유로 흔히 H.261을 Px64(P=1..30)이라 부르기도 한다. H.261에는 영상회의 응용을 지원하기 위한 특별한 기능으로, 전송 채널에서 발생하는 오류를 발견하고 수정하는 순방향 오류 수정 코드가 포함되어 있고 화면을 분할하는 등 회의에 필요한 기능과 도큐먼트 카메라를 사용 중이라는 표시 기능 등을 가지고 있다. 그러나 ISDN의 보급이 지연되어 영상전화 단말이 상용화되지 못하자 이에 대한 대안으로 채널 대역폭 64Kbps 이내에서도 사용이 가능한 H.324 표준 규격[4]을 제안하였고, 이 H.324표준의 비디오 압축/복원 표준으로 H.263을 채택하였다.

영상전화는 영상 데이터 뿐만 아니라 음성이나 오디오 정보를 제한된 전송 용량에 많은 양을 제공하기 위해 디지털 오디오 신호의 압축 기술을 필요로 한다.

ITU내 SG15의 음성전문가 그룹은 여러 표준 가운데에서 저속 영상전화의 오디오 표준 규격으로서 G.723.1을 채택하였다. 〈표2〉에 오디오 코덱 표준의 종류와 압축 기술를 요약하였고, 이 외에도 MPEG, GSM, VSELP, QCELP, MELP, MBE등의 기법이 있다

〈표2〉 오디오 코덱 표준의 종류

표 준	압축기술	설 명
G.711	PCM(Pulse Code Modulation)	48, 56, 64Kbps(3.1KHz)
G.722	SB-ADPCM(Sub-band ADPCM)	48, 56, 64Kbps(7KHz)
G.723.1	ACELP, MP-MLQ(Multipulse Maximum Likelihood Quantisation)	5.3, 6.3Kbps
G.726	ADPCM(Adaptive Differential Pulse Coded Modulation)	32Kbps
G.728	LD-CELP(Low Delay Code Excited Linear Prediction)	16Kbps(3.1KHz)
G.729	CS-ACELP(Conjugate-Structure Algebraic Code Excited Linear Prediction)	8Kbps

H.323, H.324 시스템의 구현은 펜티엄 이상급의 프로세서의 출현에 따라 일반 PC상에서 소프트웨어만으로도 충분히 그 요구 성능을 만족하고 있다. 이론적으로 모션 추정에 필요한 계산량은 대략 1.1GIPS정도(QCIF, 30 프레임)이지만 수정된 알고리즘으로 이러한 문제를 해결하고 있다. 그러나 일반 PC가 아닌 일반 전화기나 무선 핸드폰 등과 같은 고가의 프로세서를 필요로 하지 않는 시스템에서는 별도의 VLSI 기술에 의한 전용 코덱 칩의 구현이 필수적으로 요구되고 있는 실정이다. 본 논문에서는 ETRI가 개발한 H.263비디오 코덱의 세부 동작 원리와 구조, 구현을 분석하고 H.263 프로세서에 기반한 H.323 영상전화 응용시스템에 대해 언급하고자 한다. 2장에서는 비디오 압축/복원 표준인 H.261 및 H.263의 차이 및 특징에 대해 분석하고, 3장에서는 오디오 압축/복원 표준인 G.723.1에 대해 간단히 서술하였다. 제4장에서는 ETRI에서 개발한 H.263의 특징, 구조 및 구현 과정에 대해, 5장에서는 H.263의 FPGA 구현을 통한 H.323 기반의 영상전화 응용시스템의 구현 과정을, 6장에서 결론을 맺는다.

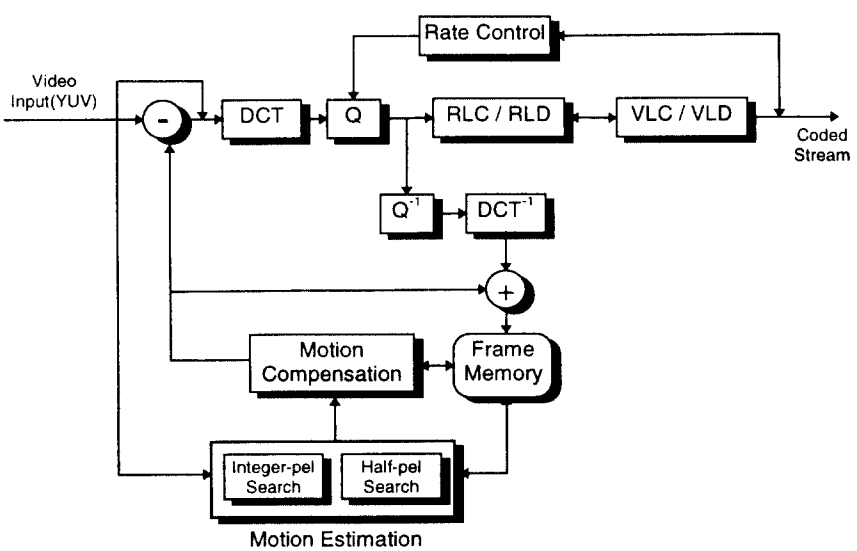
2. 비디오 압축/복원 표준

H.263은 기존의 ISDN 상에서 비디오 코덱인 H.261에 기반을 두고 새롭게 정의되었다. 초기는 64Kbps이하의 낮은 전송율을 목적으로 했으나 지금은 그 제약이 없어진 상태이고, 다양한 범위의 비트율에 모두 적용이 가능하다. H.263 비디오 코덱[3]의 일반적인 기능과 흐름은 〈그림1〉에 나타내었다. 대부분의 기능들은 기존의 MPEG이나 H.261의 기능과 거의 동일하다. 그러나 각 기능별로 요구되는 성능 특성이 차이가 나고, 그 차이를 요약하면 〈표3〉과 같다.

H.261[1]에서는 정수 픽셀 정밀도와 루프 필터로 모션 보상을 수행하나, H.263은 부동소수 정밀도에서 모션 보상을 한다. H.263은 H.261에는 없는 4가지 고급 모드를 지원하는데 물론 이는 선택사양이다. 처리 화면 크기도 H.261에서는 QCIF, CIF였던 반면, H.263은 SQCIF, QCIF, CIF, 4CIF, 16CIF의 5가지로 확대되었다. 확대된 이유는 MPEG표준과 같은 높은 비트율과 경쟁하기 위해서다.

H.263의 고급 선택 사양으로 무제한 모션 벡터(unrestricted motion vector), 문법 기반 연산 코

〈그림1〉 비디오 코덱의 일반적인 압축/복원 기능과 흐름



딩(syntax based arithmetic coding), advanced 예측 모드(prediction mode) 및 PB 프레임 기능이 추가로 제공된다.

무제한 모션 벡터 기법은 화면의 외곽에도 모션벡터를 허용하여 작은 화면에서 화면의 경계면을 따른 움직임이 있는 경우 효율적이며, 또한 모션벡터의 범위를 확장하여 카메라의 움직임이 큰 경우에도 효과적이다.

Advanced 예측모드 기법은 중복된 블록 모션 보상 방법으로 P프레임을 위해 사용된다. 한 개의 16x16 벡터 대신 4개의 8x8 벡터가 사용되고, 이때 모션 벡터는 무제한 모션 벡터기법을 적용 받는다. 4

〈표3〉 H.261과 H.263의 비교 분석

	H.261	H.263
전송율	Px64Kbps(P=1,..,30)	64Kbps 이하
지원해상도	Default : QCIF, Optional : CIF	Default : SQCIF, QCIF Optional : CIF, 4CIF, 16CIF
ME	탐색범위 : Max(+15,-15) 탐색단위 : 16x16 정확도 : integer pixel 유니트 : 1MV/MB	탐색범위 : Max(+15.5,-16) 탐색단위 : 16x16, 8x8 정확도 : integer pixel, half pixel 유니트 : 1MV/MB, 4MV/MB
Loop Filter	O	X
DCT	8x8	H.261 과 같음

Quantization	Intra :
	DC : QC(0,0)=C(0,0)//8
	AC : QC(i,j)=C(i,j)//2QP
	Inter :
	If QP=odd QC(i,j)//QP
	If QP=even
	QC(i,j)=[C(i,j)+sign(C(i,j))] //2QP
	QP : depend on buffer fullness
	QP = [1..31]
	Intra MB : uniform quantizer without deadzone
	Inter MB : uniform quantizer with deadzone
		Intra :
		DC : QC(0,0)=C(0,0)//8
		AC : QC(i,j)=sign(C(i,j))x(IC(i,j)I/2QP)
		Inter :
		QC(i,j)=
		sign(C(i,j))x((IC(i,j)I-QP/2)/2QP)
		QP = [1..31]

Invserse Quantization	If QP=odd, for QC(i,j)>0 REC(i,j)=[QP(2QC(i,j)+1)] If QP=odd, for QC(i,j)<0 REC(i,j)=[QP(2QC(i,j)-1)] If QP=even, for QC(i,j)>0 REC(i,j)=[QP(2QC(i,j)+1)-1] If QP=even, for QC(i,j)<0 REC(i,j)=[QP(2QC(i,j)-1)+1] QP=[1..31] REC(i,j)=0 for QC(i,j)=0 Intra DC : REC(0,0)=QC(0,0)*8	REC(I,j)=0 for QC(i,j)=0 If not (REC(i,j)=0), QP=odd REC(i,j)= sign(QC(i,j))x(2QPxI QC(i,j)I+QP) If not (REC(i,j)=0), QP=even REC(i,j)= sign(QC(i,j))x(2QPxI QC(i,j)I+QP-1)
Entropy Coding	DCT coeff : VLC + FLC Intra coded DC : FLC(8bit)	DCT coeff : VLC+ FLC Intra coded DC : FLC(8bit)
	Intra Coded & Inter coded AC : 2-D(Run, Level) huffman coding MBA, MTYPE, MVD, CBP : VLC(1-D huffman code)	Intra Coded & Inter coded AC : 3-D(Last, Run, Level) huffman coding MCBPC, MODB, CBPY, MVD : VLC(1-D huffman code)
고급옵션	X	unrestricted motion vector syntax based arithmetic coding advanced prediction mode PB frame

개의 벡터를 사용하면 비트량은 많아지나 훨씬 좋은 예측 성능을 보인다. 이를 통해 블록킹 단점을 보완하여 주관적 화질을 개선한다.

문법 기반 연산 코딩 기법은 VLC 코딩 대신 산술 부호화를 사용하는 것으로 SNR이나 복원 영상은 같으나 부호화율이 interframe에서는 3~4%, intraframe에서는 10%의 높아진다.

PB 프레임 기법은 P화면, B화면의 두 화면을 하나의 단위로 부호화하는 방법으로 상대적으로 비트율의 증가없이 프레임율을 증가시킨다. 이러한 추가 선택 사양을 통해 동일한 양자화 값에서 H.261보다 비트율은 절반 이하로, SNR은 2~3dB 정도 향상 시킬 수 있다[7].

3. 오디오 압축/복원 표준 G.723.1

오디오 압축/복원 표준 G.723.1은 미국의 DSP 그룹의 True Speech를 개선시킨 것을 확정하였다(1995년). 이 기술은 DSP 그룹 외에도 이스라엘의 AudioCode, 부호화기술인 ACELP를 개발한 캐나다의 셔블록 대학이 공동으로 개발한 것으로 6.3Kbps, 5.3Kbps의 전송 속도로 고음질의 음성 데이터를 압축 전송할 수 있으며, 28.8Kbps 모뎀에서 20%만의 대역으로 충분하다. 6.3Kbps의 경우 MP-MLQ, 5.3Kbps에는 ACELP를 채택하고 있다.

G.723.1[9,10]은 프레임 단위로 비트율의 변경이 가능하고, 옵션으로 가변 비트율도 지원할 수 있다. 프레임 단위의 처리에 걸리는 시간은 37.5msec로 프레임 순수 처리에 30ms, 추가로 발생하는 지연에 7.5ms가 소요된다. 여기서 지연은 구현 지연, 전송 지연, 멀티플렉싱을 위한 버퍼 지연을 합한 값이다. 이외의 파라미터인 복잡도 및 비트율과 관련하여 G.723.1과 G.729를 비교 분석[6]하면 〈표4〉와 같다.

4. H.263 VLSI 구현

4.1 내부 구조

ETRI의 저 전송율 비디오 코덱 프로세서는 H.324(Video Teleconferencing Standards) 표준 규격에서 정의된 H.263 비디오 압축/복원 규격을 만족한다. 내부 압축/복원 모듈로는 ME, MC, DCT/IDCT, Q/IQ, RLC/D가 있으며 고유 구조의 RISC프로세서를 내장하여 칩 내부의 기능 블록 제어 및 소프트웨어 다운로드를 수행한다. 외부의 프레임 메모리(512KBx16)와 직접 인터페이스되며 GPIO포트를 지원하여 키패드 연결 및 내부 디버그를 지원한다. 비디오 입출력은 CCIR656 포맷을 통해 이루어진다.

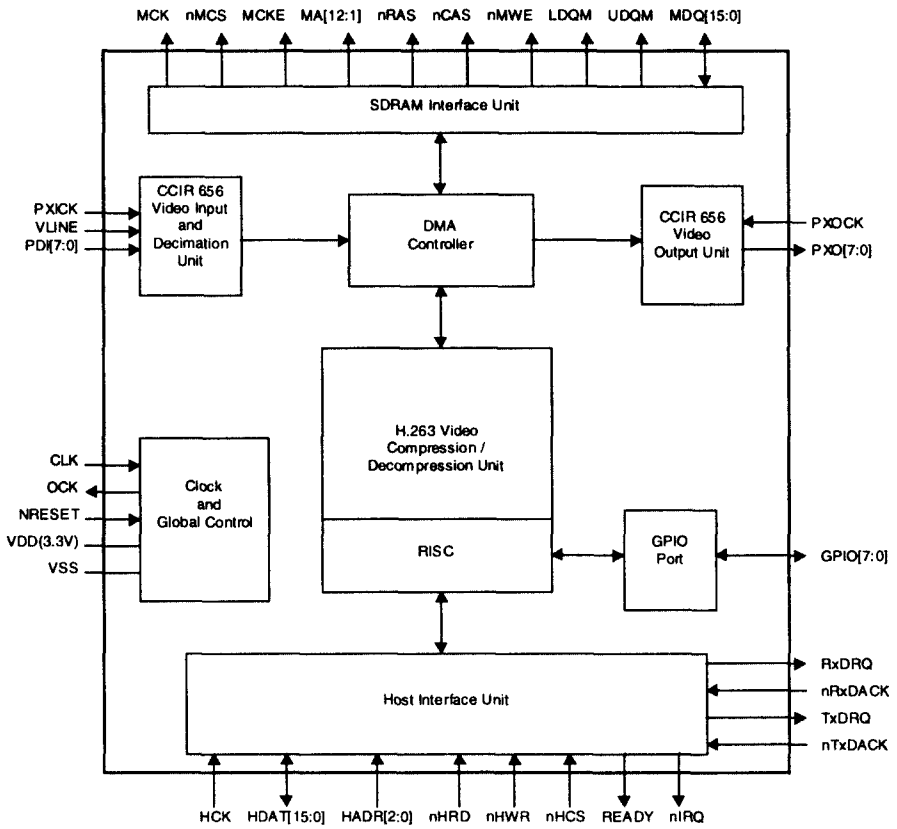
H.263의 압축/복원 기능을 구현할 ASIC칩인 비디오 코덱 프로세서의 기능 블록도는 〈그림2〉와 같다.

〈표4〉 G.723.1과 G.729의 성능비교

		Speech-coder standard			
attribute	parameter	G.729	G.729.A	G.723.1	G.723.1
Bit rate	Kbps	8	8	6.3	5.3
Delay	Frame(msec)	10	10	30	30
	Subframe(msec)	5	5	7.5	7.5
	Algorithmic delay(msec)	15	15	37.5	37.5
complexity	MIPS	22	12	18	23
	RAM(16-bit)	2.7K	2K	2.2K	2.2K

MIPS : estimated for general-purpose, fixed point DSP

〈그림2〉 비디오 코덱 프로세서 블럭도



비디오 코덱 프로세서는 비디오 디바이스로부터 데이타를 입력 받아 변환하는 픽셀 처리 모듈과 SDRAM에 있는 비디오 데이터를 칩 밖으로 바로 출력하는 픽셀 출력 모듈, H.263 압축/복원 모듈과 프레임 제어와 메모리 관리를 위한 DMA제어기로 구성된다.

움직임 예측 및 보상 모듈, DCT/IDCT변환 및 (역)양자화기, 지그재그 변환 및 RLC/D 기능을 포함하는 H.263 압축/복원 모듈의 블록도를 〈그림3〉에 나타내었다.

H.263 코덱을 구성하는 기능 블럭의 주요 기능은 다음과 같다.

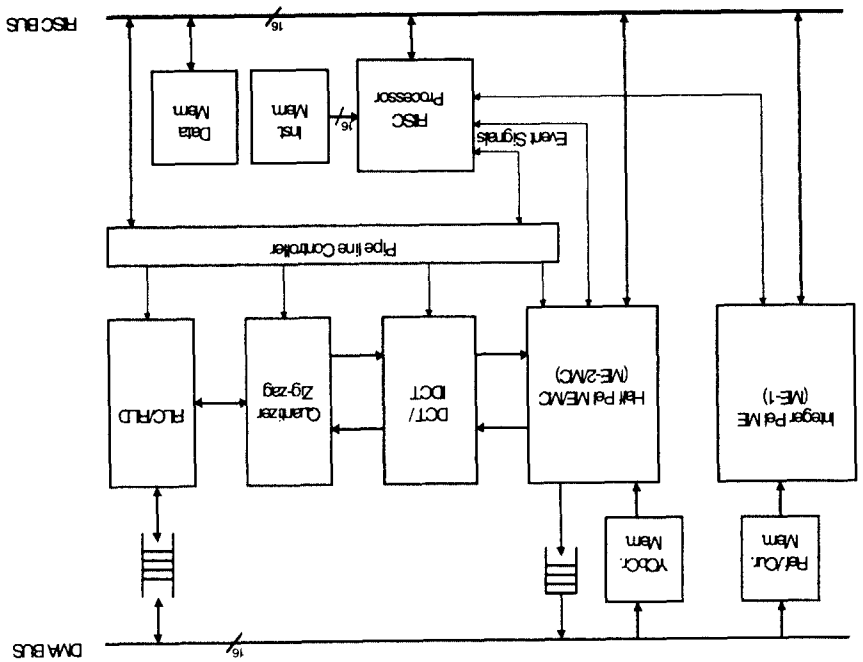
• 픽셀 프로세싱 유니트

카메라의 색신호(RGB)를 휘도와 두개의 색차 신호로 표현되는 Y/Cr/Cb 형식으로 CCIR656 표준에 따른 상호 변환 기능과 영상 형식 CIF/QCIF 상호 변환 기능, 영상 압축/복원 처리를 위한 영상 프레임 저장 기능을 포함 한다.

• 움직임 예측/보상 유니트

ME-1(Motion Estimation-1)/ME-2, MC (Motion Compensation) 블럭으로 구성된다. 동영상 전송시 시간적 중복성(Temporal redundancy)을 제거 하거나 보상하여 압축 효과를 높힌다. ME-1 블럭은 정수 화소 단위의 움직임 추정과 매크로 블럭의 모드(Intra, Inter) 결정 기능을 수행한다. 움직임 추정을 위한 탐색 영역은 (15, (7, 탐색 단위는 16x16, 8x8 단위의 현재와 이전 프레임간 전역탐색(Full Search)를 수행 한다. Inter 모드인 경우,

〈그림3〉 H.263 압축/복원 모듈



ME-2 블럭은 ME-1에서 구한 정수 화소 단위의 움직임 벡터를 반 화소 단위로 정밀 탐색 기능과 움직임 보상을 위한 마크로 블럭 단위의 현재 프레임의 영상을 움직임 벡터를 이용 재구성한다.

• 변환 유니트

DCT/IDCT, 양자화/역양자화기, RLC/D 블럭으로 구성된다. DCT 블럭은 영상신호를 압축/복원하기 위한 영상신호를 공간 영역에서 주파수 영역으로 변환 시키는 역할을 한다. 압축 모드 동작시 정방향과 역방향 DCT 동작을 동시에 수행하며, 복원 모드 동작 시 역방향 DCT로 동작된다. 연산부는 고속 알고리즘, 분산산술 방식을 혼합 설계하여 구현하였으며, 동작 주파수는 16MHz. 이다. 양자화/역양자화기는 정방향 DCT 출력 12-bit 데이터 제산 기능과 역방향 DCT을 위한 승산 기능을 수행 한다

• RISC 유니트

내부 기능 블럭의 초기화 및 전반적인 스케쥴링 기능, 에러 진단 기능 등을 담당한다. RISC 프로세서는 인스트럭션 버스와 데이타 버스가 분리된 하버드 버스 구조를 갖고, 4단계 파이프라인으로 구성되어 있다.

• DMA 제어 유니트

비디오 프로세서의 각 기능 블럭과 외부 프레임 메모리 사이의 데이타 전송 기능을 수행한다. 16 개의 전송 채널을 사용하고 있으며, 채널 중재 방법은 고정 및 순환 우선권 방식을 모두 지원한다.

4.2 H.263 제어 및 디바이스 드라이버

H.263 비디오 코덱 프로세서를 구동하기 위해 하드웨어 플랫폼의 마더 보드 위에 확장보드 형태로 장착되며, 호스트 프로세서에 의해 모든 동작이 제어된다.

압축 및 복원 모드에서 출력값과 입력값으로 사용되는 압축 코드(Run/Amplitude Code와 블록 헤더

정보)는 호스트 버스를 통하여 전송된다. 호스트 프로세서는 압축 모드에서 압축된 코드를 전달 받아서 최종 VLC 코드를 생성한다. 반대로 복원 모드에서 호스트 프로세서는 H.263 VLC 코드를 해석하여 블록 헤더 정보와 Run Amplitude 코드를 비디오 코덱 프로세서에 전달한다. 압축과 복원 모드는 프레임 단위로 바뀔 수 있다. 이를 지원하기 위해 RISC OS 환경에서의 디바이스 드라이버 구현이 필요하다.

이 드라이버의 역할은 크게 인터럽트 처리 기능, DMA를 통한 비디오 스트림의 실시간 전달, 포트에 의한 프로세서 내부 레지스터의 접근으로 분류된다. 또한 별도의 응용 프로그램으로 VLC/D 모듈과 입출력 버퍼의 상태 및 압축 비트율에 따라 양자화 값, 프레임의 생략 여부를 결정하는 버퍼 및 율 제어 모듈 및 영상전화를 위한 응용 소프트웨어 모듈로 구성된다.

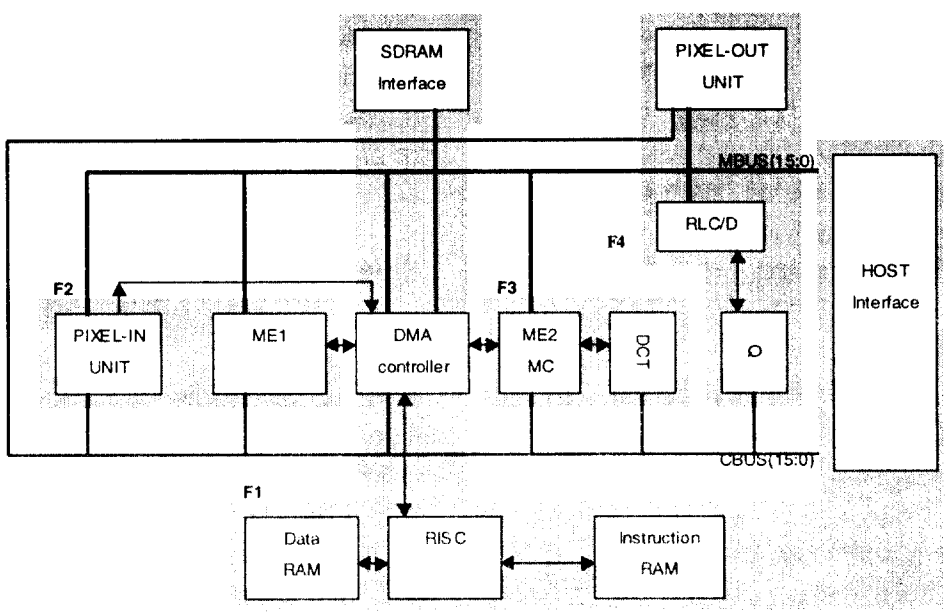
4.3 FPGA 구현

H.263 비디오 코덱의 각 기능 블록 회로들을 각 FPGA에 프로그래밍하여 하드웨어 플랫폼 상에서 그 동작을 검증하기 위한 CodecBD는 4개의 EPF10K100ARC240-1 FPGA(Altera사)칩으로 구성되고, 이 4개의 FPGA칩을 프로그래밍하기 위한 8개의 EPROM이 사용된다.

4개의 FPGA칩에 대해 살펴보면, 〈그림4〉에서 F1 FPGA는 DMA제어기(DMAC), 호스트 인터페이스, RISC 기능이 구현되어 있고, F2 FPGA는 픽셀의 입력 처리와 ME1(Motion Estimation) 기능이 구현되었다. F3 FPGA는 DCT(Discrete Cosine Transform), ME2, MC(Motion Compensation)로 구성되고, F4 FPGA는 RLC(Run Length Coding), ZZ(Zig-Zag Scan), Q(Quantizer)로 이루어져 있다.

FPGA칩 외에 SDRAM(512KB x 16)이 사용되어 내부 버스를 통하여 압축 비디오 데이터, 복원된 비디오 데이터 및 기능블럭에서 수행된 결과들을 보관한다. 이때 DMA제어기의 제어에 따라 외부 SDRAM에서 해당 기능 블록으로 또는 기능 블록에서 SDRAM으로 데이터가 이동한다.

〈그림4〉 FPGA칩간 연결 및 구성도



• F1 FPGA

호스트 인터페이스, DMA제어기, SDRAM 제어기, RISC 프로세서, Address Decoder, 클럭 분주기로 구성되어 제어 및 인터페이스 기능을 제공한다.

클럭 분주 회로는 54MHz의 입력 클럭을 받아서 27MHz, 13.5MHz의 클럭을 출력하고 시스템 클럭으로 27MHz를 사용한다. 그리고 고유 구조의 RISC 프로세서를 내장함으로 칩 내부의 기능 블록 제어 및 소프트웨어 다운로드를 용이하게 하였다. DMA제어기는 비디오 프로세서의 각 기능 블록과 외부 프레임 메모리 사이의 데이터 전송 기능을 수행한다.

• F2 FPGA

ME1과 픽셀 입력처리를 담당한다. 픽셀 입력 처리는 비디오 압축을 위한 첫번째 입력단계로 CCIR656 포맷의 비디오 코드를 입력 받아 F1 FPGA칩의 DMA제어기와 연결 동작하여 외부 프레임 메모리에 저장한다. ME1는 계산된 움직임 벡터와 해당 SAD 값을 F1 FPGA칩의 RISC로 전달한다.

• F3 FPGA

ME2, MC, DCT, Transform Module Controller로 구성되어 반화소 모션 벡터를 생성하고 transform module의 전처리부와 제어기능, 움직임 보상 기능을 가지며, DCT 처리를 수행한다. ME2 블록은 Inter모드에서 ME1에서 구한 정수 화소 단위의 움직임 벡터를 반 화소 단위로 정밀 탐색 기능과 움직임 보상을 위한 매크로 블록 단위의 현재 프레임의 영상을 움직임 벡터를 이용 재구성한다.

• F4 FPGA

양자화기, 역양자화기, 지그재그 변환, RLC/D, 픽셀 출력 유니트로 구성된다. 픽셀 출력은 CCIR656 포맷으로 이루어진다.

• FPGA칩의 하드웨어 재원

FPGA10K칩의 재원은 로직셀이 10만게이트, 내부메모리는 24,576비트이다. 〈표5〉는 F1, F2, F3, F4의 4개 FLEX10K 칩에서 사용하는 외부 인터페이스 신호의 수 및 내부에서 사용하는 하드웨어 량을 수치로 나타내고 있다.

〈표5〉 F1,F2,F3,F4의 사용 하드웨어 재원

	F1	F2	F3	F4
입력핀수	30	33	43	52
출력핀수	51	12	38	30
양방향핀수	82	32	32	32
내부메모리사용비율(%)	95	68	58	52
로직셀사용비율(%)	78	76	88	60

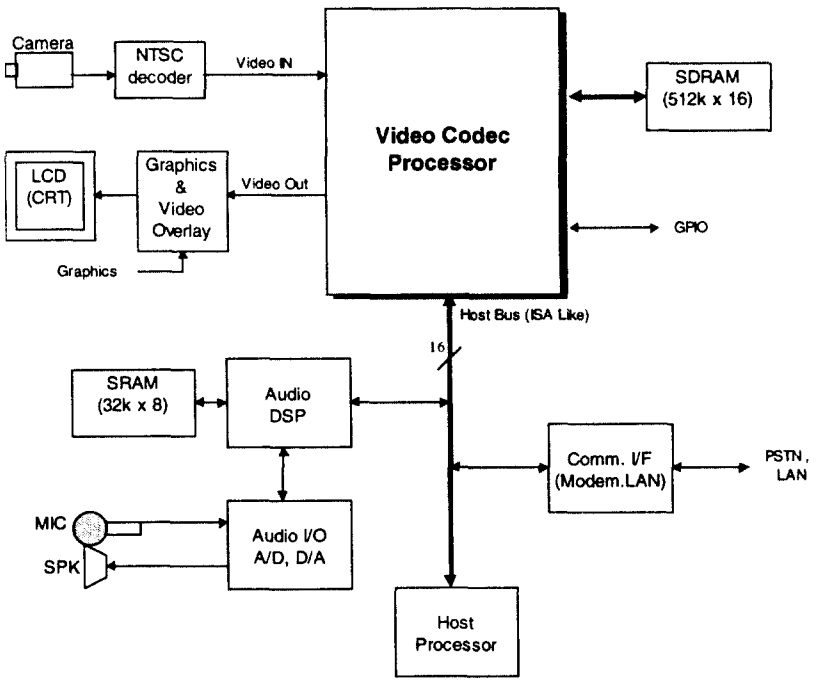
4.4 처리 성능 및 하드웨어 재원

저전력의 시스템 ASIC 기술에 바탕을 두고 설계된 코덱 프로세서로 휴대단말에 효과적으로 사용이 가능하다. 이 프로세서는 H.263 표준을 만족하며 지원해상도는 SQCIF, QCIF이며 각각의 해상도에서 30 frame/sec, 15 frame/sec의 성능을 나타내며 출력 비트율은 16~24Kbps이다. 단일칩으로 구현할 경우 0.5(CMOS 공정에서 집적도는 15만 게이트 수준이며 40MHz의 동작주파수와 3.3V에서 동작한다. FPGA로 구현한 경우 FLEX10K 4개를 사용해야 하고 27MHz, 3.3V에서 동작이 검증되었다.

5. 영상전화 응용시스템의 구조 설계 및 구현

영상전화는 QoS(Quality of Service)가 보장되지 않는 LAN을 이용한 멀티미디어 단말 표준인 H.323 프로토콜을 지원하기 위해 비디오/오디오 코덱의 하드웨어 구현을 바탕으로 호 설정 및 회의 설정 절차를 위한 H.225 및 H.245 프로토콜 블록을 RISC OS 플랫폼상의 C++언어로 구현되었다. 비디오 코덱

〈그림5〉 영상전화 응용시스템 구성 예



프로세서를 중심으로 영상전화 응용시스템을 구축하는 경우의 시스템 구조도는 〈그림5〉와 같다.

H.323 구현을 위한 하드웨어는 2개의 서브-보드(ACodecVIO, CodecBD)로 구분되어 핸디콤비 하드웨어 플랫폼상에 장착된다. 핸디콤비(HandyComBi)는 ETRI, 본 부서에서 개발한 지능화된 사용자 인터페이스를 갖춘 멀티미디어 입출력 전용 단말로 StrongARM SA110, 233MHz CPU를 탑재하고 있다.

ACodecVIO보드는 G.723.1오디오 코덱 처리를 위한 CT8020칩, 마이크, 스피커 입출력을 위한 A/D, D/A칩세트와 비디오 코덱 처리를 위한 비디오 카메라 또는 디스플레이 유니트와 입출력을 위한 NTSC 인코더/디코더 칩들이 탑재되어 있다.

CodecBD는 4개의 FPGA로 구성되어 H.263 표준의 각 기능 블럭들이 분할되어 구현되어 H.263 코덱 기능을 처리한다. 이때 4개의 FPGA 프로그래밍을 위한 8개의 EPROM과 프레임 저장을 위해 SDRAM이 사용되었다.

5.1오디오 코덱부

오디오의 압축/복원은 CT8020[1] 칩을 채택하였고, 오디오의 레코딩과 플레이백을 위한 모드로 구분된다. 레코딩 모드에서 마이크를 통한 오디오 신호는 A/D 변환기를 통해 디지털 신호로 변환되어 CT8020에 입력되고 압축된 데이터는 외부 8K SRAM에 저장된다.

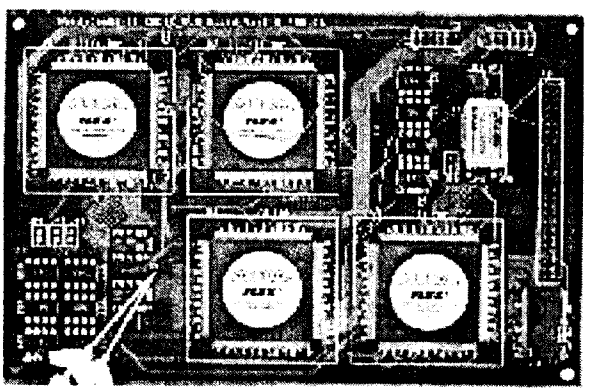
H.225는 이 데이터를 호스트버스를 통해 읽어 비디오 데이터와 멀티플렉싱한다. 플레이백 모드에서는 데이터가 SRAM에 저장된 후 DMA 동작을 통해 CT8020에서 처리되고 D/A 변환기를 거쳐 스피커를 통해 출력된다. CT8020칩은 6.4Kbps의 비트율, 에코 캔슬러 기능을 포함한 상태에서 동작한다.

5.2 비디오 코덱부

자체 설계한 ETRI의 H.263 코덱을 FPGA형태로 구현한 보드(〈그림6〉)를 장착하여 비디오 인코딩과 디코딩을 프레임단위로 순차적으로 수행한다. 비디오

카메라 또는 디스플레이 유니트와 A/D, D/A 변환 칩세트를 통해 직접 인터페이스된다.

〈그림6〉 비디오 코덱보드의 모습

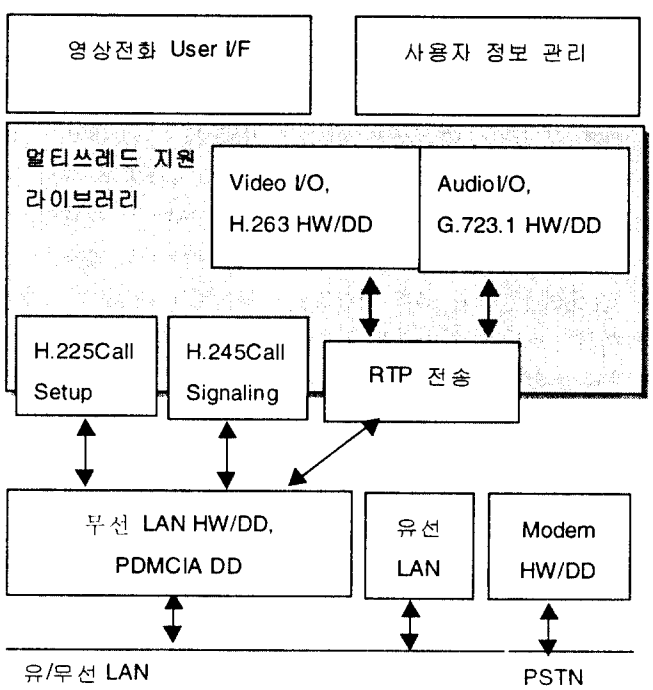


5.3 H.323 소프트웨어 구조

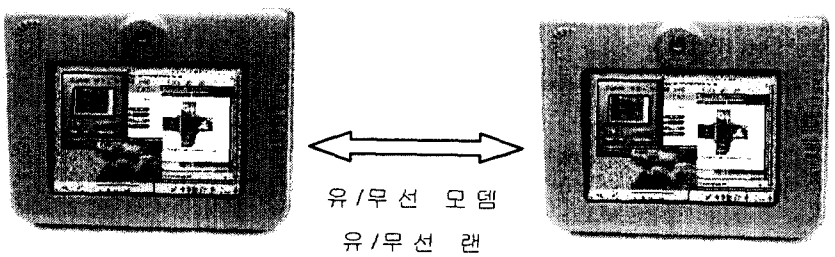
영상전화 응용을 위해서 H.323을 지원하는 각 프로토콜 스택이 구현 되었으며, 오디오 데이터 입출력 및 압축, 복원을 담당하는 오디오 디바이스 드라이버, 비디오 캡쳐 및 압축/복원을 위한 비디오 코덱칩 및 비디오 디바이스 드라이버가 구현되었다. 〈그림7〉은 영상전화 응용 시스템의 구성도를 나타내고 있다. 본 영상전화 응용 시스템은 이동 중인 핸디콤비 사용자에게 무선랜을 이용한 2자간 영상전화를 가능하도록 하며, 다음과 같은 모듈들로 구성된다.

- 영상전화 사용자 인터페이스 및 사용자 관리 모듈
- 영상데이터 출력 모듈
- 멀티쓰레드 지원 라이브러리를 이용한 멀티쓰레드 기능 모듈
- H.225 Call setup 프로토콜 처리 모듈
- H.245 Signalling 프로토콜 처리 모듈
- 실시간 전송 프로토콜(RTP) 프로토콜 처리 모듈
- H.263 비디오 스트림 압축/복원 드라이버 모듈
- 오디오 입출력 루틴 및 G.723.1 압축/복원 드라이버 모듈

〈그림 7〉 영상전화 응용시스템의 기능 모듈 구성도



〈그림 8〉 핸디 콤비에서의 영상전화 모습



기타로 무선랜지원 드라이버 모듈은 상용 PCMCIA 무선랜 카드를 이용하여 핸디콤비 시스템의 운영체제에 맞도록 하드웨어 드라이버 모듈을 구현하였다. PCMCIA용 무선 랜 하드웨어 및 드라이버는 IEEE 802.11 표준이 지원된다.

본 영상전화 응용시스템은 〈그림8〉처럼 다양한 통신 수단, 유/무선 모뎀, 유/무선 랜상에서 동작이 가능하며 유선랜 기반하에서 SQCIF 이미지를 양방향으로 초당 10프레임 처리가 가능하다. 이렇게 낮은 프레임율은 자체 프로세서의 능력보다는 H.323 프로그램내의 S/W 처리로 인한 지연에 기인하는 것으로 분석되었다.

6. 결 론

일반 전화기나 무선 핸드폰 등과 같이 고가의 프로세서를 필요로 하지 않는 단말들을 위한 저전력, 저가의 전용 코덱 칩의 구현은 필수적이다. 이런 추세에 맞추어 ETRI의 본 부서에서 H.263 코덱을 구현하였다. 처리 성능은 SQCIF의 경우 초당 30프레임, QCIF의 경우 초당 15프레임의 처리가 가능하며 출력 비트율은 16~24Kbps이다. H.263 코덱은 FPGA를 사용한 보드형태와 단일칩 형태로 구현을 하고 있으며 보드는 본 부서에서 개발한 핸디콤비 시스템에 탑재되어 기능이 검증되었다. 핸디콤비상에서 H.323을 수행시킨 결과 SQCIF를 초당 10프레임 처리가 가능하였다. 단일칩으로 구현할 경우 0.5(CMOS 공정에서 집적도는 15만 게이트 수준이며 40MHz의 동작주파수와 3.3V에서 동작한다. FPGA로 구현한 경우 FLEX10K 4개를 사용해야 하고 27MHz, 3.3V에서 동작이 검증되었다.

※ 참고문헌

1. ITU-T Draft H.263, "Video Coding for Low Bitrate Communication", July. 5, 1995.

2. ITU-T Rec. H.323, "Visual telephone systems and equipment for local area networks which provide a non guaranteed quality of service", Nov, 1996.

3. ITU-T Rec. H.261, "Video Codec for Audiovisual services at px64 kbit/s", Mar, 1993.

4. ITU-T Rec. H.324, "Line transmission of non-telephone signals - terminal for low bit rate multimedia communication", July, 1995.

5. ITU-T Rec. H.320, "Narrow-band visual telephone systems and terminal equipment", Mar, 1996.

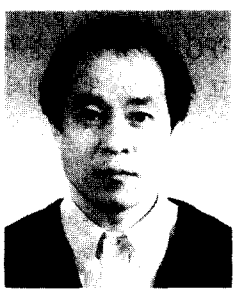
6. Videoconferencing goes to POTS, http://www.ednmag.com/reg/1997/052297/11df_02.cfm

7. H.263 Video Coding, http://rice.ecs.soton.ac.uk/peter/h263/h263.html

8. AT&T Technical Journal, "Screen-Based Multimedia Telephony", Sep/Oct, 1995.

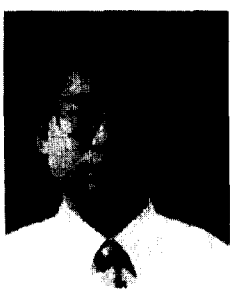
9. ITU-T Draft G.723, "Dual rate speech coder for multimedia communications transmitting at 5.3 & 6.3 Kbit/s", Oct, 1995.

10. ITU-T Rec. G.723.1, "Dual rate speech coder for multimedia communications transmitting at 5.3 & 6.3 Kbit/s", Nov, 1996.



민 병 기

1980 서울대학교 전자공학과 학사
1982 한국과학기술원 전기 및 전자공학과 석사
1991 프랑스 ENST(Telecom. Paris) 공학박사
현재 한국전자통신연구원 컴퓨터 · 소프트웨어기술연구소 멀티미디어연구부 클라이언트구조연구팀 책임연구원
관심분야 : VLSI 구조, ASIC 설계



김 기 현

1989 경북대학교 전자공학과 학사
1991 경북대학교 컴퓨터공학과 석사
현재 한국전자통신연구원 컴퓨터 · 소프트웨어기술연구소 멀티미디어연구부 클라이언트구조연구팀 선임연구원
관심분야 : VLSI 구조, ASIC 설계, 영상처리

최 장 식

1991 전북대학교 컴퓨터공학과 학사
1998 전북대학교 컴퓨터공학과 석사
현재 한국전자통신연구원 컴퓨터 · 소프트웨어기술연구소 멀티미디어연구부 클라이언트구조연구팀 선임연구원
관심분야 : VLSI 구조, ASIC 설계, 영상처리

김 상 범

1985 고려대학교 전자공학과 학사
1987 고려대학교 전자공학과 석사
현재 한국전자통신연구원 컴퓨터 · 소프트웨어기술연구소 멀티미디어연구부 클라이언트구조연구팀 선임연구원
관심분야 : VLSI 구조, ASIC 설계, 영상처리